

UNIVERSIDADE FEDERAL DO PARANÁ

ALISSON ANDREY PUSKA

DISMISS: UMA ABORDAGEM PARA ANÁLISE SOCIOTÉCNICA DA DESINFORMAÇÃO  
DIGITAL

CURITIBA / PR

2023

ALISSON ANDREY PUSKA

DISMISS: UMA ABORDAGEM PARA ANÁLISE SOCIOTÉCNICA DA DESINFORMAÇÃO  
DIGITAL

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Ciência da Computação no Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, da Universidade Federal do Paraná.

Área de concentração: Ciência da Computação.

Orientador: Dr. Roberto Pereira.

CURITIBA / PR

2023

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)  
UNIVERSIDADE FEDERAL DO PARANÁ  
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

Puska, Alisson Andrey

Dismiss: uma abordagem para análise sociotécnica da desinformação digital / Alisson Andrey Puska. – Curitiba, 2023.

1 recurso on-line : PDF.

Tese (Doutorado) - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Informática.

Orientador: Roberto Pereira

1. Desinformação. 2. Semiótica. 3. Comunicação de massa. 4. Desinformação digital. I. Universidade Federal do Paraná. II. Programa de Pós-Graduação em Informática. III. Pereira, Roberto. IV. Título.

Bibliotecário: Elias Barbosa da Silva CRB-9/1894

## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação INFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **ALISSON ANDREY PUSKA** intitulada: **DISMISS: UMA ABORDAGEM PARA ANÁLISE SOCIOTÉCNICA DA DESINFORMAÇÃO DIGITAL**, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 28 de Agosto de 2023.

Assinatura Eletrônica

29/08/2023 09:22:52.0

ROBERTO PEREIRA

Presidente da Banca Examinadora

Assinatura Eletrônica

29/08/2023 09:51:57.0

KATIA ROMERO FELIZARDO SCANNAVINO

Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO  
PARANÁ)

Assinatura Eletrônica

01/09/2023 12:42:43.0

SILVIA AMÉLIA BIM

Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO  
PARANÁ)

Assinatura Eletrônica

30/08/2023 16:57:04.0

ANDREY RICARDO PIMENTEL

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)



## **AGRADECIMENTOS**

Em primeiro lugar, eu gostaria de me agradecer por acreditar em mim mesmo, por todo o meu trabalho duro, por não tirar nenhum dia livre, folga ou férias da minha tese. Eu quero me agradecer por nunca desistir, por sempre ser generoso e tentar dar mais do que eu recebo. Por sempre tentar fazer mais o certo e menos o errado, por sempre ser eu. Por sempre acreditar que o sonho é possível, que os tempos ruins passarão, que é só uma fase. Ainda, quero me agradecer por perseverar no luto de meu pai, de meus tios, e meus amigos na pandemia. Quero agradecer por ter fé em um futuro promissor, mesmo que tudo parecesse desmoronar ao meu redor. Quero agradecer a mim por superar os meus desafios, mesmo com as dificuldades de uma desabilidade cognitiva operacional que foi descoberta no meio do doutorado. Por conseguir enfrentar as dificuldades emocionais e psicológicas que a distância daqueles que amo e a COVID-19 ocasionaram. Pelo amadurecimento que esses anos me trouxeram, pela clareza na visão da realidade, e pelo aprendizado excepcional que essa experiência me deu. Assim, nesse agradecimento, junto as palavras de Snoop Dogg e Edi Rock (Racionais MC's) com as minhas.

Em segundo lugar, quero agradecer a minha família. Ao meu amado pai (Luiz Cícero Puska), que viu seu filho ingressar no doutorado mas não pode ver seu filho receber o título de doutor por conta de um câncer, por todo seu apoio incondicional e por estruturar o meu desenvolvimento. A minha amada mãe (Miriam Lúcia Chaves Puska), que sempre esteve do meu lado, por me segurar enquanto eu caía, por me incentivar enquanto eu estava no chão, e por comemorar comigo quando eu me levantava. A minha amada irmã (Raquel Cristina Puska), que faz sempre muito mais do que ela pode para cuidar da família, por me impulsionar e me inspirar todos os dias.

Em terceiro lugar, quero agradecer a minha segunda família, a todos os meus queridos amigos que sempre buscaram me entender, e que sempre foram um suporte formidável em minhas crises. Além disso, quero agradecer aos meus colegas de laboratório, ao meu orientador, a CAPES e ao Dinf da UFPR pelo auxílio e apoio, pelo amparo dado em suas diferentes formas.

## RESUMO

Essa tese aborda o desafio de entender e lidar com a desinformação digital como um fenômeno sociotécnico, ou seja, que envolve tanto aspectos das tecnologias utilizadas para comunicação quanto do contexto humano/social em que a desinformação ocorre. Os resultados de nosso mapeamento sistemático da literatura mostraram que projetistas de intervenções para mitigação da desinformação têm dificuldades em lidar com a natureza sociotécnica do fenômeno, tendem a utilizar abordagens disciplinares focadas nos aspectos técnicos da desinformação e abordam o fenômeno de forma segmentada. Essas dificuldades podem levar os projetistas à ignorarem aspectos relevantes para o entendimento do fenômeno e à soluções com potenciais prejudiciais, como a censura ou avisos invasivos. Nesse sentido, essa tese investiga meios para apoiar projetistas a compreenderem o fenômeno pela perspectiva sociotécnica, ajudando a caracterizar casos de desinformação digital e auxiliando no entendimento abrangente de problemas. Como solução, essa tese apresenta a Dismiss - uma abordagem para análise Sociotécnica de Desinformações Digitais. A Dismiss é fundamentada na Semiótica Organizacional, composta pelo Modelo Conceitual do Ciclo de Vida da Desinformação Digital, artefatos e materiais de apoio que amparam a análise sociotécnica da desinformação. A abordagem representa uma ferramenta epistêmica projetada para proporcionar a reflexão de seus utilizadores sobre as circunstâncias em que a desinformação ocorre, auxiliando na compreensão da origem e consequências da desinformação digital. A Dismiss foi avaliada de forma construtiva ao longo de seu desenvolvimento, usando métodos de grupo focal (11 encontros), estudos em pequena escala (7 casos), e oficinas de análise sociotécnica de casos de desinformação digital com representantes do público-alvo (3 oficinas). Os resultados dos grupos focais e estudos em pequena escala informaram o refinamento da abordagem, sua estrutura, componentes e métodos de aplicação. Os resultados das oficinas indicam a utilidade percebida da abordagem em apoiar a compreensão da desinformação como um fenômeno sociotécnico. Os resultados também indicaram aspectos que podem ser aprimorados na Dismiss, como a quantidade de passos, a explicação de artefatos, e a densidade dos materiais de apoio, informando melhorias.

Palavras-chave: Desinformação, semiótica organizacional, fenômeno sociotécnico, ferramenta epistêmica, ciclo de vida da desinformação, método de articulação de problemas.

## ABSTRACT

This thesis addresses the challenge of understanding and dealing with digital misinformation as a sociotechnical phenomenon, meaning that it involves both aspects of the technologies used for communication and the human/social context in which misinformation occurs. The results of our systematic literature review showed that designers of interventions for mitigating misinformation face difficulties in dealing with the sociotechnical nature of the phenomenon. They tend to employ disciplinary approaches focused on the technical aspects of misinformation and often address the phenomenon in a fragmented manner. These difficulties can lead designers to overlook relevant aspects for understanding the phenomenon and result in potentially harmful solutions, such as censorship or invasive warnings. In this regard, this thesis investigates means to support designers in comprehending the phenomenon from a sociotechnical perspective, helping to characterize cases of digital misinformation and aiding in a comprehensive understanding of the issues. As a solution, this thesis presents Dismiss - an Approach for Sociotechnical Analysis of Digital Misinformation. Dismiss is grounded in Organizational Semiotics, comprised of the Conceptual Model of the Digital Misinformation Lifecycle, artifacts, and supporting materials that underpin the sociotechnical analysis of misinformation. The approach serves as an epistemic tool designed to facilitate users' reflection on the circumstances in which misinformation occurs, assisting in understanding the origins and consequences of digital misinformation. Dismiss was constructively evaluated throughout its development, utilizing focus group methods (11 meetings), small-scale studies (7 cases), and workshops for the sociotechnical analysis of digital misinformation cases with representatives of the target audience (3 workshops). The results from the focus groups and small-scale studies informed the refinement of the approach, its structure, components, and application methods. The workshop results indicate the perceived utility of the approach in supporting the understanding of misinformation as a sociotechnical phenomenon. The results also highlighted aspects that can be improved in Dismiss, such as the number of steps, artifact explanations, and the density of supporting materials, providing insights for enhancements.

**Keywords:** Misinformation, organizational semiotics, sociotechnical phenomenon, epistemic tool, disinformation life cycle, problem articulation method.

## LISTA DE FIGURAS

|      |                                                                                                 |    |
|------|-------------------------------------------------------------------------------------------------|----|
| 1.1  | Visão geral dos materiais e métodos da pesquisa. . . . .                                        | 18 |
| 2.1  | Uma representação de desinformação accidental. Fonte: Autor. . . . .                            | 22 |
| 2.2  | Um caso de estelionato (desinformação deliberada). Fonte: G1 (G1, 2023a).. . .                  | 23 |
| 2.3  | Escada Semiótica Adaptada. Inspirada: Ferreira (2015).. . . . .                                 | 26 |
| 2.4  | Diagrama de Partes Interessadas. Traduzido de Stamper (1994) . . . . .                          | 27 |
| 2.5  | Exemplo de enquadramento. . . . .                                                               | 30 |
| 3.1  | Linhas de pesquisa sobre desinformação na Computação. . . . .                                   | 36 |
| 3.2  | Tipos de desinformação encontrados no mapeamento.. . . . .                                      | 43 |
| 3.3  | Mapa de partes interessadas encontradas no mapeamento.. . . . .                                 | 47 |
| 3.4  | Esquema de categorização dos artigos.. . . . .                                                  | 49 |
| 3.5  | Aspectos sociotécnicos encontrados na literatura. . . . .                                       | 50 |
| 4.1  | Componentes da Dismiss. Fonte: Autor. . . . .                                                   | 59 |
| 4.2  | Etapas da Dismiss. Fonte: Autor.. . . . .                                                       | 60 |
| 4.3  | Catálogo dos tipos de desinformação. Fonte: Autor.. . . . .                                     | 62 |
| 4.4  | Diagrama de Partes Interessadas. Adaptado de (Kolkman, 1993). . . . .                           | 64 |
| 4.5  | Adaptação da Escada Semiótica. Fonte: Inspirado em (Guimarães et al., 2008). .                  | 66 |
| 4.6  | Estrutura do ciclo da desinformação digital. . . . .                                            | 70 |
| 4.7  | Exemplo de um caso de desinformação pela perspectiva do ciclo de vida da desinformação. . . . . | 71 |
| 4.8  | Representação do objeto desinformativo original. . . . .                                        | 72 |
| 4.9  | Identificação de Objetos Desinformativos. . . . .                                               | 73 |
| 4.10 | Identificação de Canais de Comunicação. . . . .                                                 | 73 |
| 4.11 | Identificação de Partes Interessadas. . . . .                                                   | 74 |
| 4.12 | Identificação de Partes Interessadas. . . . .                                                   | 74 |
| 4.13 | Identificação de Partes Interessadas. . . . .                                                   | 75 |
| 4.14 | Identificação de Partes Interessadas. . . . .                                                   | 76 |
| 4.15 | Objeto desinformativo para exemplificação da Estratégia 2. Fonte: WhatsApp do autor. . . . .    | 78 |
| 4.16 | Exploração do contexto com base no objeto. . . . .                                              | 79 |
| 4.17 | Exploração do contexto com base no ciclo de vida - Busca no Instagram.. . . .                   | 79 |
| 4.18 | Exploração do contexto com base no ciclo de vida - Busca no Site Agora Notícia.                 | 80 |
| 4.19 | Exploração do contexto com base no ciclo de vida - Busca no Site Agora Notícia.                 | 80 |

|      |                                                                                                                                           |     |
|------|-------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4.20 | Exploração do contexto com base no ciclo de vida - Busca no Site Antena Política.                                                         | 81  |
| 4.21 | Exploração do contexto com base no ciclo de vida - Busca no Google. . . . .                                                               | 81  |
| 4.22 | Exploração do conteúdo do objeto para identificar Partes Interessadas. . . . .                                                            | 82  |
| 4.23 | Possíveis fontes para observar partes interessadas no FRC preenchido. . . . .                                                             | 83  |
| 4.24 | Exploração do contexto para identificar Partes Interessadas- Ferramentas úteis. .                                                         | 83  |
| 4.25 | Partes Interessadas mapeadas no DPI. . . . .                                                                                              | 84  |
| 4.26 | Representação da Análise sociotécnica do Ciclo de Vida da Desinformação usando a Escada Semiótica. . . . .                                | 85  |
| 4.27 | Análise sociotécnica pela perspectiva da criação da desinformação - Sistema de Informação Humano. . . . .                                 | 86  |
| 4.28 | Análise sociotécnica pela perspectiva da criação da desinformação - Plataforma Tecnológica. . . . .                                       | 86  |
| 4.29 | Análise sociotécnica pela perspectiva da difusão da desinformação - Sistema de Informação Humana. . . . .                                 | 88  |
| 4.30 | Análise sociotécnica pela perspectiva da difusão da desinformação - Sistema de Informação Humano. . . . .                                 | 88  |
| 4.31 | Análise sociotécnica pela perspectiva do consumo da desinformação - Sistema de Informação Humano. . . . .                                 | 89  |
| 4.32 | Análise sociotécnica pela perspectiva do consumo da desinformação - Sistema de Informação Humana. . . . .                                 | 89  |
| 5.1  | Evolução da representação visual do ciclo de vida da desinformação. . . . .                                                               | 97  |
| 5.2  | Coleção de objetos desinformativos analisados pelo autor. . . . .                                                                         | 98  |
| 5.3  | Representação do objeto desinformativo (phishing). . . . .                                                                                | 102 |
| 5.4  | Partes interessadas (RSA Phishing). . . . .                                                                                               | 102 |
| 5.5  | Propriedades semióticas no objeto desinformativo. . . . .                                                                                 | 103 |
| 5.6  | Representação do objeto desinformativo original. . . . .                                                                                  | 106 |
| 5.7  | Variantes do objeto desinformativo (Farsa WhatsApp). . . . .                                                                              | 108 |
| 5.8  | Transformações do objeto desinformativo. . . . .                                                                                          | 109 |
| 5.9  | Partes da nota informativa Nº 9/2020-SE/GAB/SE/MS. . . . .                                                                                | 110 |
| 5.10 | Exemplos de falácias lógicas: apelo à autoridade e à ignorância. Fonte: Material Oficial (Pág. 4). . . . .                                | 111 |
| 5.11 | Recomendações do CFM - Fonte: Material Oficial (Pág. 5) . . . . .                                                                         | 112 |
| 5.12 | Desinformação em APPs de Saúde.. . . .                                                                                                    | 112 |
| 5.13 | Área de discussões sobre notícias. Fonte: Gazeta do Povo (do Povo, 2021). . . .                                                           | 113 |
| 5.14 | Postagens em redes sociais (Março/Agosto de 2020). Fonte: AosFatos (Fatos, 2020). . . . .                                                 | 114 |
| 5.15 | Matéria da BBC sobre comunidade de profissionais da saúde pró-tratamento preventivo (Setembro de 2020). Fonte: BBC (Alvim, 2020). . . . . | 114 |
| 5.16 | Site sobre a temática de tratamento preventivo. . . . .                                                                                   | 115 |

|      |                                                                                                      |     |
|------|------------------------------------------------------------------------------------------------------|-----|
| 5.17 | Outdoors em Curitiba. Fonte: Gazeta do Povo (Do Povo, 2021). . . . .                                 | 115 |
| 5.18 | Desinformação para análise - Oficina I. . . . .                                                      | 118 |
| 5.19 | Sobre a utilidade percebida da Dismiss. . . . .                                                      | 119 |
| 5.20 | Sobre a facilidade de uso dos formulários de registro de dados da Dismiss. . . . .                   | 120 |
| 5.21 | Formulário que mapeia o ASE - Preenchido por P4. . . . .                                             | 121 |
| 5.22 | Sobre a complexidade conceitual e teórica da Dismiss. . . . .                                        | 122 |
| 5.23 | Formulário que mapeia o ASE - Preenchido por P2. . . . .                                             | 122 |
| 5.24 | Sobre utilidade e facilidade de uso do Modelo Conceitual do Ciclo de Vida da Desinformação. . . . .  | 123 |
| 5.25 | Sobre utilidade dos materiais de apoio. . . . .                                                      | 124 |
| 5.26 | Sobre utilidade e facilidade de uso do manual de aplicação da Dismiss. . . . .                       | 125 |
| 5.27 | Isca-clique no Youtube - Caso 1. . . . .                                                             | 127 |
| 5.28 | Farsa no WhatsApp (Oferta de emprego) - Caso 2. . . . .                                              | 127 |
| 5.29 | Sobre a utilidade percebida da Dismiss. . . . .                                                      | 128 |
| 5.30 | Sobre a facilidade de uso da Dismiss em geral. . . . .                                               | 129 |
| 5.31 | Sobre a facilidade de uso dos materiais da Dismiss. . . . .                                          | 130 |
| 5.32 | Sobre a facilidade de uso dos materiais da Dismiss. . . . .                                          | 131 |
| 5.33 | Desinformação para análise - Oficina III. Fonte: Twitter. . . . .                                    | 133 |
| 5.34 | Sobre a facilidade de uso dos materiais da Dismiss. . . . .                                          | 134 |
| 5.35 | Diagrama de Identificação de Partes Interessadas Preenchido. . . . .                                 | 135 |
| 5.36 | Escada Semiótica preenchida para o estágio de Consumo do Ciclo de Vida. . . . .                      | 136 |
| 5.37 | Quadro de sugestão de problemas preenchido. . . . .                                                  | 136 |
| 5.38 | Sobre a facilidade de uso dos materiais da Dismiss. . . . .                                          | 137 |
| A.1  | Relação da tríade semiótica. . . . .                                                                 | 181 |
| A.2  | Diferentes enquadramentos na comunicação que levam ao desentendimento (acidental). . . . .           | 190 |
| A.3  | Diferentes enquadramentos na comunicação que levam ao desentendimento (deliberado). . . . .          | 191 |
| A.4  | Representação da Cebola Organizacional. . . . .                                                      | 192 |
| A.5  | Principais critérios para categorização de desinformação encontrados na literatura revisada. . . . . | 216 |
| A.6  | Temas abordados pelos estudos. . . . .                                                               | 221 |
| A.7  | Aspectos de um caso de desinformação (evento) na DO. . . . .                                         | 254 |
| A.8  | Aspectos do estágio de criação na CO. . . . .                                                        | 255 |
| A.9  | Aspectos do estágio de criação na CO. . . . .                                                        | 255 |
| A.10 | Aspectos da difusão na CO. . . . .                                                                   | 257 |
| A.11 | Aspectos do estágio de Difusão na DO. . . . .                                                        | 257 |
| A.12 | Aspectos do consumo na Cebola Semiótica. . . . .                                                     | 260 |

|      |                                                                                |     |
|------|--------------------------------------------------------------------------------|-----|
| A.13 | Aspectos do estágio de criação na CO. . . . .                                  | 260 |
| A.14 | Aspectos do estágio de Reiteração na DO. . . . .                               | 262 |
| A.15 | Aspectos de um caso de desinformação (evento) na DO.. . . .                    | 263 |
| A.16 | Representação do objeto desinformativo (phishing). . . . .                     | 264 |
| A.17 | Partes interessadas (RSA Phishing). . . . .                                    | 265 |
| A.18 | Aspectos do campo de informações relevantes para o evento desinformativo.. . . | 266 |
| A.19 | Aspectos do campo de informações relevantes para o evento desinformativo.. . . | 266 |
| A.20 | Exemplo de Unidade Semânticas - Caso RSA. . . . .                              | 267 |
| A.21 | Exemplo de Diagrama Ontológico de atividade Spear Phishing. - Caso RSA. . .    | 267 |
| A.22 | Propriedades semióticas no objeto desinformativo. . . . .                      | 268 |
| A.23 | Representação do objeto desinformativo (Farsa WhatsApp).. . . .                | 272 |
| A.24 | Aspectos do campo de informações relevantes para o evento desinformativo.. . . | 273 |
| A.25 | Partes interessadas (Covid e pH) . . . . .                                     | 273 |
| A.26 | Análise de Farsa no WhatsApp.. . . .                                           | 274 |
| A.27 | Variantes do objeto desinformativo (Farsa WhatsApp). . . . .                   | 277 |
| A.28 | Ciclo de vida da desinformação - Reiteração. . . . .                           | 278 |



## LISTA DE TABELAS

|      |                                                                                            |     |
|------|--------------------------------------------------------------------------------------------|-----|
| 2.1  | Conceitos dos fundamentos teóricos . . . . .                                               | 34  |
| 2.2  | Conceitos dos fundamentos teóricos . . . . .                                               | 35  |
| 3.1  | Desafios encontrados na literatura . . . . .                                               | 38  |
| 3.2  | Chave-base de pesquisa . . . . .                                                           | 41  |
| 3.3  | Estatísticas do mapeamento sistemático. . . . .                                            | 42  |
| 3.4  | Critérios de categorização encontrados na literatura.. . . .                               | 43  |
| 3.5  | Problemas de concisão no entendimento dos tipos de desinformação . . . . .                 | 44  |
| 3.6  | Tipos comuns de desinformação . . . . .                                                    | 46  |
| 3.7  | Exemplos de características dos estágios do fenômeno da desinformação. . . . .             | 55  |
| 3.8  | Comparação entre diferentes abordagens sociotécnicas do fenômeno da desinformação. . . . . | 57  |
| 4.1  | Formulário de apoio: registro de caso de desinformação. Fonte: Autor.. . . .               | 61  |
| 4.2  | Formulário de apoio: Registro de Partes Interessadas - Adaptação do DPI. . . .             | 65  |
| 4.3  | Exemplos de aspectos relevantes em casos de desinformação.. . . .                          | 67  |
| 4.4  | Quadro de Aspectos do Meio de Comunicação. Fonte: Autor.. . . .                            | 68  |
| 4.5  | Quadro de Princípios de Influência. Fonte: Autor. . . . .                                  | 68  |
| 4.6  | Formulário de apoio: registro de caso de desinformação. Fonte: Autor.. . . .               | 74  |
| 4.7  | Identificação de Partes Interessadas. . . . .                                              | 75  |
| 4.8  | Formulário de Registro de Aspectos Sociotécnicos (Análise do Objeto) - Preenchido. . . . . | 76  |
| 4.9  | Formulário de Registro de Caso - Preenchido. . . . .                                       | 82  |
| 4.10 | Fomulário de Registro de Partes Interessadas. . . . .                                      | 84  |
| 4.11 | Exploração do contexto com base no ciclo de vida - Busca em outros canais.. . .            | 90  |
| 4.12 | Formulário de Análise Sociotécnica preenchido.. . . .                                      | 91  |
| 4.13 | Requisitos para abordar a desinformação digital como um fenômeno sociotécnico. . . . .     | 92  |
| 5.1  | Detalhamento dos encontros com o grupo focal. . . . .                                      | 94  |
| 5.2  | Resumo dos Estudos em Pequena Escala . . . . .                                             | 99  |
| 5.3  | Resumo dos aspectos sociotécnicos identificados na análise. . . . .                        | 104 |
| 5.4  | Resumo das Oficinas . . . . .                                                              | 117 |
| A.1  | Catálogo de Falácias Lógicas . . . . .                                                     | 195 |
| A.2  | Catálogo de vieses cognitivos. . . . .                                                     | 196 |
| A.3  | Estratégias de mitigação de desinformação digital.. . . .                                  | 206 |

|      |                                                                            |     |
|------|----------------------------------------------------------------------------|-----|
| A.4  | Questões de pesquisa . . . . .                                             | 208 |
| A.5  | Chave-base de pesquisa . . . . .                                           | 208 |
| A.6  | Critérios para seleção e exclusão dos artigos . . . . .                    | 209 |
| A.7  | Formulário de extração de dados . . . . .                                  | 210 |
| A.8  | Estatísticas do mapeamento sistemático . . . . .                           | 211 |
| A.9  | Revisões Sistemáticas sobre desinformação . . . . .                        | 212 |
| A.10 | Critérios de categorização encontrados na literatura.. . . .               | 216 |
| A.11 | Descrição das categorias de desinformação . . . . .                        | 220 |
| A.12 | Quantidade de artigos por tipo de desinformação . . . . .                  | 220 |
| A.13 | Problemas de concisão no entendimento dos tipos de desinformação . . . . . | 221 |
| A.14 | Tipos de desinformação . . . . .                                           | 226 |
| A.15 | Distribuição dos estudos por sub-disciplina . . . . .                      | 227 |
| A.16 | Classificação da dimensão Mundo Social. . . . .                            | 227 |
| A.17 | Classificação da dimensão Pragmática . . . . .                             | 228 |
| A.18 | Classificação da dimensão Semântica . . . . .                              | 230 |
| A.19 | Classificação da dimensão Sintática . . . . .                              | 230 |
| A.20 | Classificação das dimensões Empírica e Mundo Físico . . . . .              | 233 |
| A.21 | Tipos de interações, tempo e direção da comunicação . . . . .              | 259 |
| A.22 | Resumo dos aspectos sociotécnicos identificados na análise . . . . .       | 269 |
| A.23 | Resumo dos aspectos sociotécnicos identificados na análise . . . . .       | 275 |

## LISTA DE ACRÔNIMOS

|      |                                                  |
|------|--------------------------------------------------|
| AE   | Análise de Enquadramento                         |
| CO   | Cebola Organizacional                            |
| DPI  | Diagrama de Partes Interessadas                  |
| ES   | Escada Semiótica                                 |
| ESE  | Escada Semiótica Estendida                       |
| FRAS | Formulário de registro de aspectos sociotécnicos |
| FRED | Formulário de registro de caso desinformativo    |
| FRPI | Formulário de registro de partes interessadas    |
| GQM  | <i>Goal-Question-Metric</i>                      |
| OMS  | Organização Mundial de Saúde                     |
| QP   | Questões de Pesquisa                             |
| SI   | Sistema de Informação                            |
| SO   | Semiótica Organizacional                         |
| TIC  | Tecnologias de informação e comunicação          |

## SUMÁRIO

|          |                                                                                             |           |
|----------|---------------------------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>INTRODUÇÃO . . . . .</b>                                                                 | <b>16</b> |
| 1.1      | PROBLEMA DE PESQUISA E OBJETIVOS . . . . .                                                  | 17        |
| 1.2      | MATERIAIS E MÉTODOS . . . . .                                                               | 18        |
| 1.3      | CONTRIBUIÇÕES . . . . .                                                                     | 19        |
| 1.4      | ESTRUTURA DA TESE . . . . .                                                                 | 20        |
| <b>2</b> | <b>FUNDAMENTAÇÃO TEÓRICA. . . . .</b>                                                       | <b>21</b> |
| 2.1      | A DESINFORMAÇÃO DIGITAL COMO UM FENÔMENO SOCIOTÉCNICO. . . . .                              | 21        |
| 2.1.1    | Os elementos de um caso de desinformação . . . . .                                          | 23        |
| 2.2      | SEMIÓTICA ORGANIZACIONAL E OS ASPECTOS SOCIOTÉCNICOS . . . . .                              | 25        |
| 2.2.1    | A Escada Semiótica . . . . .                                                                | 25        |
| 2.2.2    | As Partes Interessadas . . . . .                                                            | 26        |
| 2.2.3    | Um Entendimento para Desinformação Inspirado na Semiótica Organizacional. . . . .           | 27        |
| 2.3      | ANÁLISE DE ENQUADRAMENTO E O FENÔMENO DA DESINFORMAÇÃO . . . . .                            | 29        |
| 2.4      | GLOSSÁRIO . . . . .                                                                         | 34        |
| 2.5      | EXPLICAÇÕES E DESDOBRAMENTOS TEÓRICOS. . . . .                                              | 35        |
| <b>3</b> | <b>REVISÃO DA LITERATURA . . . . .</b>                                                      | <b>36</b> |
| 3.1      | ESTUDO EXPLORATÓRIO . . . . .                                                               | 36        |
| 3.2      | MAPEAMENTO SISTEMÁTICO DA LITERATURA . . . . .                                              | 40        |
| 3.2.1    | QP01 - Quais são os tipos de desinformação estudados? . . . . .                             | 42        |
| 3.2.2    | QP02 - Quais as partes interessadas consideradas pelas pesquisas? . . . . .                 | 46        |
| 3.2.3    | QP03 - Como as pesquisas abordam aspectos sociotécnicos do fenômeno? . . . . .              | 49        |
| 3.2.4    | QP04 - Como as pesquisas abordam a estrutura de um caso de desinformação digital? . . . . . | 53        |
| 3.3      | AMEAÇAS À VALIDADE DO MAPEAMENTO . . . . .                                                  | 56        |
| 3.4      | CONSIDERAÇÕES FINAIS . . . . .                                                              | 56        |
| <b>4</b> | <b>DISMISS - UMA ABORDAGEM PARA ANÁLISE SOCIOTÉCNICA DA DESINFORMAÇÃO. . . . .</b>          | <b>58</b> |
| 4.1      | VISÃO GERAL . . . . .                                                                       | 58        |
| 4.1.1    | Estrutura da Dismiss . . . . .                                                              | 59        |
| 4.1.2    | Etapas 1 - Familiarização . . . . .                                                         | 61        |
| 4.1.3    | Etapas 2 - Análise sociotécnica do caso . . . . .                                           | 65        |
| 4.1.4    | Etapas 3 - Identificação, Análise e Entendimento de Problemas . . . . .                     | 67        |

|          |                                                                                                |            |
|----------|------------------------------------------------------------------------------------------------|------------|
| 4.2      | UM MODELO CONCEITUAL DO CICLO DE VIDA DA DESINFORMAÇÃO                                         | 69         |
| 4.3      | ESTRATEGIAS DE APLICAÇÃO . . . . .                                                             | 71         |
| 4.3.1    | Análise sociotécnica de objetos desinformativos (Profundidade) . . . . .                       | 72         |
| 4.3.2    | Análise sociotécnica do Ciclo de Vida da Desinformação . . . . .                               | 77         |
| 4.4      | CONSIDERAÇÕES SOBRE A DISMISS . . . . .                                                        | 92         |
| <b>5</b> | <b>AVALIAÇÃO E RESULTADOS ALCANÇADOS. . . . .</b>                                              | <b>93</b>  |
| 5.1      | AVALIAÇÃO QUALITATIVA POR GRUPO FOCAL. . . . .                                                 | 93         |
| 5.1.1    | Resultados e discussões . . . . .                                                              | 94         |
| 5.2      | ESTUDOS EM PEQUENA ESCALA . . . . .                                                            | 98         |
| 5.2.1    | Análise do esquema de <i>phishing</i> - Farsa na RSA . . . . .                                 | 101        |
| 5.2.2    | Análise da mensagem - Farsa no WhatsApp . . . . .                                              | 105        |
| 5.3      | OFICINAS DE ANÁLISE SOCIOTÉCNICA DE DESINFORMAÇÃO . . . . .                                    | 117        |
| 5.3.1    | Oficina I. . . . .                                                                             | 118        |
| 5.3.2    | Oficina II . . . . .                                                                           | 126        |
| 5.3.3    | Oficina III . . . . .                                                                          | 132        |
| 5.3.4    | Discussão sobre as oficinas . . . . .                                                          | 138        |
| 5.4      | CONSIDERAÇÕES SOBRE AS AVALIAÇÕES . . . . .                                                    | 139        |
| 5.4.1    | Nota sobre os estudos . . . . .                                                                | 141        |
| <b>6</b> | <b>CONCLUSÃO . . . . .</b>                                                                     | <b>142</b> |
| 6.1      | TRABALHOS FUTUROS . . . . .                                                                    | 144        |
|          | <b>REFERÊNCIAS . . . . .</b>                                                                   | <b>146</b> |
|          | <b>APÊNDICE A – APÊNDICES . . . . .</b>                                                        | <b>179</b> |
| A.1      | APÊNDICE 1 - SOBRE AS DECISÕES TEÓRICAS PARA ELABORAÇÃO DA DISMISS . . . . .                   | 179        |
| A.1.1    | As bases teóricas da Semiótica Organizacional. . . . .                                         | 179        |
| A.1.2    | Sobre a veracidade e as implicações . . . . .                                                  | 183        |
| A.1.3    | Modelos de Comunicação, Semiótica Organizacional e os Desafios da Análise Pragmática . . . . . | 185        |
| A.1.4    | A Análise de Enquadramento. . . . .                                                            | 189        |
| A.1.5    | A Cebola Organizacional . . . . .                                                              | 191        |
| A.1.6    | Eixos complementares que sustentam a análise sociotécnica de desinformação . . . . .           | 193        |
| A.2      | APÊNDICE 2 - DESDOBRAMENTOS . . . . .                                                          | 201        |
| A.2.1    | Vulnerabilidades Humanas no Fenômeno da Desinformação Digital . . . . .                        | 201        |
| A.2.2    | Uma classificação para estratégias de mitigação da desinformação digital. . . . .              | 205        |
| A.3      | APÊNDICE 3 - PROTOCOLO DO MAPEAMENTO SISTEMÁTICO . . . . .                                     | 208        |
| A.3.1    | Resultados e Discussões . . . . .                                                              | 211        |
| A.3.2    | Relatório: Mapeamentos e Revisões sistemáticas da literatura. . . . .                          | 211        |

|       |                                                                                                             |     |
|-------|-------------------------------------------------------------------------------------------------------------|-----|
| A.3.3 | Relatório: QP01.a: O que a literatura define como <i>misinformation</i> e <i>disinformation</i> ? . . . . . | 215 |
| A.3.4 | Relatório: QP01.b: Quais os tipos de desinformação estudados e os temas? . . .                              | 220 |
| A.3.5 | Relatório: QP03 - Como as pesquisas abordam os aspectos sociotécnicos? . . .                                | 226 |
| A.3.6 | Estudos Relacionados . . . . .                                                                              | 237 |
| A.4   | APÊNDICE 4 - DISSMISS FALSE INFORMATION: A VALUE MATTER . . .                                               | 239 |
| A.5   | APÊNDICE 5 - WHATSAPP AND FALSE INFORMATION: A VALUE-ORIENTED EVALUATION . . . . .                          | 243 |
| A.6   | APÊNDICE 6 - DESENVOLVIMENTO DO MODELO CONCEITUAL DO CICLO DE VIDA DA DESINFORMAÇÃO DIGITAL . . . . .       | 254 |
| A.7   | APÊNDICE 7 - RELATÓRIOS DAS AVALIAÇÕES . . . . .                                                            | 264 |
| A.7.1 | Sobre os estudo em pequena escala . . . . .                                                                 | 264 |
| A.7.2 | Análise de uma farsa no WhatsApp. . . . .                                                                   | 272 |
| A.7.3 | Análise Semiótica do Ciclo de Vida . . . . .                                                                | 273 |
| A.8   | APÊNDICE 8 - PROTOCOLO DAS OFICINAS I E II . . . . .                                                        | 281 |
| A.8.1 | O estudo - ( 120 min) . . . . .                                                                             | 281 |
| A.8.2 | Questionários . . . . .                                                                                     | 282 |
| A.9   | APÊNDICE 9 - PROTOCOLO DA OFICINA III . . . . .                                                             | 284 |
| A.9.1 | O estudo - ( 120 min) . . . . .                                                                             | 284 |
| A.9.2 | Introdução. . . . .                                                                                         | 284 |
| A.9.3 | Fase 1: Introdução à Dismiss (20 minutos). . . . .                                                          | 284 |
| A.9.4 | Fase 2: O estudo de caso (60 minutos) . . . . .                                                             | 285 |
| A.9.5 | Fase 3: Avaliação de Aceitação de Tecnologia (40 min) . . . . .                                             | 285 |
| A.10  | APÊNDICE 10 - TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE) . . . . .                                   | 287 |
| A.11  | APÊNDICE 11 - PLANO DE AVALIAÇÃO DA DISMISS . . . . .                                                       | 289 |
| A.12  | APÊNDICE 13 - MANUAL DA DISMISS. . . . .                                                                    | 294 |

## 1 INTRODUÇÃO

A desinformação digital é um fenômeno emergente nas Tecnologias Digitais de Informação e Comunicação (TIC) (da Silva Gomes e Dourado, 2019). Esse fenômeno ocorre quando há problemas na expressão e interpretação de informações em mídias digitais, podendo ser acidental ou deliberado. A desinformação acidental ocorre sem a intenção de enganar e pode ser causada por erros de memória, desatenção, ignorância, corrupção de dados, defeitos físicos (tela quebrada), entre outros (Tudjman e Mikelic, 2003). A desinformação deliberada envolve a intenção de enganar, como notícias falsas, estelionatos, difamação, *hacking* ou vírus (Tudjman e Mikelic, 2003).

Independentemente do tipo (deliberada ou acidental), a desinformação digital envolve aspectos sociais, como cultura, valores, crenças, leis e comportamentos, e aspectos técnicos, como procedimentos, protocolos e tecnologias de comunicação (Puska et al., 2019). Como resultado, diferentes grupos sociais em diferentes locais podem ter suas próprias formas de comunicação. A interação humana com (e através) da tecnologia de comunicação pode ser influenciada por aspectos como valores humanos (Cho et al., 2016), condições sociais e econômicas (Kshetri e Voas, 2017), nível educacional e traços culturais (Chandra et al., 2017), o que afeta a maneira como as pessoas usam as TICs para se comunicar, produzir, compartilhar e consumir informações.

Marshall McLuhan, filósofo canadense, afirma que a tecnologia como meio de comunicação é tão importante quanto o conteúdo da mensagem transmitida por ela (McLuhan, 1974). Ele argumenta que a tecnologia molda a forma como as pessoas se expressam e se associam. Dessa forma, a tecnologia pode restringir ou aprimorar a capacidade de expressão e interpretação humanas e controlar o alcance da comunicação. No fenômeno da desinformação digital, a tecnologia pode facilitar a difusão de desinformação, dificultar a verificação de mensagens e favorecer a transformação de conteúdos, atrapalhando sua detecção (Puska et al., 2020). Portanto, entender o fenômeno da desinformação digital requer uma compreensão tanto dos aspectos socioculturais quanto técnicos (Puska et al., 2019). Para lidar com essa complexidade, esta tese defende abordar a desinformação digital como um fenômeno sociotécnico. Por fenômeno sociotécnico, a tese se refere ao projeto de sistemas técnicos, as capacidades de processamento humano, as condições socioeconômicas, a legislação, os valores humanos e os aspectos culturais, que devem ser compreendidos e considerados interligados ao abordar a desinformação digital (Cafezeiro et al., 2017).

Para compreender como a literatura aborda a natureza sociotécnica do fenômeno da desinformação digital foi realizado um mapeamento sistemático da literatura em Ciência da Computação. O mapeamento revelou diversas abordagens para o fenômeno, desde trabalhos sobre detecção e bloqueio automático de desinformação (Bondielli e Marcelloni, 2019), à estratégias de correção e educação das pessoas (Soler e Busilo, 2019). Os resultados da análise destes trabalhos mostram três aspectos desafiadores do fenômeno e as limitações ao abordar a natureza sociotécnica da desinformação digital:

1. Aspectos sociotécnicos fragmentados: o primeiro aspecto desafiador da desinformação digital é a sua natureza sociotécnica, que envolve aspectos técnicos, humanos, sociais e organizacionais de forma inter-relacionada (Puska et al., 2019). No entanto, muitas abordagens disciplinares têm focado apenas nos aspectos técnicos (Fernandez e Alani, 2018), enquanto outras pesquisas consideram apenas conjuntos pré-determinados de aspectos sociotécnicos (Jindal et al., 2020), o que limita o entendimento das



características relevantes e de como suas inter-relações impactam a interação situada com a desinformação digital em diferentes contextos de comunicação.

2. Diversidade de partes interessadas: o segundo aspecto desafiador é que o fenômeno da desinformação envolve diversas partes interessadas, realizando diferentes tarefas, em diferentes contextos e com diferentes tecnologias (Starbird et al., 2018). Entretanto, as soluções tendem a abordar as partes interessadas de modo genérico, levando a uma visão superficial da relação entre elas, as tecnologias e a desinformação digital.
3. Estrutura de casos segmentada: O terceiro aspecto desafiador é que a estrutura de um caso de desinformação consiste de vários estágios inter-relacionados (Shin et al., 2018), mas muitas análises os consideram de modo segmentado ou por modelos que restringem a análise de aspectos sociotécnicos inter-relacionados (Fernandez e Alani, 2018), ocultando aspectos relevantes em consideração a uma visão ampla da criação, difusão e consumo de desinformação.

Embora avanços significativos tenham sido feitos, estes desafios representam limitações para o entendimento do fenômeno que levam a estratégias de mitigação pouco adaptáveis a outros contextos (Fernandez e Alani, 2018), e a soluções com o potencial prejudicial à interação das pessoas através de TIC's, como a censura e avisos invasivos (Nekmat, 2020; Wang et al., 2014).

O fenômeno da desinformação engloba a diversidade de aspectos sociotécnicos, contextos e partes interessadas, diferentes estágios inter-relacionados, e as diferentes tecnologias e meios de comunicação. Essa complexidade dificulta a identificação de fatores essenciais para compreensão das circunstâncias em que a desinformação digital ocorre, como, por exemplo, o papel de diferentes partes interessadas nos estágios da vida de uma desinformação (criação, consumo, etc.) e suas formas particulares de ver e entender uma mesma mensagem digital. No entanto, nosso mapeamento mostrou — até a redação desta tese — que a literatura carece de métodos e ferramentas epistêmicas que auxiliem os projetistas interessados na desinformação digital a entenderem os problemas relacionados ao fenômeno de forma abrangente<sup>1</sup> e que considerem a sua natureza sociotécnica. Isso leva a limitações que restringem a análise e definição de problemas, ignorando aspectos relevantes. Assim sendo, lidar com a natureza sociotécnica do fenômeno da desinformação digital constitui um problema desafiador em aberto, justificando a necessidade de mais estudos no tema.

## 1.1 PROBLEMA DE PESQUISA E OBJETIVOS

Esta tese considera as limitações no entendimento da natureza sociotécnica do fenômeno da desinformação identificadas — 1) a abordagem fragmentada de aspectos sociotécnicos, 2) a caracterização genérica das partes interessadas, e 3) análises segmentadas dos estágios de vida da desinformação, como partes de um problema conceitual (Oulasvirta e Hornbæk, 2016) — relacionado aos desafios no entendimento do fenômeno. O objetivo geral dessa pesquisa foi investigar e desenvolver uma abordagem para apoiar os projetistas no entendimento da desinformação como um fenômeno sociotécnico, auxiliando na caracterização do contexto de casos de desinformação e na compreensão abrangente de problemas. Para isto, os objetivos específicos são:

1. Delinear um panorama das pesquisas que investigam a desinformação digital e a abrangência com que abordam o fenômeno;

---

<sup>1</sup>Por abrangente, esta tese considera abordar as partes interessadas, os estágios do ciclo de vida e aspectos sociotécnicos inter-relacionados.

2. Investigar uma abordagem que auxilie no entendimento sociotécnico de casos de desinformação digital;
3. Investigar como abordar a estrutura de um caso de desinformação de modo integrado, considerando os diferentes estágios da vida de uma desinformação e suas características;
4. Avaliar a capacidade da abordagem em resolver o problema de pesquisa com o público-alvo;
5. Produzir e avaliar recursos para apoiar a aplicação da abordagem pelos projetistas;
6. Disponibilizar a abordagem, seus artefatos e tutoriais de aplicação para uso aberto.

## 1.2 MATERIAIS E MÉTODOS

O método de pesquisa é baseado na investigação e resolução de problemas construtivo-conceituais. Esse tipo de investigação tem a finalidade de desenvolver, esclarecer e modificar conceitos, ideias e artefatos para a formulação de abordagens posteriores (Oulasvirta e Hornbæk, 2016). A Figura 1.1 apresenta os passos, materiais e métodos usados para alcançar os objetivos.

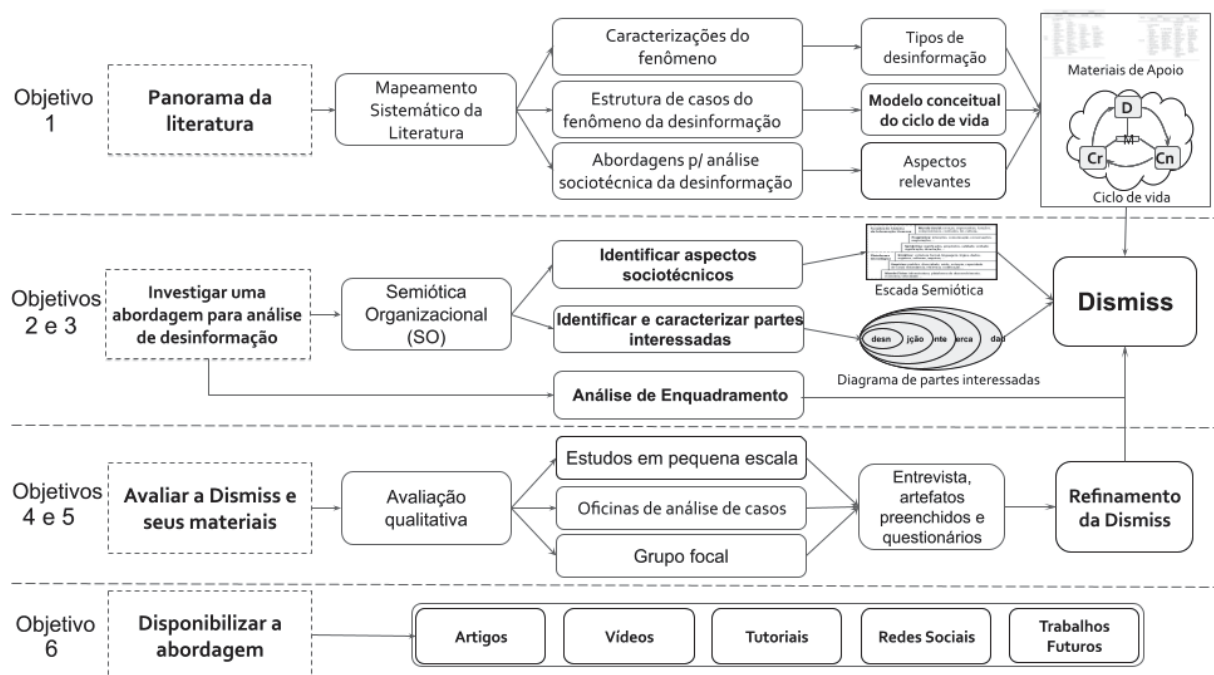


Figura 1.1: Visão geral dos materiais e métodos da pesquisa.

Os objetivos específicos foram alcançados da seguinte forma:

- Objetivo 1) - Realizamos um mapeamento sistemático da literatura (Capítulo 3). O protocolo do estudo e os resultados foram avaliados em seções baseadas no método de grupo focal <sup>2</sup>. O mapeamento revelou aspectos relevantes para abordar a desinformação

<sup>2</sup>Como parte de um método de pesquisa construtivo, pesquisadores em Interação Humano-Computador avaliaram diferentes estágios da pesquisa em sessões baseadas no método de grupo focal (detalhados no Capítulo 5).

digital de modo sociotécnico e sistêmico, como os estágios da vida de uma desinformação, os tipos de desinformação digital e suas características, e os aspectos e mecânicas sociais relevantes que devem ser considerados na análise de desinformações;

- Objetivo 2 e 3) - Com base nos resultados do mapeamento, propusemos a DISMISS (Capítulo 4 ) como uma abordagem que considera os aspectos relevantes a serem analisados em diferentes níveis de abstração na perspectiva sociotécnica. Os resultados do mapeamento também forneceram informações para a construção do Modelo Conceitual do Ciclo de Vida da Desinformação, o qual ajuda a lidar com as dificuldades ao se abordar a estrutura de um caso de modo interligado.
- Objetivo 4) - Como parte de um método de pesquisa construtivo, a DISMISS foi avaliada por grupo focal, em estudos de pequena escala realizados pelo autor e oficinas de análise de casos de desinformação com representantes do público-alvo (utilidade e facilidade de uso percebidas). Os resultados auxiliaram a desenvolver e refinar a DISMISS.
- Objetivo 5) - Com base nos resultados dos objetivos 2 e 3, o autor desenvolveu um manual e tutoriais para apoiar a aplicação da abordagem. Esse material de apoio foi avaliado por grupo focal (utilidade percebida) e em oficinas para a análise sociotécnica de desinformação digital, possibilitando aprimoramentos.
- Objetivo 6) - A divulgação da Dismiss, com seus métodos e ferramentas, se dá por meio de um canal no Youtube<sup>3</sup> criado para a divulgação da Dismiss, apresentando vídeos explicativos sobre a abordagem, exemplos de análises sociotécnicas de casos de desinformação, além de tutoriais sobre a fundamentação teórica e outras possíveis aplicações da Dismiss. Além disso, o desenvolvimento dos conceitos e componentes da abordagem estão disponibilizados em artigos publicados em revistas e conferências .

### 1.3 CONTRIBUIÇÕES

As seguintes contribuições são destacadas:

- Contribuições teóricas:
  - O mapeamento sistemático fornece um panorama sobre as abordagens, desafios e oportunidades de pesquisa no estudo do fenômeno da desinformação.
  - Um entendimento para a desinformação digital considerando a natureza sociotécnica do fenômeno.
  - A caracterização de um conjunto de tipos de desinformação e seus aspectos sociotécnicos.
  - Um modelo conceitual da estrutura de um caso de desinformação (Ciclo de Vida).
- Contribuições metodológicas:
  - A Dismiss: abordagem para análise sociotécnica de casos de desinformação.
  - Artefatos de apoio a aplicação da Dismiss.
  - A disponibilização da Dismiss para uso aberto.

---

<sup>3</sup><https://www.youtube.com/@alissonpuska>

- Publicações:

- *DisMiss False Information: A Value Matter* (Puska et al., 2019).
- *WhatsApp and false information: a value-oriented evaluation* (Puska et al., 2020).
- *Pathway to a human-values based approach to tackle misinformation online* (Piccolo et al., 2020).
- *Exploring Digital Misinformation as a Sociotechnical Phenomenon: Insights from a Small-scale Study* (Puska e Pereira, 2023).

#### 1.4 ESTRUTURA DA TESE

O Capítulo 2 apresenta os fundamentos teóricos do fenômeno da desinformação digital e da Semiótica Organizacional considerados nesta pesquisa. O Capítulo 3 aborda o mapeamento da literatura. O Capítulo 4 apresenta a abordagem, o ciclo de vida da desinformação e os artefatos da S.O adaptados ao contexto do problema de pesquisa. O Capítulo 5 apresenta um resumo das avaliações. Capítulo 6 mostra os resultados alcançados e os próximos passos.

## 2 FUNDAMENTAÇÃO TEÓRICA

Essa tese adota a perspectiva da Semiótica Organizacional para estabelecer os fundamentos para investigação da desinformação digital. Além da SO, essa tese também é inspirada no método de Análise de Enquadramento (AE) — um método multidisciplinar surgido na sociologia usado para analisar como pessoas entendem situações e atividades, e como estruturam o entendimento de suas experiências (Goffman, 1974). Esse capítulo aborda a perspectiva da tese em relação a natureza sociotécnica da desinformação digital e apresenta os fundamentos da SO e da AE usados no desenvolvimento da Dismiss.

### 2.1 A DESINFORMAÇÃO DIGITAL COMO UM FENÔMENO SOCIOTÉCNICO

A desinformação digital é um fenômeno endógeno à qualquer Tecnologia de Informação e Comunicação (TIC), decorrente de alguma forma de "manuseio incorreto" de informações (Luciano, 1996). Isso quer dizer que ela surge da interação entre pessoas através de TIC's, sendo caracterizada por problemas que afetam tanto entendimento humano — como a falta de clareza, de objetividade, e de completude de um conteúdo (Luciano, 1996)) — quanto a operação destas tecnologias (como a ocorrência de ruído, interferência, *hacking*, perda de dados, etc. (Starbird et al., 2019)). Dessa forma, uma desinformação pode ser causada por aspectos técnicos dos dispositivos usados para comunicação e por aspectos sociais relacionados aos fatores humanos e socioculturais dos participantes da comunicação (Puska et al., 2019).

Esta tese entende que uma desinformação digital pode ser caracterizada em duas categorias: desinformação acidental e deliberada (Tudjman e Mikelic, 2003). A desinformação acidental pode ser definida como informação deturpada que causa confusão, geralmente atribuída a erros do indivíduo, como falhas de memória ou erros de digitação (Hinz, 2007), ou a erros técnicos como ruído ou interferência no canal de comunicação (Zhou e Zhang, 2007).

A Figura 2.1 mostra um exemplo de desinformação acidental. O Enquadramento 1 apresenta uma publicação do autor no seu canal de comunicação pessoal (a sua conta na plataforma Instagram), e o Enquadramento 2 apresenta um reenquadramento da mesma publicação em outro canal de comunicação (a conta do grupo de pesquisa do autor na mesma plataforma). A primeira publicação alcança uma audiência particular ao autor e que interpretam as mensagens com base na relação que possuem com ele (esperam postagens pessoais e de abrangência limitada). Por outro lado, o segundo quadro, que adiciona somente o elemento "Origem/Fonte" como sendo a conta do grupo de pesquisa, terá um alcance diferente e atingirá outro tipo de audiência. Essa reutilização da mensagem original, mesmo preservando as informações de autoria do conteúdo, adiciona outros elementos para o enquadramento da mensagem que influenciam o entendimento dela. Por exemplo, a finalidade da conta do grupo de pesquisa (divulgação científica) e as expectativas da audiência (esperar um evento abrangente). Nesse caso, nem o autor, nem o grupo de pesquisa tiveram a intenção de causar confusão nas diferentes audiências do conteúdo.



Enquadramento 1



Enquadramento 2

Figura 2.1: Uma representação de desinformação accidental. Fonte: Autor.

Por outro lado, a desinformação deliberada corresponde à informação falsa criada com a intenção de enganar, gerar erros de decisão, e manipular os sistemas de crenças dos indivíduos (Tudjman e Mikelic, 2003). O aspecto intencional é obrigatório e caracteriza o uso da desinformação em diferentes contextos da Internet para o engano/manipulação, como o *hacking* cognitivo (Cybenko et al., 2002) que utilizando fatores humanos e engenharia social para persuadir ou enganar os usuários, ataques de negação de serviço que geram falhas de armazenamento e distribuição (dados corrompidos ou incompletos), e atrapalham a obtenção de uma informação clara (Zhou e Zhang, 2007), ou até a exploração de aspectos técnicos relacionados a utilização de um determinado software, como alguns tipos de padrões de design abusivos (Baroni et al., 2021).

A Figura 2.2 mostra um recorte de uma notícia sobre um golpe (estelionato) que utilizou o WhatsApp e desinformação para causar prejuízo econômico. O golpista explora a natureza anônima e pouco segura da plataforma WhatsApp (Puska et al., 2020), além de aspectos persuasivos para se passar por uma parte interessada que quer vender um ingresso para um show badalado. Aspectos sociotécnicos como a falta de recursos para verificar a validade do ingresso ou a identidade e credibilidade do suposto vendedor (a forma como o usuário se posiciona para os outros, a linguagem que utiliza, etc.), além da urgência (Cialdini, 1993) e existência de compradores buscando um ingresso para o evento eminente, entre outros, são fatores importantes de serem observados e entendidos.





Figura 2.2: Um caso de estelionato (desinformação deliberada). Fonte: G1 (G1, 2023a).

É possível perceber que o fenômeno engloba tanto aspectos acidentais quanto deliberados que vão desde aspectos sociais (p.ex. as diferenças entre relacionamentos pessoais e profissionais) até aspectos técnicos (p.ex. a qualidade das informações disponíveis no canal de comunicação para avaliação de credibilidade/veracidade) na utilização de TIC's por pessoas. Os exemplos servem para demonstrar a diversidade de características que existem em diferentes níveis de abstração que podem constituir um caso de desinformação. Isso representa a complexidade de investigar e compreender a natureza sociotécnica do fenômeno da desinformação.

### 2.1.1 Os elementos de um caso de desinformação

Como identificado no resultado do mapeamento da literatura (Capítulo 3), existem duas formas de abordar o fenômeno para estudá-lo (Zhang et al., 2019a): abordagens em nível de postagem/mensagem (profundidade) — abordam uma única mensagem de mídia digital em profundidade -, e abordagens em nível de evento (largura) — abordam um grupo de mensagens que constituem um evento. Como simplificação, esta tese entende que uma ocorrência de desinformação pode ser observada em um contexto de comunicação, e esse contexto configura um caso de desinformação, podendo ser analisado em profundidade e em largura.

Um caso de desinformação pode ser caracterizado por três elementos básicos que delineiam o contexto de um caso de desinformação digital:

1. Partes Interessadas: são os participantes do caso, diferentes atores — cada um com seu entendimento, propósito e interesses — que interagem com a desinformação em diferentes circunstâncias.
2. Canais de Comunicação: são os meios de comunicação por onde as partes interessadas se comunicam, onde interagem, transmitem e recebem mensagens.
3. Objetos Desinformativos: são as mensagens digitais que se difundem pelos canais de comunicação e que podem ou não representar desinformações para as partes interessadas.

As Partes Interessadas trazem diversos aspectos relevantes para compreender o fenômeno. Elas vivem em um determinado contexto (Shu et al., 2019), têm relacionamentos e reconhecem autoridades (Wilder e Vorobeychik, 2018), têm posicionamentos e interesses (Hao, 2021), etc. Além disso, seres humanos não são bons em identificar mentiras e falsidades (Caulfield et al., 2019; Kasra et al., 2018), sendo suscetíveis à falhas de julgamento que afetam a percepção de risco e avaliação de veracidade (Halbach et al., 2019), ligadas a estratégias usadas diariamente para agilizar a tomada de decisão (Pohl e Pohl, 2004). As estratégias são influenciadas por aspectos contextuais e subjetivos na percepção do que cada indivíduo reconhece em cada situação, como autoridades (Zannettou et al., 2017), conhecimento/posicionamento prévio (Cholvy, 2014), afinidades (Huang et al., 2019), discriminação de gênero/raça (Armstrong e McAdams, 2009), o tipo de publicação (Halbach et al., 2019), costumes ao consumir informações digitais (Torres et al., 2018a), dentre outros. Esses aspectos de influência podem ser explorados deliberadamente (Aigner et al., 2017; Kovic e Laissue, 2016) ou ocorrer acidentalmente (Shen et al., 2019; Valdez e Ziefle, 2018). Essa amalgama de aspectos (fatores humanos, sociais e culturais) demonstra a complexidade de se compreender as partes interessadas de um caso de desinformação, elemento desafiador para a solução de problemas no fenômeno.

O Canal de Comunicação limita ou aprimora a capacidade das partes interessadas de se comunicar (McLuhan, 1974; Zhou e Zhang, 2007), estabelece relações de poder no controle da interação (Vargo et al., 2018; Rehm, 2017), coloca as partes interessadas em diferentes papéis (Chen et al., 2019; Zhou e Zhang, 2007), etc. O canal de comunicação molda a interação das partes interessadas de diferentes modos, proporcionando diferentes "Affordances Tecnológicas" <sup>1</sup> (Starbird et al., 2019; Murimi, 2018), podendo ser físicas ou sociais (Stamper e Liu, 1994). Por exemplo, a sincronicidade, direcionalidade e dispersão (Zhou e Zhang, 2007) do canal de comunicação podem colocar as partes interessadas em papéis diferentes na comunicação (Dawson e Innes, 2019), com capacidades diferentes de controle (Pearce e Malhotra, 2022), levando a vantagens e desvantagens. Por exemplo, utilizar um canal de comunicação que carece de mecanismos de controle de difusão de desinformação digital e ferramentas de detecção e avisos aos usuários (Puska et al., 2020; Konstantinou et al., 2019; Gao et al., 2018). Além disso, cada meio tem limitações para a formatação da mensagem (Janicka et al., 2019), possibilitando imagens, sons, e vídeos, ou limitando a quantidade de caracteres aceitas, o tamanho de um arquivo de imagem ou o tempo de um vídeo. Esses diferentes aspectos mostram os impactos do canal de comunicação nas capacidades de difusão da desinformação e na interação das partes interessadas com uma desinformação digital, indicando a relevância desse elemento para compreender os problemas do fenômeno. Essa tese considera o Canal de Comunicação como sendo um complexo de elementos que engloba dispositivos (p.ex. notebook, televisão, celular, etc.), tecnologias (p.ex. 3G, wifi, ethernet, etc.), e plataformas de comunicação (p.ex. WhatsApp, twitter, skype, etc.).

O Objeto Desinformativo possui aspectos que influenciam a percepção de risco (Fan et al., 2019; Xiong et al., 2018), o julgamento de valor (Baxter et al., 2019) e avaliação de veracidade das partes interessadas (Popat et al., 2017). Esses aspectos são atributos que podem ser vistos como "Affordances do Objeto Digital", podendo ser físicas ou sociais. Por exemplo, atributos físicos que afetam a motivação para interagir com o objeto digital como a quantidade de caracteres, o tamanho de um áudio/vídeo, e as cores usadas (Li e Chong, 2019). Atributos que organizam o conteúdo e guiam o fluxo de leitura (Gunawan e Suwandi, 2020; Shen et al., 2019), como a diagramação, o posicionamento de figuras, índices, e listas. Atributos que chamam a atenção e enfatizam o conteúdo, como rótulos, títulos e subtítulos (Özgöbek e Gulla,

<sup>1</sup>Nessa tese, o termo "*affordances*" significa "aquilo que, na interação com o sujeito, o ambiente potencialmente proporciona, oferece ou convida a agir em razão das propriedades ou características de reconhecimento que possui, ou deveria possuir, para satisfazer certas necessidades de uso" (McCabe, 1982).



2017). Exemplos de atributos sociais como aspectos de significado — o uso de uma linguagem característica de uma comunidade (Addawood et al., 2019), uso de imagens com famosos (Galeotti, 2019), ou palavras que evocam sensações (Zannettou et al., 2019; Chen et al., 2015c).

Esta tese contribui conceitualmente com a perspectiva de que antes de se propor soluções para os problemas do fenômeno da desinformação é preciso entender a desinformação de modo abrangente, considerando seus elementos e como os aspectos sociotécnicos estão inter-relacionados. Por isso é preciso identificar os elementos básicos de um caso e caracterizá-los para realizar uma análise sociotécnica abrangente e ampliar o entendimento sobre as circunstâncias em que a desinformação ocorre. Para isso, essa tese adota as lentes e artefatos da Semiótica Organizacional (SO).

## 2.2 SEMIÓTICA ORGANIZACIONAL E OS ASPECTOS SOCIOTÉCNICOS

A SO é a disciplina que estuda como a informação pode ser melhor usada no contexto de atividades organizadas (Stamper et al., 2000). Ela se inspira na Semiótica Peirceana (Peirce e Buchler, 1902) para examinar a natureza, as características e os recursos da informação em organizações humanas (Liu, 2000). Seu objetivo está relacionado com o desenvolvimento de enquadramentos teóricos, métodos e técnicas para a compreensão, análise, modelagem, concepção e implementação de Sistemas de Informação (SI) (Stamper et al., 2000). Ela dispõe de métodos e artefatos para análise das práticas de uma determinada organização, capazes de identificar aspectos sociotécnicos relevantes para o entendimento de limitações no desenvolvimento e uso de SIs. Como o fenômeno da desinformação pode ser abordado como uma atividade organizada, a SO representa uma perspectiva teórica adequada para investigar o fenômeno de modo sociotécnico.

A SO considera que qualquer atividade organizada é impactada pelo uso e interpretação de signos (Liu, 2000). Na SO, Stamper (Stamper, 1991) explica o signo como "algo que representa alguma coisa para alguém em algum aspecto ou capacidade em algum contexto social". A SO entende o signo como a menor partícula concreta de uma representação da realidade. Um signo pode denotar qualquer coisa, de conhecimento e informação, até comportamentos e objetos no ambiente de interação (Andersen, 1997). Essa perspectiva possibilita estudar os elementos de um caso de desinformação, os signos que compõem um objeto desinformativo, como um signo pode representar uma desinformação, e quais aspectos impactam a expressão e interpretação dos participantes da comunicação, para construir um entendimento abrangente do fenômeno da desinformação.

### 2.2.1 A Escada Semiótica

Um signo carrega propriedades sobre o que ele representa (Stamper, 1991). A SO propõe a Escada Semiótica (ES) — Figura 2.3 — como um artefato para a análise das propriedades de signos e SI's (Stamper, 1991). A Escada Semiótica divide as propriedades dos signos em duas dimensões semióticas: as relacionadas ao Sistema de Informações Humanas e as presentes na Plataforma Tecnológica (Liu, 2000). O Sistema de Informação Humano aborda os aspectos subjetivos da percepção ou concepção de um signo, como a interpretação e a construção de significado. O Sistema de Informação Humana divide-se nos níveis Social, Pragmático e Semântico. A Plataforma Tecnológica considera como a informação é formatada, armazenada, e transmitida, dividindo-se nos níveis Sintático, Empírico e Físico.

|                              |                                                                                                                                                                                                                                                                                                                        |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sistema de Informação Humana | <b>Mundo Social:</b> compreende os efeitos dos signos e as mudanças que causam no mundo social. Considera relacionamentos, expectativas, compromissos, crenças, leis. No fenômeno da desinformação: proximidade afetiva, hierarquia, autoridades, deveres, direitos, popularidade, etc;                                |
|                              | <b>Pragmático:</b> é onde a interação ganha propósito e sentido. Considera a função da comunicação e da mensagem, como as intenções que diferenciam a desinformação deliberada ou acidental, as motivações para compartilhamento, campanhas de desinformação, as intenções e expectativas;                             |
|                              | <b>Semântico:</b> consiste da essência da subjetividade individual da interpretação. Por exemplo, nesse nível estão aspectos sobre gênero textual e suporte textual, linguagem sentimental, temática do conteúdo das mensagens, posicionamentos, veracidade, validade, entre outros;                                   |
| Plataforma Tecnológica       | <b>Sintático:</b> compreende a dinâmica das estruturas e padrões dos signos. Aspectos relacionados à características estruturais da desinformação e do meio de comunicação. No fenômeno: a diagramação do conteúdo, estruturas lógicas da rede social, elementos de destaque na mensagem, etc;                         |
|                              | <b>Empírico:</b> considera a formação de estruturas e padrões com base nas capacidades de indivíduos e canais. No fenômeno: padrões de propagação, aspectos relacionados à interação entre a tecnologia e partes interessadas (como a sincronidade e direcionalidade do canal), padrões entre mensagens, entre outros. |
|                              | <b>Mundo Físico:</b> envolve qualquer diferença fisicamente discernível no estado, ação ou característica de indivíduos e canais. Por exemplo, o tipo de hardware, distância entre as partes interessadas, o poder computacional do dispositivo, custos, etc;                                                          |

Figura 2.3: Escada Semiótica Adaptada. Inspirada: Ferreira (2015).

Esta tese emprega a ES como um artefato de apoio para o estudo sociotécnico do fenômeno da desinformação. Ela ampara análise da desinformação digital, flexibilizando e estruturando a identificação de aspectos sociotécnicos em diferentes níveis de abstração no contexto em que a desinformação ocorre.

### 2.2.2 As Partes Interessadas

No contexto da SO, o signo está em um meio compartilhado por diversos participantes que compartilham o seu significado e o usam para executar tarefas, se comunicar, se organizar, etc(Liu, 2000). Esses participantes são chamados de partes interessadas e o conjunto de partes interessadas que compartilham um entendimento coletivo para os signos é chamado de organização. No fenômeno da desinformação as partes interessadas de uma organização podem realizar diversas tarefas: quem cria a desinformação, quem consome (como a audiência), quem difunde (como um *bot* no Facebook), quem corrige/mitiga(como os jornalistas), quem controla o canal de comunicação (como o próprio Facebook), etc.

Dessa forma, o fenômeno da desinformação envolve uma diversidade de partes interessadas que possuem diferentes perspectivas, demandas, expectativas, e entendimentos subjetivos à cada uma. Nesse sentido, é fundamental identificar as diferentes partes que estão em um caso de desinformação digital para entender os aspectos relevantes do contexto em que a desinformação ocorre. A SO propõe o Diagrama de Identificação de Partes Interessadas (DPI) como um artefato para o mapeamento de partes interessadas e seus papéis (Stamper, 1994). O DPI apoia o levantamento de diferentes partes interessadas que podem afetar e ser afetadas de maneira direta ou indireta pela desinformação. Esse artefato permite construir uma visão abrangente dos participantes de um caso de desinformação, identificando autoridades, partes influenciadoras, audiência, etc. A Figura 2.4 apresenta o DPI e seus grupos.

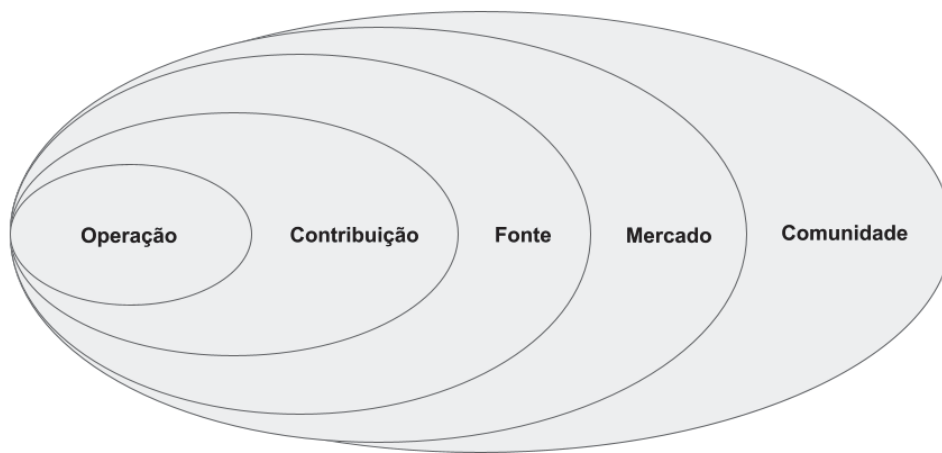


Figura 2.4: Diagrama de Partes Interessadas. Traduzido de Stamper (1994)

O DPI considera as partes interessadas de uma determinada organização agrupando-as em em cinco grupos:

- **Operação:** O DPI original agrupa nesse nível as partes interessadas envolvidas diretamente no funcionamento do projeto de um SI. Para o desenvolvimento dessa tese, esse nível considera a ocorrência de desinformação no lugar do projeto — o caso de desinformação.
- **Contribuição:** Esse nível agrupa as partes interessadas relacionadas ao andamento do projeto e desenvolvimento da solução. Esse nível foi adaptado para considerar os atores envolvidos na concepção e disponibilização inicial da desinformação, como quem cria, financia, planeja, prepara e disponibiliza.
- **Fonte:** Nesse nível estão aqueles que fornecem informações e/ou são fonte de dados ao problema ou a sua solução, ou que ainda fazem o uso deles. Esse nível corresponde a atores que interagem com a desinformação diretamente, disseminando e consumindo, como a audiência pretendida da desinformação e os atores que fornecem informações para concepção e difusão da mensagem em uma rede social, e simpatizantes.
- **Mercado:** No nível Mercado estão as partes interessadas que possuem alguma relação de parceria ou concorrência com o projeto. Nessa tese, o nível 'Mercado' associa as partes interessadas parceiras e concorrentes da desinformação, como redes sociais, provedores de Internet, influenciadores, e checadores de fatos.
- **Comunidade:** O nível comunidade consiste das partes interessadas que podem influenciar ou ainda sofrer influência da solução. Ela se refere às partes interessadas que são impactadas indiretamente pela desinformação como clientes de uma marca sendo difamada ou legisladores que precisam lidar com os problemas do fenômeno.

### 2.2.3 Um Entendimento para Desinformação Inspirado na Semiótica Organizacional

Para estudar o fenômeno da desinformação por uma perspectiva sociotécnica, é necessário um entendimento adequado para o termo "desinformação". Nos resultados do mapeamento da literatura apresentado no Capítulo 3, *não há consenso da comunidade acadêmica de Computação*

*em relação ao conceito de desinformação, predominando um conceito genérico como "informação falsa" que pode ser acidental ou deliberada.* Abordar a influência que a subjetividade de cada circunstância tem na percepção de uma desinformação não é uma tarefa trivial (Cho et al., 2019), e não foi encontrado um entendimento para o termo "desinformação" que deixe explícita a importância da perspectiva de quem interage com a desinformação em um caso do fenômeno. Dessa forma, é importante discutir um entendimento para "desinformação" que operacionalize o estudo do fenômeno pela perspectiva sociotécnica considerando as pessoas que se comunicam e enfatize sua experiência (e importância).

A desinformação pode ser entendida como um signo que não corresponde a um estado das coisas na realidade que um indivíduo percebe ou entende (Eco, 1979)<sup>2</sup>. Esta tese entende que o significado de algo depende da perspectiva subjetiva de um indivíduo, assim como ocorre para os outros indivíduos que estão experienciando este mesmo algo. O "significado" é algo ao mesmo tempo individual e compartilhado, que se transforma durante a interação, sendo influenciado por diversos fatores, como as experiências prévias, o conhecimento do indivíduo, as suas capacidades de interação com/e do seu ambiente (*affordances*) (Liu, 2000), e a sua configuração social (como casos do efeito manada) (Nuruddin Hidayat et al., 2020). Assim, aquilo que um indivíduo de uma organização percebe como verdade pode não ser percebido por outro com o mesmo significado. Dessa forma, essa tese entende que a veracidade de uma mensagem está para um indivíduo, e é intimamente relacionada às circunstâncias que a mensagem é observada, além de se transformar com o tempo. Nesse contexto, a objetividade factual do conteúdo de uma mensagem pode ter pouco impacto na interpretação das partes interessadas que interagem com ela.

Considerando a perspectiva de intencionalidade apresentada em (Tudjman e Mikelic, 2003), as ideias de signo de Stamper (Stamper, 1973, 1991) e de Eco (1979) essa tese entende a desinformação acidental como o uso não intencional de um signo que não corresponde a um estado das coisas na realidade que um indivíduo (em uma determinada configuração social) percebe ou entende em uma certa circunstância. Do mesmo modo pode-se definir desinformação deliberada como o uso intencional de signos que não correspondem a um estado das coisas na realidade que um indivíduo (em uma determinada configuração social) percebe ou entende em uma certa circunstância, com um propósito particular em manipular o entendimento deste indivíduo para obter algum tipo de ganho.

Pela perspectiva adotada nesta tese, uma desinformação no formato digital não é só um conjunto de aspectos físicos (digitais) que representam um objeto digital, mas um agregado de aspectos humanos relacionados a percepção, julgamentos, emoções, intenções, entre outros. Esses aspectos são exemplos de aspectos relevantes para entendimento do fenômeno, como motivações, sentidos e expectativas que constituem elementos sociais da desinformação que vão além de um fato como percebido objetivamente. Portanto, a desinformação no formato digital é um agrupamento de aspectos perceptivos e funcionais que implicam uns aos outros sob a "ideia de um objeto digital"<sup>3</sup>, ou seja, uma representação, signo ou conjunto de signos<sup>4</sup>. Com base nessa perspectiva, podemos instrumentalizar o estudo do fenômeno sem que a noção de verdade seja removida do objeto digital (signos) em análise, mas posta entre parênteses como sendo a perspectiva de quem experiencia o signo em um certo contexto de comunicação social.

<sup>2</sup>A compatibilidade das teorias, técnicas e conceitos usados na construção da Dismiss são discutidos no Apêndice A.1.1 que trata das decisões teóricas do autor para essa tese.

<sup>3</sup>mais sobre objetos digitais em Spiegelberg (1994).

<sup>4</sup>mais sobre representações constituídas de signos em (Peirce e Buchler, 1902).

### 2.3 ANÁLISE DE ENQUADRAMENTO E O FENÔMENO DA DESINFORMAÇÃO

A desinformação digital engloba uma diversidade de aspectos que impactam como as partes interessadas percebem e avaliam um conteúdo e como são impactadas pela desinformação (Valdez e Ziefle, 2018). A forma como uma mensagem é apresentada, as palavras usadas, as imagens, as cores, a ideia apresentada/concluída, são aspectos da mensagem que se conectam com as diferentes partes interessadas de diferentes formas. Esses aspectos levam a um determinado entendimento, ou seja, enquadram a mensagem de um determinada forma para um determinado alguém. A ideia de enquadramento da sociologia da interação (Goffman, 1974), significa as diferentes perspectivas (vieses) empregadas na cobertura de notícias, diferentes enquadramentos usados na comunicação (D'angelo, 2018). Enquadramentos, de acordo com Goffman (1974), são estruturas interpretativas que as pessoas usam para entender e se comportar em situações sociais. Essas estruturas são compostas por expectativas, suposições e crenças e podem ser implícitas ou explícitas, variando de acordo com o contexto social e as circunstâncias. Goffman também argumenta que as pessoas podem usar molduras para enquadrar a maneira como os outros as percebem através de sua comunicação e comportamento estratégico.

Nessa tese, o termo enquadramento é usado para referenciar os aspectos sociotécnicos que constituem um caso de desinformação (o contexto do caso: a mensagem, os canais de comunicação e as partes interessadas) e que se relacionam com o entendimento de diferentes partes interessadas de diferentes formas. Na literatura de Computação, a análise de enquadramento foi empregada para tratar o fenômeno da desinformação (Verdoliva, 2020; Hamborg et al., 2019, 2017), sendo um método útil para ser adaptado para análise sociotécnica de casos de desinformação.

O método de Análise de Enquadramento (AE) estuda a organização da experiência de interação social (como uma experiência é organizada em uma determinada situação) (Mendonça e Simões, 2012). Ele compreende a forma como uma informação é organizada em um determinada circunstância e quais são os elementos na comunicação que influenciam a interpretação dos participantes (Goffman, 1974). Na Teoria do Enquadramento esse são chamados de meta-fatores da comunicação (Mendonça e Simões, 2012). Para esta tese, os meta-fatores estão relacionados com a natureza sociotécnica do fenômeno (como tratado por trabalhos na área de Computação que abordam a multimodalidade (Jindal et al., 2020)) e influenciam a forma como uma desinformação é construída, armazenada, compartilhada, consumida e mitigada. Por exemplo, aspectos do meio de comunicação como a diagramação, o tamanho da fonte, cores, imagens, e recursos de interação enquadram um conteúdo de um determinada forma. Esses aspectos serão percebidos pelas partes interessadas e dão um apoio ao entendimento da mensagem por uma determinada perspectiva/ideia. Por isso, esta tese entende os meta-fatores como sendo aspectos sociotécnicos<sup>5</sup>.

A AE auxilia a entender a estrutura funcional do comportamento social pela ótica da interação (Goffman, 1974), sendo aplicável não só para análise de notícias, mas pra qualquer outro tipo de mensagem digital em qualquer contexto. Além disso, a perspectiva de enquadramento traz uma orientação conceitual para sintetizar a quantidade de aspectos sociotécnicos que constituem as circunstâncias em que um objeto desinformativo é observado e dar um apoio para as interpretações e conclusões de um analista. Essa tese usa a AE como um procedimento-síntese, inspirado em D'angelo (2018), como método de análise de objetos de desinformação digital (como analisar, o que analisar, porque analisar, etc.), que guia a identificação de aspectos

---

<sup>5</sup>Nessa tese, os elementos que constituem um caso de desinformação e suas características são observados como indícios sobre a perspectiva de quem cria um objeto desinformativo, a ideia que seu conteúdo passa, para qual audiência ele foi criado, onde ele foi difundido, a relação do conteúdo com o ambiente das partes interessadas, qual audiência o está consumindo, etc.



sociotécnicos capazes de caracterizar a interação de uma parte interessada específica com um objeto desinformativo específico.

A Figura 2.5 apresenta um exemplo de um enquadramento de uma desinformação digital que foi analisada nesta tese<sup>6</sup>.



Figura 2.5: Exemplo de enquadramento.

No primeiro quadro temos uma imagem de uma pessoa famosa, um político, fazendo uma expressão estressada, apontando o dedo no ar como se estivesse enfatizando um ponto de vista. A imagem vai evocar diferentes sensações, dependendo das características do contexto onde é observado. No Enquadramento 2, foi adicionado um texto que induz a uma conclusão sobre a imagem, sendo um título de tamanho proporcional à imagem e com destaque negrito. No quadro 3, foi adicionado um elemento que referencia uma fonte de informações, o site "agora notícias". Além disso, no mesmo quadro também foi adicionado um rótulo que resume a ideia passada no título e dá um apoio pra interpretação da mensagem que objeto digital passa. No quadro 4, foi adicionada outra figura com a imagem de uma pessoa desconhecida encapuzada, portando uma arma e apontando-a para o espectador da mensagem. Ainda, um elemento que referencia um grupo do Instagram "conservadoresbr", como se endossasse a mensagem.

Como a Figura 2.5 demonstra, um objeto digital de desinformação apresenta uma complexidade de aspectos em diferentes níveis de abstração que constroem uma ideia e dão um apoio para interpretação da audiência. Nesse sentido, é possível investigar, através dos elementos que constituem um caso de desinformação e suas inter-relações, qual a perspectiva de quem cria, de quem consome e de quem compartilha, observando os aspectos sociotécnicos em cada contexto de interação (diferentes canais e partes interessadas que possuem e proporcionam diferentes capacidades de interação).

A AE emprega a análise qualitativa de conteúdo, observando elementos visuais, textuais, simbólicos, de organização do conteúdo, e a relação entre os elementos e o ambiente da

<sup>6</sup>A análise foi realizada utilizando a Dismiss e pode ser vista em detalhes no link <https://youtu.be/LfMrysNp9p8>

audiência (Matthes e Kohring, 2008). Seu objetivo é identificar a relação entre aspectos contextuais da comunicação com elementos da mensagem que impactam na forma como uma mensagem é observada, seja por quem a constrói com um propósito, seja por quem a consome com uma expectativa (Matthes e Kohring, 2008). A AE foi adotada para auxiliar a lidar com a subjetividade de interpretações das partes interessadas, auxiliando a entender o posicionamentos sobre a veracidade percebida por elas.

A AE possui diversas abordagens metodológicas, sendo usada extensivamente pelos acadêmicos da área de Comunicação e Jornalismo para análise de diversos tipos de objetos digitais, de discursos políticos e notícias sensacionalistas, até propagandas de produtos e desinformações (Matthes e Kohring, 2008). *No entanto, os métodos de AE encontrados na literatura carecem de clareza conceitual e padronização operacional (D'angelo, 2018; Matthes e Kohring, 2008), sendo dependentes da experiência do analista e vulneráveis ao viés interpretativo.* As abordagens de AE mais rigorosas seguem uma linha inspirada na análise de conteúdo e no método *Grounded Theory* (D'angelo, 2018), acarretando em custos para a aplicação, com várias regras para codificação e rodadas de avaliação (como o método apresentada em Starbird et al. (2019), trabalho relacionado a esta tese). Dessa forma, um dos desafios em usar o método AE para a análise sociotécnica de desinformação é que ele não aborda a exploração do contexto da comunicação — por exemplo, o que se deve fazer para identificar diferentes partes interessadas ou como elas podem ser classificadas — de modo sistemático e de fácil reprodução. Por isso, os artefatos da Semiótica Organizacional serão usados para o levantamento de partes interessadas (Diagrama de Partes Interessadas) e identificação de aspectos sociotécnicos (Escada Semiótica) de um caso de desinformação, visando apoiar a aplicação do método AE e torná-lo mais confiável e com resultados reprodutíveis, auxiliando na caracterização do contexto de um caso de desinformação e na justificação de conclusões<sup>7</sup>.

Fazendo uma analogia com a Escada Semiótica, o foco da AE pode ser entendido como o Sistema de Informação Humana das organizações humanas. Cientistas da Comunicação analisam conteúdos midiático para compreender como uma mensagem pode ser interpretada, quais elementos influenciam a interpretação da audiência, qual o propósito que a mensagem foi construída, e assim por diante (D'angelo, 2018). Ao adotar a perspectiva da SO com a AE para construção da Dismiss, integrando a flexibilidade da Escada Semiótica para analisar práticas e SI's, com a abrangência do Diagrama de Identificação de Partes Interessadas no levantamento de participantes para além do obvio de locutor e audiência, e a perspectiva de que cada participante cria um enquadramento para o que observa, resulta em uma abordagem com potencial apoiar o estudo do fenômeno da desinformação para construir um entendimento abrangente de um caso de desinformação. No melhor do conhecimento do autor, não foram encontrados trabalhos que abordam a AE pela ótica da SO.

Como uma definição operacional, esta tese entende o enquadramento como o conjunto de aspectos sociotécnicos em um caso de desinformação que dão "apoio de significado" de como um parte interessada observa um objeto digital de desinformação em uma determinada situação. O Apêndice A.1.1 aborda em detalhes como a Dismiss integra os conceitos e artefatos da SO com os conceitos e técnicas da AE para a análise sociotécnica de casos de desinformação, bem como a preferencia da AE como método-síntese em detrimento de outros (como *Grounded Theory*) encontrados na literatura. Nesse sentido, o Diagrama de Identificação de Partes Interessadas pode ser empregado para identificar as partes interessadas e seus papéis no contexto do caso de desinformação. A partir da perspectiva das partes interessadas, é possível usar a ES para identificar os aspectos sociotécnicos que caracterizem o contexto em que a desinformação ocorre e identificar quais deles tem o potencial de impactar a criação, difusão e consumo de desinformação.

<sup>7</sup>A compatibilidade da AE e da SO é discutida em detalhes nos Apêndices A.1.1



Para guiar a identificação de aspectos sociotécnicos relevantes, pode-se conduzir a análise de enquadramento inspirado na estrutura e diretrizes descritos em (Linstrom e Marais, 2012):

- Determinar a unidade de análise: nesta tese a unidade de análise é o objeto desinformativo, podendo ser um ou vários.
- Determinar o tipo do objeto: a tipo do objeto auxilia a delimitar o que observar na análise.
- Escolher uma mídia: nesta tese a mídia representa o complexo de elementos do caso "canal de comunicação" que engloba dispositivos, tecnologias, e plataformas de comunicação.
- Delimitar o contexto espaço-temporal do caso: nesta tese o contexto espaço-temporal define um universo do qual a amostra dos objetos, partes interessadas e canais de comunicação serão considerados para a análise.
- Analisar o enquadramento do objeto: identificar aspectos sociotécnicos usando a ES. Em caso de mais objetos, identificar aspectos sociotécnicos que estejam inter-relacionados nos objetos para detectar padrões.

Nesta tese, a AE é usada somente para guiar a observação de aspectos sociotécnico pelas lentes da Escada Semiótica, e para sintetizar os aspectos identificados da análise sociotécnica de objetos desinformativos que auxiliam no entendimento do contexto de um caso de desinformação. A Dismiss não incorpora outros elementos da AE, como os tipos de enquadramento e aspectos das estratégias de interação.

### *2.3.0.1 Princípios da influência, vieses cognitivos e falácias lógicas*

Aspectos técnicos, humanos, sociais e organizacionais contribuem para interpretação de um signo de uma determinada forma. Sendo assim, é preciso critérios para identificar aspectos sociotécnicos relevantes para tentar entender os possíveis enquadramentos usados pelas partes interessadas que interagem com o objeto desinformativo e como as características do meio de comunicação influenciam a interação. Considerando o entendimento de desinformação para esta tese, a veracidade da mensagem está para uma parte interessada, não sendo um aspecto objetivo da comunicação de desinformação. Assim, a percepção da veracidade para um indivíduo é definida por aspectos sociotécnicos incorporados do contexto da comunicação, impactando a experiência de interação com o objeto desinformativo e a interpretação de algo como sendo verdade para um indivíduo particular. Dessa forma, a verdade é um enquadramento empregado por um indivíduo em um contexto de interação. Nesse sentido, buscamos amparo na literatura que estuda influência para mapear as características na comunicação que podem influenciar a percepção de credibilidade e veracidade da audiência e que fazem parte tanto da desinformação accidental, quanto são explorados ativamente na ocorrência de desinformação deliberada. Para isso, adotamos as lentes da Teoria da Persuasão (Cialdini, 1993).

A influência e a persuasão constituem o processo de induzir alguém a fazer, reagir, pensar, ou acreditar em algo que se deseja (Cialdini, 1993). Na comunicação de desinformação, um parte interessada induz uma determinada interpretação (ou é induzida a uma determinada interpretação) ao enquadrar uma mensagem de uma determinada forma em um objeto digital desinformativo, seja por acidente ou de modo deliberado. Nesse sentido, aspectos de influência são aqueles que tocam no que é importante para as partes interessadas, como emoções, crenças, valores, e no que e como as pessoas estão pensando em um determinado grupo em um determinado momento (Cialdini, 1993). Por isso, é necessário caracterizar as partes interessadas e mapear o contexto da situação de comunicação onde a desinformação ocorre para identificar aspectos sociotécnicos relevantes e caracterizar a desinformação digital de modo abrangente.

Os fundamentos da influência e persuasão são baseados em manipular a atenção, construir confiança, identificar o estado emocional e os interesses da audiência, e explorar esses aspectos ao se comunicar (Hogan, 2010). Segundo Cialdini (1993), os princípios da influência são:

- **Afinidade:** Pessoas tendem a gostar daquilo que as representa, que se parece com elas. Partes interessadas mal intencionadas exploram elementos como o uso de uma linguagem significativa para uma audiência, uma aparência agradável, usar opiniões positivas, a similaridade na estrutura/diagramação de um conteúdo, o tamanho de uma mensagem, as cores, etc.
- **Aprovação social:** Pessoas querem se sentir validadas pelo que as outras pessoas estão fazendo ao seu redor, em especial quando em dúvida. Eventos ocorrendo no ambiente de um grupo têm impacto no que o grupo percebe como comportamento correto. Elementos desse princípio são homofilia, confiança, consensos, popularidade de um tema ou pessoa, valores morais, credences, etc. Esse princípio ganha força na presença de incerteza, induzindo a tendência de um indivíduo em agir como o grupo está agindo — o efeito manada.
- **Autoridade:** Pessoas tendem a seguir figuras de autoridade. Partes interessadas mal-intencionadas exploram elementos como a popularidade de indivíduos para o grupo, como celebridades e influenciadores digitais. Outros exemplos são pessoas públicas, especialistas, pessoas que são referência nas atividades que realizam, como médicos, cientistas etc. Esse princípio é potencializado por aspectos ligados a incerteza e a ignorância, como a escolaridade, a experiência etc.
- **Comprometimento:** Pessoas querem que seus compromissos sejam consistentes com os seus valores. Comprometimento tem base no que as pessoas estão fazendo, e no que as outras pessoas pensam que ela está fazendo. Assim, partes interessadas mal-intencionadas exploram aspectos como a coerência, a fluência, domínio de um assunto, a expectativa da audiência de ter suas necessidades atendidas, ou defender as ideologias do grupo para gerar afinidade e comprometimento.
- **Escassez:** A tendência das pessoas de temerem perder algo ser pior que a possibilidade de deixar de ganhar. Algo que se torna (ou pode se tornar em breve) escasso ou indisponível, tem mais valor do que quando era abundante — como as corridas por papel higiênico no início da pandemia (Gagnani, 2020), a inexistência de uma vacina/conhecimento sobre o vírus, ou um determinado grupo que luta por direitos. Agentes mal-intencionadas exploram temas como a perda a liberdade (censura, restrições do direito de ir e vir etc.).
- **Reciprocidade:** É a expectativa inerente de quando se recebe algo com valor positivo, um sentimento de dívida surge. Explorada por mensagens ou informações não solicitadas, que geram uma ligação com a necessidade de responder e se engajar, ou atender a pedidos posteriores. Esse princípio é usado para construir um senso de consenso. Alguns elementos explorados são a exposição seletiva, percepção seletiva, e controle do fluxo de informações.
- **Unidade:** A tendência de buscar nos outros, traços de identidade que nos definem, e levam a formação de grupos com indivíduos ideologicamente alinhados. É um princípio que deriva da interação social e da criação de laços de confiança entre indivíduos em

grupos. Uma parte interessada mal-intencionada explora aspectos tribais da sociedade, como gírias, sentimentos de familiaridade, proximidade, ou "nós" contra "eles".

Os Princípios da Influência estão relacionados a fatores humanos, sociais e organizacionais. Eles são baseados (não somente) em vieses cognitivos — erros sistemáticos nas estratégias que seres humanos usam para fazer sentido da realidade — que influenciam a forma como os estímulos da realidade são interpretados (Cialdini, 1993). Esse é um aspecto fundamental do fenômeno da desinformação digital, os vieses cognitivos auxiliam na caracterização de aspectos que influenciam as pessoas a reagirem a uma desinformação de uma certa maneira. Além dos vieses cognitivos, os Princípios de Influência estão relacionados com as falácias lógicas — são raciocínios falsos que parecem verdadeiros, geralmente induzindo conclusões na argumentação (Cialdini, 1993; Pohl e Pohl, 2004). As falácias lógicas podem ser identificadas no conteúdo da desinformação, e são indícios do enquadramento usado para organizar uma ideia e se relacionam com os vieses cognitivos. Os quadros com algumas falácias lógicas e vieses cognitivos são encontrados nos Apêndice A.1.6.2.

Ao analisar a ocorrência de desinformação na comunicação, os Princípios da Influência servem como eixos complementares para análise sociotécnica de desinformação. Eles devem ser vistos como dispositivos que podem ser empregados ou não pelas partes interessadas para criar seus enquadramentos quando interagem através de TIC's. Esses elementos embasam o estudo do fenômeno no sentido de reduzir o viés de interpretação do analista ao observar o caso de desinformação, e mitigar os impactos do "achismo" em relação a um determinado aspecto observado. Além disso, eles amparam as discussões sobre os posicionamentos epistêmicos do analista, relacionados a interpretação dos resultados da análise sociotécnica de casos de desinformação digital.

## 2.4 GLOSSÁRIO

A Tabela 2.1 apresenta um resumo dos conceitos abordados na fundamentação teórica.

Tabela 2.1: Conceitos dos fundamentos teóricos

| Termo               | Conceito                                                                                                                                                                                                            | Teoria                                      |
|---------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------|
| <i>Affordance</i>   | Repertórios de comportamentos que os agentes podem desenvolver em relação ao meio que experimentam. <i>Affordances</i> podem ser físicas, como um botão, ou sociais, como um valor tido como aceitável em um grupo. | Teoria das <i>Affordances</i>               |
| Enquadramento       | O conjunto de aspectos sociotécnicos que influenciam ("amparam o significado") como um parte interessada observa um objeto digital de desinformação em uma determinada situação.                                    | Análise de Enquadramento                    |
| Desinformação       | Uso de signos que não correspondem a um estado real das coisas no ambiente. Pode ser acidental (sem intenção de enganar) ou intencional (com um propósito de enganar a audiência).                                  | Semiótica de Eco & Semiótica Organizacional |
| Informação          | Entidade subjetiva que pode ter diferentes interpretações de acordo com as circunstâncias.                                                                                                                          | Semiótica Organizacional                    |
| Objeto digital      | Um conjunto de atributos físicos, perceptivos, avaliativos, emocionais, intencionais, etc.                                                                                                                          | Multimodalidade                             |
| Organizações        | Uma organização é um conjunto de indivíduos que compartilham consensos (padrões de significados atrelados aos signos) que permitem que as pessoas se comportem de maneira organizada (Gazendam e Liu, 2006).        | Semiótica Organizacional                    |
| Partes interessadas | São os participantes do fenômeno da desinformação.                                                                                                                                                                  | Semiótica Organizacional                    |

## 2.5 EXPLICAÇÕES E DESDOBRAMENTOS TEÓRICOS

Como articulado por Harrison et al. (2011), entendo que para avaliar uma pesquisa precisa-se saber não somente quais métodos foram usados, mas como e porque foram usados — "precisamos saber como o conhecimento foi gerado" (Harrison et al., 2011). Nesse sentido, é preciso considerar alguns aspectos teórico-filosóficos desafiadores de se trabalhar com o fenômeno da desinformação digital, especialmente no que diz respeito a complexidade de abordar pontos como a subjetividade na interpretação do que é verdade e do que é mentira (a epistemologia do ponto de vista), as transformações de significado dentro de comunidades, e a adoção de novas crenças — pontos que apesar de não serem tratados diretamente pela Dismiss, acredito sere fundamental estarem presentes no desenvolvimento de uma ferramenta epistêmica que aborda a natureza sociotécnica do fenômeno da desinformação. Assim, discuto como alguns conceitos e perspectivas das teorias utilizadas na Dismiss foram adaptados para o estudo do fenômeno da desinformação e como podem ser explorados em trabalhos futuros para estudar outros aspectos do fenômeno não abordados pela Dismiss.

A Tabela 2.2 apresenta um resumo de como foram abordados os desafios e quais são os desdobramentos teóricos.

Tabela 2.2: Conceitos dos fundamentos teóricos

| Conceito                                     | Descrição                                                                                                                                                                                                                                                                                                                                                                                                                       | Apêndice                 |
|----------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|
| <i>Subjetividade</i>                         | Nessa tese, a natureza da realidade é considerada como subjetivista extrema. A verdade é a negociação coletiva de perspectivas individuais de diferentes observadores, que possuem diferentes capacidades, e que pertencem a uma determinada comunidade.                                                                                                                                                                        | A.1.1.1                  |
| Semiose e normas                             | De que modo é abordado a percepção, a significação, e o que rege o juízo de valor, a racionalidade, os estados mentais, as intenções, as atitudes, as invariantes comportamentais, etc.                                                                                                                                                                                                                                         | A.1.1.2 A.1.1.3<br>A.1.2 |
| Modelos de comunicação                       | Quais os desafios e limitações de se utilizar os modelos de comunicação empregados na Semiótica Organizacional, como a Teoria do Ato de Fala e a Lógica Deontica para estudar o fenômeno da desinformação.                                                                                                                                                                                                                      | A.1.3                    |
| Compatibilidade entre teorias                | Como a Semiótica Organizacional, a Teoria do Enquadramento, a Análise de Enquadramento, a Teoria da Persuasão e os Princípios de Influência são compatíveis.                                                                                                                                                                                                                                                                    | A.1.4 A.1.6.1            |
| Vulnerabilidades Humanas                     | Um desdobramento da fundamentação teórica que inter-relaciona Semiótica Organizacional, Semiose, Normas, Análise de Enquadramento, e a Teoria da Persuasão. Vulnerabilidades Humanas são invariantes-comportamentais que podem ser identificadas e exploradas ativamente por um agente mal intencionado para ganho próprio. Discute os fatores humanos que podem influenciar a suscetibilidade para consumir uma desinformação. | A.2.1                    |
| Categorização de estratégias de mitigação    | Um desdobramento do mapeamento sistemático. A mitigação envolve as atividades relacionadas ao combate da desinformação e as categorias de intervenções para reduzir seus efeitos.                                                                                                                                                                                                                                               | A.2.2.1                  |
| Aspectos sociotécnicos e suas inter-relações | Um desdobramento dos estudos em pequena escala realizados na fase de avaliação da tese em conjunto com os resultados do mapeamento sistemático. Apresenta a discussão da relevância de aspectos sociotécnicos em cada nível da Escada Semiótica e das suas inter-relações, indicando possíveis caminhos para avançar no entendimento do fenômeno.                                                                               | A.3.5.7                  |
| Mapeamento Sistemático                       | O relatório técnico do mapeamento sistemático. Apresenta o protocolo em detalhes, os dados levantados, e as discussões das questões de pesquisa. Resultados que levaram a construção do conceito de desinformação usado nessa tese, da necessidade de criar um catalogo de tipos de desinformação, e a motivação para criar o modelo do ciclo de vida da desinformação                                                          | A.3                      |

### 3 REVISÃO DA LITERATURA

A literatura na área de Ciência da Computação foi verificada em busca de um panorama das iniciativas para lidar com a desinformação digital, entender os desafios relacionados à sua natureza sociotécnica e oportunidades de pesquisa. Em maio de 2018, foi feita uma busca exploratória pelos termos "misinformation", "disinformation", "fake news", "false information" e "rumor" nas bases da IEEE, ACM e SCOPUS, que revelou mapeamentos desatualizados (dos 39 estudos secundários encontrados o mais atual era de 2016), e revisões sistemáticas de escopo limitado ou focadas em um tipo de tarefa de mitigação, não atendendo os objetivos desta pesquisa. Para obter um panorama atualizado e amplo das pesquisas, um protocolo de mapeamento sistemático da literatura foi elaborado e aplicado<sup>1</sup>. Os resultados (749 estudos primários) sintetizam as principais linhas de pesquisa e seus desafios, revelam problemas relacionados ao conceito de desinformação, apontam deficiências na abordagem das partes interessadas e destacam limitações ao tratar dos aspectos sociotécnicos do fenômeno.

#### 3.1 ESTUDO EXPLORATÓRIO

O estudo exploratório buscou por revisões da literatura e mapeamentos sistemáticos realizadas entre 2010 e 2020 para obter uma visão abrangente das pesquisas e dos desafios do fenômeno da desinformação digital. Foram encontradas 39 revisões da literatura sobre algum tipo de desinformação digital. Esses trabalhos serviram de referência para codificação, organização e discussão dos problemas de pesquisa, das soluções propostas, e suas limitações<sup>2</sup>. Com base nos resultados do estudo, as pesquisas podem ser agrupadas em quatro linhas (Almaliki, 2019b; Fernandez e Alani, 2018; Shu et al., 2017): Detecção, Validação, Dinâmica, e Gerenciamento. A Figura 3.1 mostra uma síntese das linhas de pesquisa na literatura.

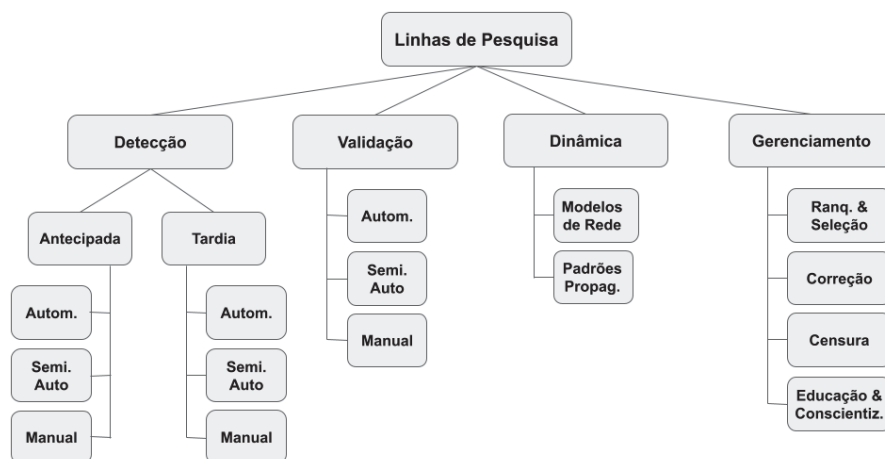


Figura 3.1: Linhas de pesquisa sobre desinformação na Computação.

As abordagens sobre Detecção se concentram na identificação de padrões que indiquem uma falsidade, se baseando na análise de conteúdo, perfil de usuários e análise de redes sociais. Os estudos de Validação focam na verificação da veracidade de mensagens, buscando fatos relacionados ao conteúdo avaliado. Os estudos de Dinâmica abordam a difusão de desinformação,

<sup>1</sup>O protocolo foi elaborado entre Junho e Agosto de 2018, foi aplicado em Dezembro de 2018 e finalizado em Dezembro de 2019. Uma atualização dos resultados foi feita até Julho de 2020.

<sup>2</sup>Detalhes desses trabalhos podem ser encontrados no Apêndice A.3.2.



como a caracterização da topologia de redes sociais e padrões de viralidade. Os estudos sobre Gerenciamento buscam desenvolver estratégias para mitigar a desinformação, como encontrar a melhor forma de apresentar correções e auxiliar a audiência a identificar falsidades.

As estratégias de Detecção podem ser automáticas, semi-automáticas e manuais, e se subdividem em detecção "precoce" e "tardia". Na detecção precoce investiga-se meios de identificar uma desinformação antes de ser amplamente difundida em redes sociais, empregando estratégias como a detecção automática de contas maliciosas (Pal e Chua, 2016), identificação de mensagens suspeitas com aprendizado supervisionado (Lozano et al., 2020) e definição de critérios para atividades suspeitas (Reuter et al., 2017). Na detecção tardia, busca-se detectar desinformações virais, empregando estratégias como o uso de mecanismos manuais para verificar uma desinformação (Saquete et al., 2020), o rastreamento da origem de uma desinformação (Almaliki, 2019b), e métricas para avaliação de credibilidade (Viviani e Pasi, 2017).

As estratégias de Validação empregam métodos de avaliação de veracidade, com abordagens automáticas, semi-automáticas e manuais. Por exemplo, definição de critérios de confiabilidade (Kumar e Geethakumari, 2014), validação de meta-dados (Bajaj et al., 2016), e validação de fatos (Khandelwal e Kumar, 2020). As estratégias automáticas (Detecção e Validação) mais exploradas são as que empregam técnicas de aprendizado de máquina e processamento de linguagem natural (Bondielli e Marcelloni, 2019).

As pesquisas sobre Dinâmica investigam aspectos da topologia das redes sociais e a identificação de padrões de propagação (Rana et al., 2019). Os estudos sobre Dinâmica se aprofundam na caracterização da estrutura de redes sociais e do estágio de difusão (Fernandez e Alani, 2018), sendo predominante as pesquisas em tarefas de controle (identificação, rastreamento e bloqueio) (Almaliki, 2019b; Shelke e Attar, 2019).

As abordagens de Gerenciamento empregam as estratégias de Detecção, Validação e Difusão para mitigação da desinformação. Estas se subdividem em Classificação & Seleção, Correções, Censura & Controle e, Educação & Conscientização. Abordagens de Classificação & Seleção empregam estratégias de Detecção e Validação para identificar falsidades e mensagens suspeitas (Manzoor e Nikita, 2019), e avaliar a prioridade de mensagens para validação manual (Viviani e Pasi, 2017). Abordagens de Correções empregam estratégias de Validação e propagação de correções, como checadores de fatos (Pal e Loke, 2019), rótulos (Gao et al., 2018) e abordagens baseadas em multidões (Kim et al., 2018). Abordagens de Censura & Controle empregam estratégias de Detecção, Validação e Dinâmica, como detecção e bloqueio de contas maliciosas (Khaled et al., 2018) e mitigação de difusão por limitação de compartilhamentos (Melo et al., 2019). Abordagens de Educação empregam estratégias de Validação e Correção, como campanhas de conscientização e educação sobre hábitos de verificação (Soler e Busilo, 2019).

### *3.1.0.1 Problemas e Desafios*

As revisões da literatura apontam diversas limitações nas linhas de pesquisa. A Tabela 3.1 apresenta um resumo das abordagens de cada linha e as limitações reportadas pelas revisões.

A efetividade das abordagens de Detecção depende de bases de dados rotulados e modelos de desinformação pouco abrangentes quanto aos aspectos multidisciplinares, sendo este um desafio para pesquisas da área (Almaliki, 2019b). Al-Sarem et al. (2019) convocam a comunidade para investigar e integrar aspectos que vão além daqueles relacionados à linguística. Fernandez e Alani (2018) afirmam que soluções de detecção fornecem alertas sem justificativa, geralmente desvinculados do contexto da audiência, além de serem pouco populares e pouco usadas. As soluções de detecção automática têm o potencial de serem enganadas por desinformação deliberada (Fernandez e Alani, 2018). Ainda, Al-Sarem et al. (2019) articulam que a qualidade técnica das bases de dados representa outro obstáculo, com dados



Tabela 3.1: Desafios encontrados na literatura

| Linha         | Complexidade                                                                                                                      | Multidisciplinaridade                                                                                                                                                                                                                                                                                                                                                         |
|---------------|-----------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Detecção      | Foco em tarefas pequenas (Fernandez e Alani, 2018);                                                                               | Tendem a considerar aspectos técnicos do fenômeno (Fernandez e Alani, 2018);<br>Focam em aspectos linguísticos (Saquete et al., 2020);<br>Abordam conjuntos de aspectos sociotécnicos pre-determinados (Jindal et al., 2020);<br>Atributos de credibilidade oriundos de modelos genéricos, propostos com base em outras mídias (TV e jornal impresso) (Saquete et al., 2020); |
|               | Poucos tipos de dados (Al-Sarem et al., 2019);                                                                                    |                                                                                                                                                                                                                                                                                                                                                                               |
|               | Acesso restrito a bases de dados rotulados (Zhang e Ghorbani, 2020);                                                              |                                                                                                                                                                                                                                                                                                                                                                               |
|               | Classificadores binários não dão conta da variedade de tipos de desinformação (Zhang e Ghorbani, 2020);                           |                                                                                                                                                                                                                                                                                                                                                                               |
|               | Pobreza metodológica das pesquisas (Lozano et al., 2020);<br>Adaptabilidade limitada à temática do conteúdo (Habib et al., 2019); |                                                                                                                                                                                                                                                                                                                                                                               |
| Validação     | Dependência de dados verificados e fontes confiáveis (Al-Sarem et al., 2019);                                                     | Tendem a focar em aspectos técnicos (Fernandez e Alani, 2018);<br>Focam em aspectos linguísticos (Saquete et al., 2020);<br>Atributos de credibilidade oriundos de modelos genéricos, propostos com base em outras mídias (TV e jornal impresso) (Saquete et al., 2020);                                                                                                      |
|               | Correções distantes do usuário (Fernandez e Alani, 2018);                                                                         |                                                                                                                                                                                                                                                                                                                                                                               |
|               | Acesso restrito a bases de dados rotulados (Zhang e Ghorbani, 2020);                                                              |                                                                                                                                                                                                                                                                                                                                                                               |
|               | Tendência de focar na detecção e não validação <i>per se</i> (Al-Sarem et al., 2019);                                             |                                                                                                                                                                                                                                                                                                                                                                               |
|               | Escopo limitado à temática (Al-Sarem et al., 2019);                                                                               |                                                                                                                                                                                                                                                                                                                                                                               |
| Dinâmica      | Predomina o foco na topologia (Rana et al., 2019);                                                                                | Tendem a focar em aspectos técnicos (Fernandez e Alani, 2018);<br>Aspectos sociais focam em atributos sociodemográficos (Rana et al., 2019);<br>Focam em motivações e relacionamentos de partes interessadas (Fernandez e Alani, 2018);                                                                                                                                       |
|               | Modelos inspirados na biologia (propagação viral) são pouco representativos do contexto (Fernandez e Alani, 2018);                |                                                                                                                                                                                                                                                                                                                                                                               |
| Gerenciamento | Foco em gerar e difundir correções (Fernandez e Alani, 2018);                                                                     | Tendem a focar em aspectos técnicos (Fernandez e Alani, 2018);<br>Poucas abordagens centradas no usuário (Fernandez e Alani, 2018);<br>Mais estudos exploratórios sobre fatores humanos de influência, como vieses cognitivos (Al-Sarem et al., 2019);                                                                                                                        |
|               | Estratégias fragmentadas, tendo pouca integração das soluções (Caulfield et al., 2019);                                           |                                                                                                                                                                                                                                                                                                                                                                               |
|               | Predominam tarefas de controle (identificação e censura) (Shelke e Attar, 2019);                                                  |                                                                                                                                                                                                                                                                                                                                                                               |

enviesados e pouca representatividade de contextos. Habib et al. (2019) apontam a necessidade de caracterizações do fenômeno que considerem o contexto e mais pesquisas qualitativas sobre os impactos nos usuários de redes sociais. Os autores afirmam que há poucos conjuntos de dados de temáticas diversas em saúde, educação e política, além de poucos estudos sobre desinformação deliberada (Habib et al., 2019). Lozano et al. (2020) criticam a pobreza de informações metodológicas nos artigos sobre avaliação e detecção automática, além de empregarem técnicas ultrapassadas, com poucos estudos de aprendizado profundo ou de transferência.

As estratégias de Validação apresentam as mesmas dificuldades encontradas pelos estudos em Detecção, como bases de dados pouco representativas e enviesadas, além dos estudos dedicados à validação não trabalharem diretamente com esse problema, tendendo a focar na

detecção (Al-Sarem et al., 2019). Reuter et al. (2017) indicam a necessidade de investigações sobre a veracidade e credibilidade, de mensagens e usuários, em diferentes plataformas de comunicação de modo integrado, sendo um desafio a identificação de aspectos representativos da desinformação e suas transformações em diferentes mídias de comunicação (*cross-platform*). Em relação às abordagens manuais, a literatura aponta a dificuldade de checadores de fatos e abordagens baseadas em multidões de lidar com o volume de dados crescente (Pal e Chua, 2016). Além disso, as validações acontecem desconectadas do local onde o usuário consome informações, sendo pouco acessadas ou difundidas (Fernandez e Alani, 2018). Ainda, a literatura cita fatores humanos e sociais que dificultam a efetividade de correções, como relações afetivas, desconfiança e vieses cognitivo, e aponta a necessidade de ampliar a investigação da influência de fatores humanos e sociais (Al-Sarem et al., 2019; Fernandez e Alani, 2018).

As pesquisas sobre Dinâmica tendem a considerar aspectos técnicos, e se baseiam em modelos inspirados na biologia (propagação viral) e modelos matemáticos. Embora funcionais, estes modelos são pouco representativos da complexidade contextual da audiência, analisando somente tendências virais, e não abordando a topologia da rede em profundidade (Rana et al., 2019). Além disso, trabalhos de difusão tendem a não considerar o comportamento do usuário, focando em dados sociodemográficos (idade, sexo, localização geográfica), não caracterizando a difusão *per se* (Fernandez e Alani, 2018). A literatura aponta a necessidade de mais estudos que caracterizem a difusão de desinformações não virais, além de abordar contextos reduzidos e a análise de tráfegos anômalos (Rana et al., 2019). Ainda, Fernandez e Alani (2018) indicam a necessidade mais pesquisas sobre as partes interessadas e como estão interessadas no fenômeno.

Os estudos de Gerenciamento apresentam uma tendência de se concentrar nos aspectos técnicos, como a classificação de desinformações de modo automático ou o controle de difusão (censura) (Fernandez e Alani, 2018). Almaliki (2019b) aponta uma demanda por novas ideias de mitigação, considerando o contexto da audiência e fomentando sua colaboração. Fernandez e Alani (2018) indicam a necessidade de mais abordagens sociotécnicas para ampliar o conhecimento de fatores de impacto inter-relacionados. Lozano et al. (2020) sugerem uma demanda por mais estudos multidisciplinares e qualitativos para aprimorar a capacidade de caracterizar o contexto social da audiência e informar aspectos relevantes para detecção. Além disso, Fernandez e Alani (2018) apontam a falta de envolvimento de empresas, governos e usuários como a principal limitação das estratégias de mitigação. Caulfield et al. (2019) citam a necessidade de estratégias que integrem diferentes dimensões do fenômeno para aprimorar a efetividade, como soluções legislativas, mecanismos automáticos de apoio a detecção, e abordagens de validação baseadas em multidões incorporadas em uma organização unificada.

Em resumo, a literatura reporta a falta de consenso nos conceitos sobre desinformação e veracidade (Lozano et al., 2020; Fard e Cunningham, 2019; Almaliki, 2019b), além de criticar a falta de conferências e revistas dedicadas para o tema (Fard e Cunningham, 2019; Almaliki, 2019b)<sup>3</sup>. As abordagens de detecção, validação e dinâmica tendem a ser pouco abrangentes, investigando pequenas tarefas com apenas um tipo de dado oriundo de uma fonte principal, e focadas na difusão e consumo de desinformação (Lozano et al., 2020; Fernandez e Alani, 2018). Existe a necessidade de mais abordagens que considerem aspectos técnicos, humanos, sociais e organizacionais de maneira integrada (Habib et al., 2019; Fernandez e Alani, 2018). Ainda, a literatura apresenta metodologias para auxiliar no desenvolvimento de métricas, critérios e modelos usados na detecção automática que são pouco adaptativos a mudanças no contexto e abordam as partes interessadas genericamente (Fernandez e Alani, 2018).

<sup>3</sup>Através de nosso mapeamento sistemático da literatura, apresentamos duas causas para os problemas de consenso taxonômico e a discussão de uma possível solução no Apêndice A.3.3

### 3.1.0.2 *Discussão*

Os trabalhos relacionados mostram a natureza complexa do fenômeno da desinformação digital, representada pela diversidade de fatores de impacto em diferentes níveis de abstração que são abordados pelos estudos encontrados. A análise revelou uma predominância de estudos que buscam elaborar soluções para um contexto específico, com uma tendência em adotar uma perspectiva técnica do fenômeno, abordando a detecção e validação de desinformação em sua maioria no idioma inglês. Existe um desafio metodológico para abordagens abrangentes do fenômeno que limita o avanço do conhecimento a respeito de sua natureza fundamentalmente sociotécnica e impacta na capacidade de resolução de problemas das intervenções propostas. Não encontramos nas revisões, abordagens que instrumentalizem a análise de desinformações e apoiem o entendimento de elementos em diferentes dimensões a serem considerados ao se pensar em intervenções. Assim, defendemos que antes de desenvolver soluções especializadas é preciso entender os elementos que constituem o problema tratado, reconhecer os fatores relevantes em cada contexto e organizar o conhecimento gerado para habilitar o desenvolvimento de intervenções efetivas e eficientes.

Existe a necessidade de investigações que entendam como ocorre a desinformação pela perspectiva de diferentes partes interessadas, olhando aspectos subjetivos do fenômeno de modo estruturado. A tendência reportada nas revisões em abordar o fenômeno de modo fragmentado leva a perda de detalhes inter-relacionados, e ao desenvolvimento de intervenções desintegradas, segmentadas em diferentes dimensões do problema complexo, sem sinergia. Por exemplo, quem constrói uma desinformação deliberada, a constrói para um público-alvo, indicando uma relação entre a criação e o consumo. Nesse sentido, existe a necessidade de compreender a natureza sistêmica do fenômeno da desinformação para investigar intervenções e integrar estratégias em diferentes estágios considerando diferentes níveis de abstração de aspectos sociotécnicos relevantes. Ainda, não encontramos uma visão geral ou categorizações para os tipos de estratégias de mitigação, ou como, quando e onde são utilizadas. Isto pode facilitar a integração de estratégias em diferentes níveis e possibilitar melhores condições de mitigar a desinformação e seus efeitos prejudiciais.

Como as revisões reportam, a literatura não apresenta um consenso para a taxonomia da desinformação digital e seus diferentes tipos (Lozano et al., 2020). As divergências nos conceitos para o termo "desinformação" prejudicam a obtenção de um panorama conciso das contribuições e desafios de pesquisa. Além disso, a maioria das revisões e mapeamentos focam em tarefas específicas para mitigação (como a detecção e validação) (Lozano et al., 2020), ou consideram um escopo temporal que limita a abrangência dos trabalhos selecionados e restringem o entendimento dos avanços no conhecimento sobre o fenômeno. Considerando as limitações das revisões analisadas, é necessário realizar um mapeamento abrangente da literatura para obter uma visão complementar das contribuições e desafios do fenômeno na área de Ciências de Computação, considerando as limitações das revisões analisadas. Por isso, realizamos um mapeamento sistemático da literatura para compreender em profundidade como aspectos sociotécnicos e as partes interessadas são abordados e quais são os tipos de desinformação digital estudados e suas características.

## 3.2 MAPEAMENTO SISTEMÁTICO DA LITERATURA

O mapeamento<sup>4</sup> foi proposto para obter um panorama atualizado e abrangente das pesquisas sobre o fenômeno da desinformação digital pela perspectiva da área de Ciências da

<sup>4</sup>O protocolo do mapeamento, detalhes sobre a aplicação e um relatório técnico completo detalhando os resultados podem ser encontrados no Apêndice A.3.

Computação no contexto das bibliotecas digitais da ACM, Scopus e IEEE. A chave de busca foi baseada nas palavras-chaves encontradas no estudo exploratório (Quadro 3.2). O mapeamento cobriu o período de 1993<sup>5</sup> até 2020.

Tabela 3.2: Chave-base de pesquisa

---

("disinformation" OR "misinformation" OR "fake news" OR  
 "false news" OR "fabricated news" OR "hoax" OR "rumor"  
 OR "false information" OR "fake information" OR  
 "fabricated information")

---

Com base nas limitações observadas no estudo exploratório foram elaboradas quatro Questões de Pesquisa (QP):

- QP01 - Quais são os tipos de desinformação estudados?
  - O estudo exploratório revelou divergências nos conceitos de "*disinformation*" e "*misinformation*" e na caracterização dos tipos de desinformação. O objetivo dessa questão de pesquisa é compreender os problemas com os conceitos de desinformação, identificar os tipos de desinformação digital (sátira, notícia falsa, etc.) e os aspectos sociotécnicos que os caracterizam.
- QP02 - Quais e como são abordadas as partes interessadas?
  - O estudo exploratório mostrou uma tendência em se abordar as partes interessadas óbvias (e.g. locutor e audiência), deixando outros atores relevantes de fora (e.g. controladores de plataformas digitais, prestadores de serviço, governos). O objetivo dessa questão de pesquisa é identificar as partes interessadas abordadas e a metodologia usada para identificá-las e caracterizá-las.
- QP03 - Como as pesquisas abordam os aspectos sociotécnicos?
  - No estudo exploratório, observamos a tendência de se abordar aspectos técnicos da desinformação digital e dificuldades de se considerar aspectos humanos e sociais de modo integrado. O objetivo dessa questão de pesquisa é identificar a abrangência dos modelos e caracterizações da desinformação digital considerados pelas soluções e como os aspectos sociotécnicos são abordados.
- QP04 - Como as pesquisas abordam a estrutura de um caso de desinformação digital?
  - O estudo exploratório revelou a tendência de se abordar o fenômeno da desinformação com um foco no consumo e na difusão de desinformação. O objetivo dessa questão de pesquisa é identificar os elementos de um caso de desinformação, como os estudos abordam a comunicação e categorizar os modelos e os estágios do ciclo de vida da desinformação.

A chave de pesquisa retornou 4591 artigos. Desses, foram selecionados 788 estudos (749 estudos primários e os 39 estudos secundários do estudo exploratório) para responder às QPs, usando um método baseado em Petersen et al. (2015). A seleção envolveu 3 pesquisadores

---

<sup>5</sup>O ano de nascimento do NCSA Mosaic que impulsionou a popularização da Web (Schatz e Hardin, 1994). O autor entende isso como sendo um marco para questões sociais do fenômeno da desinformação no meio digital.

do Laboratório de Interação Humano-Computador da UFPR e um processo de filtragem de 3 estágios baseado em critérios de seleção e exclusão. Os dados extraídos foram organizados em planilhas do Google Sheets para as fases de filtragem<sup>6</sup> e a extração de dados<sup>7</sup>. A Tabela A.8 resume os dados da seleção.

Tabela 3.3: Estatísticas do mapeamento sistemático.

| Estatísticas          | Processado | Revisão | Total      |
|-----------------------|------------|---------|------------|
| Todas as artigos      |            |         | 4591       |
| 1º filtro (excluídos) | -1776      | -1733   | -3059      |
| 2º filtro (excluídos) | -275       | -19     | -294       |
| Dados extraídos       |            |         | <b>788</b> |

### 3.2.1 QP01 - Quais são os tipos de desinformação estudados?

De 788 artigos selecionados em nossa revisão, 64 estudos definiram o que entendiam por *Misinformation* e/ou *Disinformation*. Um total de 486 artigos indicavam de forma implícita o seu entendimento sobre o tipo de informação falsa com que trabalhavam, referenciando outros estudos. Outros 238 artigos expressaram explicitamente uma definição para pelo menos uma forma de informação falsa, como sátiras e notícias falsificadas. Os resultados do mapeamento mostram a inconsistência na terminologia, com ambiguidade nas definições de *misinformation* e *disinforamtion*, o que atrapalha o entendimento do objeto de estudo dos artigos e torna a revisão da literatura morosa e improdutiva<sup>8</sup>.

#### 3.2.1.1 Tipos de desinformação

O mapeamento sobre os tipos de desinformações revelou 24 termos usados para diferenciar falsidades com pequenas particularidades. A Figura 3.2 apresenta um mapa dos tipos de desinformação encontrados nos artigos do mapeamento. Os resultados da análise indicam termos iguais usados com entendimentos diferentes, além de definições ambíguas. Por exemplo, em Khan et al. (2019) notícias falsificadas são sátiras, e em Ishida e Kuraya (2018) são farsas.

<sup>6</sup>[https://docs.google.com/spreadsheets/d/1QqHWVMC\\_CMMwLePvYGf8miM7NlM-CxYz6Z1ttDi494Q/](https://docs.google.com/spreadsheets/d/1QqHWVMC_CMMwLePvYGf8miM7NlM-CxYz6Z1ttDi494Q/)

<sup>7</sup><https://docs.google.com/spreadsheets/d/1lZORTRirkr-kKYzZXC9gV8lip3d3lfsVG7X8nY7RLbg/>

<sup>8</sup>Por isso, essa tese define um entendimento para a desinformação digital e seus tipos adotando uma perspectiva sociotécnica. Detalhes sobre esse ponto estão no Apêndice A.3.3.



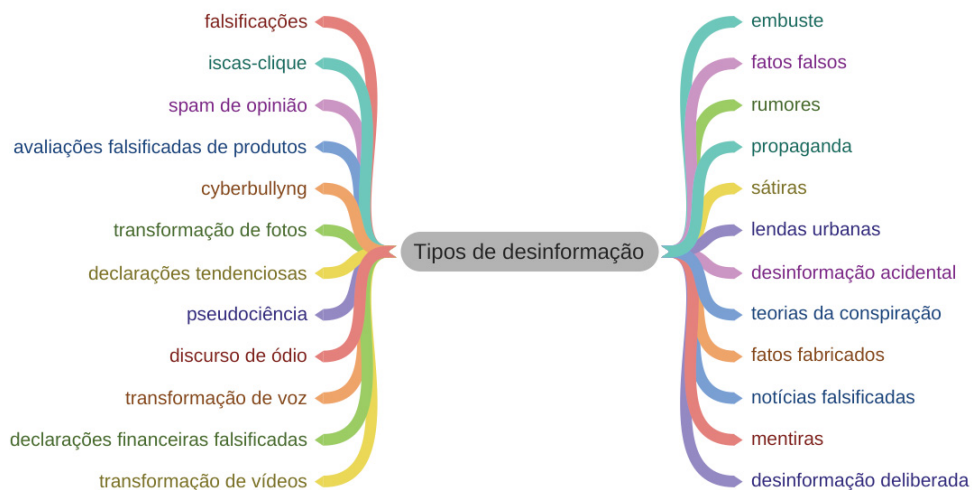


Figura 3.2: Tipos de desinformação encontrados no mapeamento.

Quanto aos atributos para os tipos de desinformação, o mapeamento encontrou 6 artigos dedicados ao estudo da taxonomia e tipologia de informações falsas (Fard e Cunningham, 2019; Lemieux e Smith, 2018; Shu et al., 2017; Kumar et al., 2016b; Zhou e Zhang, 2004; Tudjman e Mikelic, 2003). Estas pesquisas propõem uma diversidade de critérios para categorizar os tipos de desinformação. Ao todo encontramos 7 critérios usados para categorizar a desinformação que são apresentados na Tabela A.10<sup>9</sup>.

Tabela 3.4: Critérios de categorização encontrados na literatura.

| Critério         | Descrição                                                         |
|------------------|-------------------------------------------------------------------|
| Evidência        | se é baseada em evidências ou opiniões;                           |
| Significância    | é sobre o tópico de alguma urgência ou a importância percebida;   |
| Sobre indivíduos | se é sobre uma pessoa ou não ;                                    |
| Intenção         | propósito da desinformação, sendo com ou sem intenção de enganar; |
| Veracidade       | se é verdadeira ou falsa;                                         |
| Função           | o objetivo da mensagem (magoar, explicar, manipular, etc.);       |
| Motivação        | ganho monetário, pessoal, político, ou outro tipo de ganho;       |

Os critérios mais usados nos estudos mapeados são (artigos podem ter considerado mais de um critério): evidência (352), intencionalidade (263) ganho/motivação (41). A categorização de desinformação quanto a Evidência diferencia a desinformação em dois tipos, aquelas com base em fatos e aqueles baseados em opiniões. Estas são avaliadas com base na qualidade da mensagem em relação a atributos como veracidade (Kumar e Geethakumari, 2013), autenticidade (Shu et al., 2018), e confiabilidade (Fard e Cunningham, 2019) e significância (Fard e Cunningham, 2019). Quanto ao critério Intencionalidade, a desinformação pode ser acidental (*misinformation*) — quando oriunda de erro — ou deliberada (*disinformation*) — construída/usada com a intenção de enganar (Shu et al., 2018). Quanto ao Ganho, a literatura reporta o lucro financeiro (Shu et al., 2019; Buchanan e Benson, 2019; Seo et al., 2019), ganho político (Moreno e Bressan, 2019; Shu et al., 2019; Varol e Uluturk, 2018), e ganho pessoal (Buchanan e Benson, 2019; Seo et al., 2019).

<sup>9</sup>Os resultados e discussões sobre taxonomia e como a literatura aborda os diferentes critérios pode ser encontrado nos Apêndices A.3.3.1



Sobre os tópicos dos conteúdos, encontramos pesquisas sobre teorias da conspiração (Flintham et al., 2018; Volkova et al., 2017), propaganda política (Al-Rawi et al., 2019; Bedard e Schoenthaler, 2018; Karadzhov et al., 2018), famosos (Al-Rawi et al., 2019; Bedard e Schoenthaler, 2018); militar, industrial e acadêmico, (Sethi, 2017; Granik e Mesyura, 2017), finanças (Bedard e Schoenthaler, 2018), crimes (Volkova et al., 2017), preconceito e agressões (Bessi et al., 2015; Ferrara, 2015), fatos interessantes, dicas e truques (Karadzhov et al., 2018), avaliações de serviços (Sandifer et al., 2017) e produtos (Yao et al., 2017; Ratkiewicz et al., 2011), atentados, crises e catástrofes naturais (Rajdev e Lee, 2015; Oh et al., 2010; Gupta et al., 2013), e saúde (Chomphoosang et al., 2012; Hailun et al., 2014).

A Tabela A.13 resume os problemas com os principais tipos de desinformação<sup>10</sup>.

Tabela 3.5: Problemas de concisão no entendimento dos tipos de desinformação

| Tipo                  | Entendimento                                                                                   | Exemplos                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|-----------------------|------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Rumores               | Ambiguidade de entendimentos quanto aos critérios Intencionalidade e Evidência.                | <b>Informação não verificada</b> (Habib et al., 2019; Lee e Choi, 2018; Patel et al., 2017; Metaxas et al., 2015); <b>Que pode ser verdadeira ou falsa</b> (Rana et al., 2019; Buntain e Golbeck, 2017); <b>Informação falsa</b> (Qin et al., 2018; Chen et al., 2017; Zhang et al., 2015; Wang et al., 2014; Qazvinian et al., 2011); <b>Propaganda falsa</b> (Tan et al., 2019).                                                                                                                                                             |
| Farsas                | Concisão quanto a Intencionalidade.                                                            | <b>Informação enganosa e maliciosa</b> usada para enganar e manipular pessoas (Yuliani et al., 2018; Hui et al., 2018; Prasetyo et al., 2017); <b>Conteúdo anti-social</b> intencional, como difamação e <i>bullying</i> (Yu et al., 2018);                                                                                                                                                                                                                                                                                                    |
| Notícias Falsificadas | Concisão de entendimentos quanto aos critérios Intencionalidade e Evidência.                   | <b>Qualquer forma de informação falsa</b> , de farsas à sátiras (Karduni et al., 2018); Conteúdo <b>intencionalmente criado para enganar</b> (Habib et al., 2019; Della Vedova et al., 2018; Armano et al., 2018; Al-Ash e Wibowo, 2018; Figueira e Oliveira, 2017); Histórias postadas como <b>falso fato aceito como genuíno</b> (Parikh e Atrey, 2018; Gilda, 2017); <b>Declarações tendenciosas</b> (Murungi et al., 2018); <b>Fatos alternativos</b> sem base na realidade (Purnomo et al., 2017). <b>Sátiras</b> (Chandra et al., 2017). |
| Sátiras               | Ambiguidade de entendimentos quanto aos critérios Intencionalidade e o teor cômico.            | Sub-tipos: paródias, piadas e pegadinhas (Khan et al., 2019; Popat et al., 2017); <b>Desinformação acidental</b> (Cybenko e Cybenko, 2018; Ishida e Kuraya, 2018); <b>Desinformação deliberada</b> (Bedard e Schoenthaler, 2018); Histórias para <b>entretenimento</b> (Karduni et al., 2018; Campan et al., 2017).                                                                                                                                                                                                                            |
| Teoria da conspiração | Ambiguidade de entendimentos quanto ao critério Intencionalidade. Concisão quanto a Evidência. | São <b>fabricações intencionais</b> (Tacchini et al., 2017); <b>informações não confiáveis</b> para explicar eventos ou circunstâncias (Glenski et al., 2018); informações falsas, tanto deliberadas quanto acidentais, que reduzem a complexidade de eventos sociais (Bessi et al., 2015);                                                                                                                                                                                                                                                    |

<sup>10</sup>Os resultados completos podem ser encontrados nos Apêndices A.3.4

### 3.2.1.2 Discussão

A desinformação tem sido uma preocupação pelo menos desde o desenvolvimento e popularização da Web (Luciano, 1996). Embora algumas discussões sobre informações falsas datem de 1998 (Resnik, 1998), e trabalhos definindo uma taxonomia datem do início dos anos 2000 (Zhou e Zhang, 2004), não há acordo da comunidade acadêmica da Ciência da Computação sobre os tipos de informações falsas (Pang e Ng, 2017; Jiang e Wilson, 2018)<sup>11</sup>. Os resultados mostram uma diversidade de definições para os mesmos tipos de desinformação, por exemplo, "sátira" sendo intencional (Khan et al., 2019), e acidental (Cybenko e Cybenko, 2018).

O significado de alguns critérios das taxonomias são pouco claros. Por exemplo, o critério Função, que apresenta vários tipos de valores (magoar, explicar a realidade, explicar um sentimento, manipular, etc.), mas que não são discutidos, conceituados ou apresentados extensivamente. Assim, considerar o critério Função pode levar a ambiguidade na categorização dos tipos de desinformação oriunda da interpretação subjetiva de quem utilizar a taxonomia proposta. Especialmente quando consideramos que a Função de um tipo de desinformação depende de fatores circunstanciais, como o significado de ações, atitudes, propósitos e motivações, para diferentes partes interessadas de diferentes culturas. Por exemplo na desinformação deliberada, quando o estado mental de quem expressa um determinado tipo de desinformação (Intencionalidade) difere do que é expresso no conteúdo da mensagem (Função), e explora o estado mental da audiência (Intencionalidade percebida). Isso indica que as taxonomias são limitadas ao considerar a perspectiva subjetiva do fenômeno da desinformação em relação as partes interessadas que interagem com ela, sejam os pesquisadores e projetistas, seja o agente que desinforma e sua a audiência.

Cada tipo de desinformação possui diferentes atributos, também tratados na literatura como "multimodalidades"(sons, texto, imagens, vídeos, tamanho, fonte do texto, cores, etc.) (Jin et al., 2017). Seguindo a natureza sociotécnica do fenômeno, os tipos de desinformação têm atributos técnicos específicos e aspectos sociais que induzem determinado tipo de interação evocando atitudes e ações (*Affordances* do Objeto). Por exemplo, iscas-clique empregam elementos para chamar a atenção da audiência e monetizar os "cliques", apelando à frases ambíguas, e imagens, títulos e subtítulos sensacionalistas (Zannettou et al., 2019). Por isso, essa tese entende que um tipo de desinformação é um objeto digital com um conjunto de atributos físicos, humanos e sociais inter-relacionados. Dessa forma, estudar os tipos de desinformação é importante para compreender a extensão de seus atributos e como eles podem afetar a interação com a audiência, e caracterizar os impactos com base na natureza sociotécnica do fenômeno. No sentido de apoiar o estudo sociotécnico do fenômeno, reunimos as principais características encontradas nos trabalhos da literatura e apresentamos um entendimento para os principais tipos de desinformação com base na SO (Tabela A.14)<sup>12</sup>.

<sup>11</sup>Por esse motivo definimos um entendimento para "desinformação"adequado ao propósito da pesquisa (Capítulo 2).

<sup>12</sup>Uma versão refinada é empregada como material de apoio na DISMISS.

Tabela 3.6: Tipos comuns de desinformação

| Tipo                   | Intencionalidade                   | Descrição                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|------------------------|------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Rumores                | Informação não verificada          | É uma mensagem que pode ou não estar sendo usada para desinformar a audiência (Devi e Karthika, 2018). Fica implícito para as partes interessadas que a mensagem não é verificada pela fonte (a fonte não se responsabiliza pelo conteúdo). Geralmente referência a terceiros (pessoas, organizações, etc.) como autoridade certificadora da informação e requer checagem (Rana et al., 2019). Pode ser revelada como verdade ou falsidade posteriormente.                                                                                                                                                                                                                        |
| Sátiras                | Desinformação deliberada           | São mensagens usadas deliberadamente para explorar aspectos cognitivos (entendimento, memória, raciocínio, etc.) das partes interessadas, sendo informações falsas de conotação cômica empregadas para entretenimento da audiência (Campan et al., 2017). A intenção de enganar tem o objetivo implícito de entreter comicamente, onde a audiência está consciente disso ou tem "pistas" que indicam o teor cômico. Podem ser paródias, piadas e pegadinhas (Khan et al., 2019; Popat et al., 2017).                                                                                                                                                                              |
| Farsas<br>(hoaxes)     | Desinformação deliberada           | São mensagens falsas usadas deliberadamente para explorar vulnerabilidades nas partes interessadas-alvos. Geralmente exploram crenças, expectativas, interesses, desejos, necessidade e as intenções de um público-alvo (Yuliani et al., 2018). Podem fazer parte de histórias complexas, como teorias da conspiração (Santoso et al., 2017), bem contextualizadas, como esquemas de <i>phishing</i> e <i>spams</i> (Hui et al., 2018), fabricadas, como golpes de estelionato e difamação (Goolsby et al., 2013), esquemas avançados de enganação, como campanhas desinformativas e ameaças persistentes (Ahmad et al., 2019) ou um simples ato oportunístico (Yu et al., 2018). |
| Notícias Falsificadas  | Desinformação deliberada           | São um subtipo de farsas, fabricadas para enganar a audiência, manipular crenças e atitudes (Wijaya e Santoso, 2018). Elaboradas no formato textual ou multimídia. Possuem propriedades semelhantes às notícias veiculadas por autoridades certificadas, como jornais e portais de notícias (Manzoor e Nikita, 2019). A audiência não está ciente da intencionalidade da fonte da notícia falsa. Podem ser totalmente fabricadas ou uma introdução de viés (Murungi et al., 2018).                                                                                                                                                                                                |
| Teorias da Conspiração | Desinformação deliberada/acidental | Geralmente são farsas que tendem a reduzir a complexidade da realidade explorando o campo de informação da parte interessada ou comunidade, considerando aspectos sociais ou políticos significativos como sendo tramas maquiavélicas, muitas vezes concebidas por indivíduos ou organizações poderosas (Bessi et al., 2015). Também podem ser oriundas de distorções cognitivas que estão relacionadas ao conhecimento de uma parte interessada ou comunidade, como a confiança cega no julgamento de uma autoridade (Acemoglu e Ozdaglar, 2011). Tendem a formar comunidades em torno do tema com indivíduos que podem ou não estar cientes da farsa.                           |
| Isca-clique            | Desinformação deliberada           | São representações usadas com a intenção de prender a atenção da audiência e obter acessos. Frequentemente emprega linguagem sensacionalista e apelo emocional (Glenski et al., 2018). O objetivo da fonte geralmente está ligado ao ganho monetário (Zannettou et al., 2019).                                                                                                                                                                                                                                                                                                                                                                                                    |

### 3.2.2 QP02 - Quais as partes interessadas consideradas pelas pesquisas?

A caracterização das partes interessadas é um aspecto importante do estudo do fenômeno da desinformação que tem sido abordado de modo limitado pela literatura revisada. As pesquisas apresentam uma tendência de caracterizar as partes interessadas de modo genérico (136 artigos), com classificações como receptor e emissor, interlocutor e audiência, produtor e consumidor, ou uma referência ao papel de "usuário de rede social", porém com pouca profundidade (Resende et al., 2018; Traylor et al., 2017; Hassan et al., 2017). De modo geral, são duas formas com que a literatura aborda as partes interessadas<sup>13</sup>:

<sup>13</sup>Uma discussão detalhada se encontra nos Apêndices A.3.5.8

- Modelos teóricos: incluem estudos que empregam características de perfil e fatores humanos, atributos e papéis das partes interessadas de outras disciplinas, como sociologia, psicologia, e antropologia (Saquete et al., 2020). Por exemplo, Zhou e Zhang (2007) apresenta um modelo ontológico da comunicação de desinformação deliberada, e considera as partes interessadas como sendo: contribuinte, curador, leitor, administrador, e analista com base na literatura de sociologia que estuda fraudes;
- Evidências empíricas: inclui estudos que investigam fatores humanos e sociais do fenômeno da desinformação para caracterizar as partes interessadas (Almaliki, 2019a). Por exemplo, estudos que buscam entender os critérios que determinados grupos de pessoas usam para avaliar a credibilidade de informações/fontes de informações em diferentes meios de comunicação (Piccolo et al., 2020; Saquete et al., 2020).

Encontramos 42 papéis para as partes interessadas envolvidas na comunicação de desinformação, reportados em 236 artigos, oriundas de estudos empíricos, modelos e tipologias do fenômeno. A Figura 3.3 apresenta as partes interessadas classificadas usando o Diagrama de Identificação de Partes Interessadas<sup>14</sup> (DPI) - um artefato da SO usado para mapear partes interessadas em contextos de atividades organizadas.

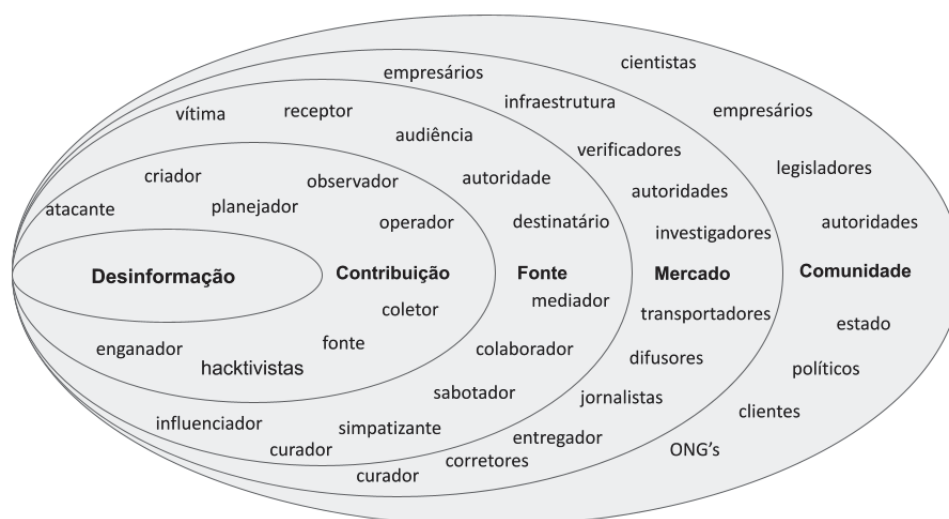


Figura 3.3: Mapa de partes interessadas encontradas no mapeamento.

Em relação a metodologia para estudar as partes interessadas, predominam estudo exploratórios que fazem uso de entrevistas e questionários abordando aspectos sociodemográficos, como sexo, idade, e escolaridade (Zhang et al., 2018; Torres et al., 2018b; Fernandez e Alani, 2018), valores (Piccolo et al., 2020), motivações de compartilhamento (Chin e Zauddin, 2019), entre outros. Outros estudos empregam técnicas etnográficas para analisar os comportamentos como, ações de compartilhamento, verificação e consumo de desinformação em grupos específicos, como estudantes e idosos, em contextos culturais diferentes (Wang e Fussell, 2020; Chin e Zauddin, 2019; Wandoko et al., 2019; Wason et al., 2019; Torres et al., 2018b), ou investigando os impactos da homofilia e proximidade afetiva (Wu et al., 2018). Ainda, algumas abordagens

<sup>14</sup>O grupo Contribuição considera agentes envolvidos na concepção da desinformação. Fonte os agentes que atuam indiretamente, como audiência sendo fonte de informações. O Mercado são os parceiros e concorrentes, como jornalistas e plataformas de comunicação. O grupo Comunidade considera agentes afetados indiretamente (Kolkman, 1993). O DPI é abordado em mais detalhes no Capítulo 2.

empregam métodos e técnicas da Psicologia, como investigações dos padrões de consumo de desinformação baseadas na personalidade (Halbach et al., 2019) e, diferenças no comportamento de avaliação de mensagens objetivas e com linguagem emocional (Flintham et al., 2018). As pesquisas que abordam as partes interessadas indiretamente utilizam metodologias de investigação complexas, usando métodos e técnicas em conjunto, como em (Dang et al., 2016b) que usa análise de rede social, análise visual, análise de conteúdo e mineração de texto, para classificar papéis de usuários.

### 3.2.2.1 *Discussão*

As pesquisas baseadas em evidências empíricas buscam entendimento de aspectos situados do fenômeno e empregam métodos e técnicas de investigação qualitativa, como questionários e grupos focais, para caracterizar as partes interessadas. Embora elas apresentem avanços para o entendimento do fenômeno, os resultados destes estudos exploratórios dificilmente são generalizáveis devido a fatores informais como cultura, crenças, e valores, e do tamanho das amostras dos estudos (Lozano et al., 2020). Além disso, aspectos circunstâncias de influência, como a pressão social e autoridades em grupos, são difíceis de investigar e variam conforme o conteúdo varia (Pal et al., 2019; Barfar, 2019). As caracterizações dos atributos das partes interessadas adotadas por modelos do fenômeno são genéricas e pouco representativas (Lozano et al., 2020).

Nas pesquisas baseadas em Modelos Teóricos, predominam pesquisas que se baseiam em aspectos técnicos para classificar o perfil do usuário (Fernandez e Alani, 2018). Por exemplo, os critérios utilizados para classificar a credibilidade de uma conta — a quantidade de compartilhamentos e postagens em uma janela de tempo (Zhang e Ghorbani, 2020) —, auxiliam a identificação de contas que se enquadram no limite estipulado para a quantidade "ideal" de compartilhamentos no período de tempo considerado (que geralmente são robôs que automatizam a difusão (Hwang et al., 2012)). Entretanto, esses critérios não servem para caracterizar a intencionalidade da criação e divulgação, qual a relação de um perfil com uma desinformação compartilhada, qual a relação da desinformação com o ambiente de uma parte interessada, ou para detectar robôs desenhados para evitar esse tipo de mecanismo (desinformação deliberada). Dessa forma, conjuntos de critérios pré-determinados podem introduzir viés na avaliação, sendo um fator que impacta no entendimento das partes interessadas que constituem o fenômeno.

Existe uma necessidade metodológica e de estrutura ao abordar as partes interessadas que constituem o fenômeno da desinformação. Modelos teóricos embasam a investigação fornecendo uma estrutura para guiar a análise das partes interessadas do fenômeno. Métodos empíricos permitem identificar indícios e mecânicas relevantes em uma determinada situação envolvendo determinadas partes interessadas. Embora estes sejam caminhos fundamentais para o desenvolvimento do conhecimento, tanto os mesmo modelos quanto o conhecimento derivado de evidências empíricas são pouco abrangentes e dificilmente generalizáveis (Saquete et al., 2020). Alguns deles introduzem viés na análise de aspectos sociotécnicos, pressupondo conjuntos de aspectos e papéis para as partes interessadas pouco representativos do contexto que caracterizam. Outros generalizam fatores de um contexto como significativos para todo o fenômeno. Por isso, um desafio de pesquisa se encontra na modelagem e caracterização situada dos participantes, em quais aspectos são relevante em uma determinada situação e um determinado contexto para se considerar em estratégias de mitigação, e quais são formas adequadas de se abordar os participantes e caracteriza-los.

Não encontramos propostas que apoiem a identificação de partes interessadas para além das óbvias ou que auxiliem na caracterização das partes interessadas para além dos atributos pré-determinados pelos modelos teóricos. Embora estudos exploratórios busquem evidências



empíricas para esclarecer o impacto de aspectos como valores, crenças e emoções, existe a necessidade de apoiar a identificação destes aspectos em relação ao contexto e ao conteúdo de uma desinformação em diferentes situações. Antes disso, existe a necessidade de apoiar a identificação de diferentes partes interessadas e como estão interessadas no fenômeno, considerando o contexto em que ocorre a interação com a desinformação. Por isso, um caso de desinformação precisa ser estudado considerando as circunstâncias de cada situação<sup>15</sup>.

### 3.2.3 QP03 - Como as pesquisas abordam aspectos sociotécnicos do fenômeno?

Os 788 estudos foram analisados em busca de detalhes sobre o estudo sociotécnico do fenômeno. Os trabalhos foram analisados usando o método de análise de conteúdo (Lazar et al., 2017) juntamente com Escada Semiótica (ES)<sup>16</sup>. A ES divide os aspectos do sistema de informação em subdisciplinas e foi usado para categorizar os estudos em relação à abrangência dos aspectos sociotécnicos considerados no estudo da desinformação. Os artigos foram categorizados com base no esquema da Figura 3.4<sup>17</sup>:

|                              |                                                                                                                                                                                                                                                                                    |             |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| Sistema de Informação Humana | <b>Mundo Social:</b> Abordagens que buscam compreender os aspectos sociais do fenômeno, como valores, a cultura, os impactos da desinformação na sociedade, os efeitos de vieses cognitivos de grupo (efeito manada, etc.), o impacto de relações afetivas e profissionais, etc;   | 57 artigos  |
|                              | <b>Pragmático:</b> Abordagens que buscam compreender as funções das mensagens, como as motivações de compartilhamento, aspectos de deliberação, difamação, infâmia, e a finalidade dos tipos de desinformação;                                                                     | 118 artigos |
|                              | <b>Semântico:</b> Abordagens que buscam compreender aspectos do significado/entendimento das informações, como a subjetividade da veracidade, análise sentimental, análise de posicionamento, ambiguidades e a relação do tema com o ambiente;                                     | 418 artigos |
| Plataforma Tecnológica       | <b>Sintático:</b> Abordagens que buscam compreender aspectos estruturais do fenômeno (conteúdo, difusão, organização, etc.), como a diagramação, elementos de ênfase, agrupamentos em redes virtuais, e identificação de gargalos na difusão;                                      | 187 artigos |
|                              | <b>Empírico:</b> Abordagens que buscam compreender aspectos da interação e dos meios comunicação, como as \textit{affordances} tecnológicas, sincronidade, direcionalidade, tempo de leitura, ruído, interferência, eficiência de classificadores, entre outros;                   | 6 artigos   |
|                              | <b>Mundo Físico:</b> Abordagens que buscam compreender aspectos físicos do fenômeno, como limitações de hardware, os custos de difusão, defeitos em dispositivos, e a localização geográfica dos participantes, modos de coleta de dados (e.g. imagem de câmera real ou deepfake); | 2 artigos   |

Figura 3.4: Esquema de categorização dos artigos.

Os resultados indicam uma tendência de pesquisas concentradas nos níveis clássicos da semiótica: Pragmático, Semântico e Sintático. Isso se reflete nas poucas pesquisas encontradas sobre os níveis Físico e Empírico. Nesse sentido, existe um entendimento limitado sobre as diferentes dimensões do fenômeno que tem o potencial de deixar aspectos relevantes ocultos. Por exemplo, considerando o critério Evidência (usado nas taxonomias de desinformação digital), observamos uma tendência em estratégias de Validação que consideram a autenticidade do conteúdo de uma mensagem com base em atributos de credibilidade da fonte de informações, ou se outras fontes corroboram o mesmo conteúdo (Lozano et al., 2020). Nesse sentido, existe espaço para estudos e análises sobre os aspectos Físicos e Empíricos, como a produção artificial

<sup>15</sup>Por esse motivo essa tese investiga como apoiar a identificação de partes interessadas, tendo adotado as ferramentas da SO para isso.

<sup>16</sup>Artefato da Semiótica Organizacional usado na análise de Sistemas de Informação (Stamper e Liu, 1994).

<sup>17</sup>O esquema detalhado da ES pode ser encontrado no Apêndice A.3.5



de imagens, se uma imagem foi coletada no mundo Físico por algum tipo de sensor (e.g. câmera), ou produzida por computador (e.g. deepfakes), e os inter-relacione com os impactos que possuem na percepção subjetiva de partes interessadas.

### 3.2.3.1 Aspectos sociotécnicos na literatura

Os resultados revelam maior concentração de estudos que abordam aspectos específicos em algum degrau da escada semiótica com pouca inter-relação. A maior quantidade de aspectos sociotécnicos abordados se encontram nos níveis Semântico (418 artigos) e Sintático (187 artigos). A literatura analisada apresenta um foco em conjuntos de características linguísticas técnicas usadas por classificadores, uma tendência em abordar o fenômeno fragmentado em estágios específicos sem considerar o todo, dificuldades em abordar aspectos do nível social e limitações na integração de aspectos sociotécnicos de diferentes níveis de abstração<sup>18</sup>.

Os estudos que abordam os aspectos sociotécnicos inter-relacionados os classificam em três níveis de abstração Zhang e Ghorbani (2020): aspectos do perfil do usuário, aspectos do conteúdo e aspectos sociais (Figura 3.5). Nessas categorias, são considerados dois tipos de aspectos: aspectos físicos e não físicos. Os aspectos físicos estão relacionados ao formato e ao meio de comunicação pelo qual a desinformação se propaga. Os aspectos não físicos são relacionados a fatores humanos, sociais e organizacionais, como opiniões, relacionamentos, autoridades, etc.

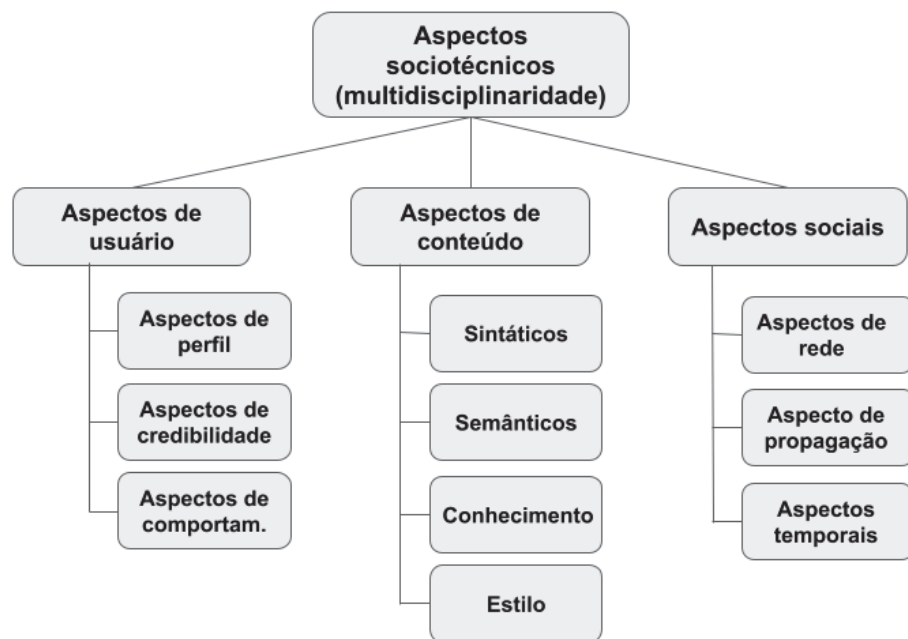


Figura 3.5: Aspectos sociotécnicos encontrados na literatura.

Aspecto de nível de usuários são atributos relacionados ao perfil de quem cria ou compartilha desinformações, e da audiência que consome. Se subdividem em:

- Atributos de perfil: como o nome, informações de geolocalização, os dados de cadastro do usuário, verificados ou não, tem descrição ou não, e assim por diante (Saquete et al., 2020; Lozano et al., 2020; Zhang e Ghorbani, 2020).

<sup>18</sup>As análises e discussões podem ser encontradas no Apêndice A.3.5.

- Aspectos de credibilidade: como a pontuação de credibilidade do usuário, o número de amigos e seguidores, a proporção entre amigos e seguidores, a número total de *posts* do usuário (Saquete et al., 2020; Lozano et al., 2020; Zhang e Ghorbani, 2020).
- Aspecto do comportamento: como a pontuação de anomalia do usuário, número de interação do usuário em uma janela de tempo, média mensal de *post* do usuário, etc (Saquete et al., 2020; Lozano et al., 2020; Zhang e Ghorbani, 2020).

Aspectos de nível de conteúdo são atributos relacionados ao conteúdo, ao formato e aspectos das plataformas de comunicação por onde se difundem.

- Aspectos linguísticos (Sintáticos): como os erros gramaticais sintáticos, número de n-gramas, frequência de termo (TF), estrutura de frases, etc;
- Aspectos semântico: como os erros gramaticais semânticos, títulos exagerados, consistência entre título e conteúdo, contradições, etc (Saquete et al., 2020; Lozano et al., 2020; Zhang e Ghorbani, 2020);
- Aspectos de conhecimento: como as bases de dados rotuladas, similaridade com conteúdos verificados, relação com do conteúdo com o ambiente da audiência, etc (Saquete et al., 2020; Lozano et al., 2020; Zhang e Ghorbani, 2020);
- Aspectos estilo: como a fração de *tweets*/postagens que contêm links externos, menções de usuários, *hashtags*, popularidade de nome de domínio, número de imagens ou vídeos, pontuação de clareza, pontuação de coerência, etc (Saquete et al., 2020; Lozano et al., 2020; Zhang e Ghorbani, 2020);

Aspectos de nível social são aspectos que refletem o padrão de difusão e a interação entre os usuários.

- Recursos baseados em rede: como o agrupamento de usuários semelhantes, localização, formação educacional, hábitos de consumo e compartilhamento (Shelke e Attar, 2019).
- Recursos baseados em distribuição: como a árvore de propagação, o grau da raiz em uma árvore, o número máximo de subárvores, o número de *retweets/reposts* para um *tweet* original, a fração de *tweets* que são retuitados por uma conta (Shelke e Attar, 2019);
- Recursos temporais: como o intervalo entre postagens, a frequência de postagem, respostas e comentários de uma conta, a hora do dia em que a informação original é postada/compartilhada/comentada, e o dia da semana da publicação (Shelke e Attar, 2019);

Foram encontrados 13 estudos sobre multimodalidade<sup>19</sup>. Predominam estudos sobre classificadores automáticos que consideram **aspectos de nível de conteúdo**, apresentando um foco em aspectos textuais (sintáticos e semânticos) inter-relacionados a aspectos visuais (imagens) (Jindal et al., 2020). Alguns trabalhos de multimodalidade abordam a relação de aspectos de nível de conteúdo, de usuário e social (Maigrot et al., 2018, 2016), considerando aspectos técnicos (como aspectos sintáticos) e sociais (como métricas de credibilidade).

<sup>19</sup>Multimodalidade considera a interação das partes interessadas com a desinformação digital como sendo "modos/modalidades" (Halliday, 2014). Exemplos de modos são, gestos, expressões faciais, imagens, textos, layout, etc. Dessa forma, estes trabalhos abordam diferentes aspectos sociotécnicos relacionados a interação.

Foram identificados 6 estudos que empregam lentes sociotécnicas. Destes encontramos a análise de um sistema sociotécnico de eleição (Caulfield et al., 2019), um estudo que adota as lentes sociotécnicas da Translucidez para promover comportamentos responsáveis de consumo e compartilhamento de informações (Wang et al., 2014), e um estudo sobre operações de informações estratégicas (Starbird et al., 2019) que adota as lentes teóricas da Estruturação, Sócio-materialidade e Cognição Distribuída para analisar casos de desinformação e identificar as táticas de comunicação. Um estudo que emprega o método de Análise de Enquadramento para identificar vieses em notícias de modo automático (Hamborg et al., 2017). Os outros estudos são revisões da literatura que abordam os avanços sociotécnicos na mitigação do fenômeno.

### 3.2.3.2 Discussões

Como encontrado nas revisões da literatura (Almaliki, 2019a; Fernandez e Alani, 2018), os estudos categorizados na ES se concentram em aspectos técnicos. Mesmo pesquisas que consideram aspectos do nível Social, Pragmático e Semântico, integram critérios de avaliação de conteúdo, da topologia, e da interação, mas não vão além de características técnicas e quantificáveis (Jindal et al., 2020). Por exemplo, Dongo et al. (2019) apresenta uma técnica para avaliação da credibilidade da fonte considerando o conteúdo, o usuário e o mundo social:

- credibilidade do texto: com base em atributos como uso de imperativo, ambiguidades e linguagem sentimental;
- credibilidade do usuário: com base em atributos como data de criação da conta e se a conta verificada;
- credibilidade social: baseada em atributos como quantidade seguidores, compartilhamentos e curtidas;

Nesse exemplo, atributos humanos e sociais são aspectos técnicos atribuídos aos indivíduos que têm relação com suas ações e características no ambiente virtual. Entretanto, não refletem aspectos subjetivos ou informais, como sua capacidade de percepção e interpretação, ou seus valores e crenças, e não são realmente significativos do indivíduo em seu contexto. Assim, as caracterizações parecem ser úteis apenas para identificar atividades suspeitas, mas não lidam bem com transformações, novidades, e fatores circunstanciais. Existe uma distância entre o conhecimento que pesquisas empíricas vêm construindo sobre os aspectos sociotécnicos, e as produções técnicas e conceituais de intervenções e modelagens para o fenômeno. Isso representa um indicio da dificuldade de identificar e inter-relacionar aspectos sociotécnicos em diferentes níveis de abstração, por exemplo entender como aspectos sintáticos e sociais podem influenciar um caso de desinformações.

As caracterizações de aspectos sociotécnicos do fenômeno tendem a focar no relacionamento entre grupos pré-determinados de atributos técnicos, humanos, e sociais oriundos de lentes teórico-metodológicas de áreas como a análise de redes sociais, psicologia, e sociologia. Isso faz com que as relações entre aspectos sociotécnicos sejam limitadas aos atributos comumente considerados nestas literaturas, restringindo a flexibilidade e a exploração de outras inter-relações. Por exemplo, as características textuais propostas pela literatura focam em aspectos técnicos e são superficiais (Zhang et al., 2019a): número de palavras, extensão do texto, ocorrência dos símbolos ? e !, *emoticons* felizes ou infelizes, pronomes na 1ª, 2ª ou 3ª pessoa, letras maiúsculas, palavras positivas e negativas, menções no Twitter, *hashtags*, urls e retuítes. Embora seja embasada teoricamente, essa abordagem acaba por pressupor aspectos relevantes para compreensão do fenômeno, em especial ela generaliza a forma como diferentes partes interessadas interpretam

e utilizam estes signos em diferentes contextos e com diferentes propósitos. Isso fragmenta o entendimento do fenômeno em aspectos quantificáveis, deixando aspectos qualitativos relevantes (como a subjetividade de interpretação de diferentes partes interessadas) negligenciados para esses modelos e caracterizações. Essas caracterizações acabam por enviesar a investigação do fenômeno e limitam a representatividade dos elementos que investigam para o desenvolvimento de uma solução. Isso representa um desafio de pesquisa relacionado a natureza sociotécnica do fenômeno e um indício da dificuldade de adaptação dessas soluções para lidar com as partes interessadas aos aspectos sociotécnicos dos diferentes contextos (Jindal et al., 2020).

As soluções que abordam a inter-relação de aspectos sociotécnicos, como a multimodalidade, requerem recursos explícitos extraídos artesanalmente do conteúdo e do contexto do evento em que a desinformação ocorre (Jindal et al., 2020). Isso implica no custo de recursos humanos na análise e em vulnerabilidades relacionadas ao viés do analista. Além disso, as abordagens que propõem a detecção automática com base em aspectos multimodais, são dependentes de bases de dados rotuladas ou requerem padrões já identificados em outros estudos para treinamento (Jin et al., 2017). Assim, a abrangência do estudo do fenômeno da desinformação fica limitada às tendências apresentadas nas pesquisas, metodologias e atributos, restringindo a exploração de novas perspectivas e características. Ainda, existe uma carência de ferramentas e metodologias que apoiem a análise situada do fenômeno para identificar aspectos sociotécnicos relevantes do contexto em que ocorre a desinformação, sendo uma limitação na exploração de aspectos para além dos apresentados na literatura analisada.

Os resultados mostram limitações na abordagem de aspectos sociotécnicos inter-relacionados, que juntamente com as limitações reportadas nos Trabalhos Relacionados indicam a dificuldade de se lidar com um fenômeno complexo como a desinformação digital. Nesse sentido, é preciso soluções com estratégias e mecanismos projetadas de modo integrado, que abordem os aspectos humanos, sociais e técnicos em diferentes níveis de abstração, para se avançar na capacidade de mitigar os efeitos prejudiciais do fenômeno.

### 3.2.4 QP04 - Como as pesquisas abordam a estrutura de um caso de desinformação digital?

De modo geral, existem duas formas de abordar o fenômeno (Zhang et al., 2019a): abordagens em nível de postagem — abordam um único *post* de mídia social em profundidade —, e abordagens em nível de evento/caso — abordam um grupo de postagens que constituem um evento/caso. Em relação ao estudo da comunicação, seja por postagem ou por evento, as pesquisas sobre o fenômeno apresentam quatro perspectivas principais para abordar o contexto da comunicação:

- Abordagens não-sistêmicas: são estudos que abordam um estágio particular da vida de uma desinformação, como estudos de detecção na difusão (Pourghomi et al., 2017), sobre o comportamento de consumo de notícias (Flintham et al., 2018), e processos de criação de notícias falsas (McCarthy, 2018);
- Abordagens sistêmicas da vida da desinformação deliberada: são estudos que abordam a vida de uma desinformação considerando estágios inter-relacionados com foco na perspectiva de quem constrói a desinformação, como os estágios de campanhas de desinformação e as estratégias de enganação (Starbird et al., 2019; Traylor et al., 2017);
- Abordagens sistêmicas da difusão: são estudos considerando estágios inter-relacionados adotando a perspectiva da difusão, como o modelo

surgimento-difusão-ressurgimento (Shin et al., 2018), surgimento-difusão-decaimento (Chua et al., 2016), ignorante-difusor-abafador (Nekovee et al., 2007) ou o modelo epidêmico suscetível-infectado-recuperado e variações (Cho et al., 2019);

- Abordagens da comunicação como um processo: são abordagens que modelam a comunicação como um processo (Transmissor-canal-receptor). Por exemplo, o modelo de Lasswell (Qiming et al., 2021), Shannon & Weaver (Zhou e Zhang, 2007), Westley & MacLean's (Vitkova et al., 2021), e a Teoria do Ato de Fala (Saha et al., 2020);

Na literatura analisada predominam abordagens não-sistêmicas, considerando um estágio particular da vida de uma desinformação (413 artigos sobre detecção na difusão ou consumo). Os resultados revelam 5 estágios da comunicação: Criação (Oehmichen et al., 2019), Difusão (Hui et al., 2018), Consumo (Starbird et al., 2019), Decaimento (Hui et al., 2018), e Transformações (Starbird et al., 2019; Shin et al., 2018; Webb et al., 2016). As pesquisas caracterizam cada estágio considerando aspectos como, as motivações das partes interessadas envolvidas na criação de uma desinformação (Jang et al., 2018), as atividades envolvidas na criação de desinformação (preparação (Oehmichen et al., 2019; Hussain et al., 2018), elaboração (McCarthy, 2018), divulgação do conteúdo falso (Boshmaf et al., 2013; Traylor et al., 2017; Raynal e Gaspard, 2010)), e a ética de intenções de usuários no compartilhamento de desinformações (Webb et al., 2016).

O Quadro 3.7 apresenta características de cada estágio<sup>20</sup>.

---

<sup>20</sup>As características e os estágios deram origem ao Modelo Conceitual do Ciclo de Vida da Desinformação Digital, apresentado no Capítulo 4.

Tabela 3.7: Exemplos de características dos estágios do fenômeno da desinformação.

| Estágio    | Características                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Criação    | Pode ser uma atividade colaborativa ou solitária. (Hussain et al., 2018);<br>Atividades colaborativas requerem planejamento e coordenação (Hussain et al., 2018);<br>Sub-atividades: preparação (Oehmichen et al., 2019), elaboração (McCarthy, 2018) e divulgação (Traylor et al., 2017; Boshmaf et al., 2013; Raynal e Gaspard, 2010);<br>Eventos como crises, catástrofes naturais e atentados terroristas são ligados ao surgimento e viralidade de desinformação (Oh et al., 2010);                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Difusão    | O formato da comunicação afeta a interação da audiência. (Khodabakhsh et al., 2018);<br>A comunicação pode ser síncrona ou assíncrona (Zhou e Zhang, 2007), bidirecional ou unidirecional (Neumann, 1996);<br>Meios assíncronos limitam o controle do interlocutor na comunicação (Kumar et al., 2016a);                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Consumo    | Aspectos da mensagem afetam a percepção de risco da audiência, como o tamanho da mensagem (Kumar et al., 2016b; Zannettou et al., 2019), se contém links (Aigner et al., 2017), se permite imagens (Taylor-Smith et al., 2018), a diagramação do conteúdo (Leite et al., 2019; Aigner et al., 2017), as cores (Taylor-Smith et al., 2018);<br>A diagramação da mensagem conduz a atenção do leitor (Shen et al., 2019);<br>Aspectos sociais como relacionamentos (autoridades, proximidade afetiva, etc.), e crenças (posicionamento político, religião, etc.) influenciam o consumo (Rath et al., 2019; Berghel, 2017a);<br>Comunidades podem possuir hábitos de consumo de informações pouco saudáveis, como consumir e compartilhar sem verificar (Kovic e Laissue, 2016);<br>Fatores humanos, como a memória, são suscetíveis a manipulação (Larson et al., 2019), ou explorar a cognição através de falácias lógicas e vieses cognitivos (Li e Santos Jr, 2012); |
| Reiteração | Após a divulgação inicial, a desinformação pode migrar tanto entre diferentes plataformas digitais de comunicação (Janicka et al., 2019; Del Vicario et al., 2017), quanto entre grupos diferentes na mesma plataforma (Melo et al., 2019);<br>As transformações podem ser tanto orgânicas, quando a audiência transforma a mensagem, quanto artificiais, quando <i>bots</i> recriam ou redirecionam o conteúdo (Starbird et al., 2019; Yu e Han, 2018);<br>Partes interessadas podem fazer ajustes para evitar detecção (Bandeli e Agarwal, 2018), ou promover mensagens já corrigidas em formato novo (Finn et al., 2015; Tang et al., 2017).                                                                                                                                                                                                                                                                                                                       |

### 3.2.4.1 Discussões

As pesquisas tendem a abordar o fenômeno da desinformação de modo segmentado com um escopo limitado a algum estágio da vida de uma desinformação. Por exemplo, os estudos de detecção de desinformação tendem a focar na difusão e consumo, considerando pouco ou nenhum aspecto da criação da desinformação. Nesse caso, a criação de uma desinformação visa um público-alvo (quem comunica algo, comunica para alguém, em especial quando desenformar deliberadamente). Dessa forma, nota-se um entendimento limitado do fenômeno que restringe a identificação de aspectos inter-relacionados nos estágios da vida de uma desinformação.

Embora as abordagens encontradas desenvolvam o conhecimento sobre o fenômeno, elas apresentam limitações para completude, abrangência e inter-relações dos aspectos sociotécnicos envolvidos em um caso de desinformação digital. Os modelos não são abrangentes para apoiar o estudo de desinformações deliberadas e acidentais <sup>21</sup>, partindo de pressupostos relacionados a intencionalidade. O ciclo de vida da desinformação deliberada foca na perspectiva do criador ou do consumidor, separadamente. Os modelos que abordam a difusão tendem a deixar as partes interessadas em segundo plano (Fernandez e Alani, 2018). Ainda, modelos que abordam a

<sup>21</sup>Um característica desafiadora do fenômeno é análise de intencionalidade e deliberação. Usuários frequentemente cometem erros e criam e espalham desinformação acidentalmente (Watney, 2018), o que representa um desafio para os classificadores, mais ainda para mitigação responsável da desinformação.



comunicação como um processo, apesar de considerarem a intencionalidade como um elemento crucial na constituição de uma mensagem (Fiske, 2010) (como a Teoria do Ato de Fala (Saha et al., 2020)), são limitadas por adotar a perspectiva do interlocutor como tendo o maior impacto no contexto para interpretação das mensagens, presumindo que a perspectiva do consumidor como sendo a mesma do locutor (Sbisà e Fabbri, 1980). Isso limita a capacidade do modelo de representar a realidade, especialmente da audiência que consome a desinformação, além de desconsiderar as partes interessadas que controlam o canal de comunicação<sup>22</sup>. Estas limitações representam barreiras para o entendimento de aspectos inter-relacionados nos estágios da vida de uma desinformação e para a "generalização"<sup>23</sup> dos modelos.

A forma como os modelos abordam o fenômeno limita a capacidade de analisar um caso de desinformação de modo abrangente (considerando aspectos sociotécnicos e a relação entre os estágios da vida de uma desinformação), e flexível (que permite explorar os diferentes aspectos sociotécnicos em diferentes níveis de abstração). Nota-se uma tendência em abordar a desinformação de modo desestruturado, focando na análise de atributos quantificáveis de um determinado estágio da vida de uma desinformação, ou de modo estruturado em modelos que generalizam suposições sobre aspectos qualitativos da desinformação. Dessa forma, é preciso entender a desinformação digital como um fenômeno sociotécnico de modo abrangente, onde aspectos sociais e técnicos estão inter-relacionados nos estágios da vida de uma desinformação.

### 3.3 AMEAÇAS À VALIDADE DO MAPEAMENTO

São considerados os seguintes tipos de validade, com base em (Petersen et al., 2015): validade descritiva, validade interpretativa e validade teórica. Com base na natureza geral deste trabalho, as ameaças de generalização (Petersen et al., 2015) foram consideradas como tendo um impacto mínimo sobre o estudo. Para reduzir as ameaças à validade descritiva, que é a forma como conduzimos a documentação, foram utilizados formulários de coleta de dados para apoiar o processo de seleção e extração de dados, com base em (Petersen et al., 2015; Wieringa et al., 2006). Eles foram projetados para manter o rigor na codificação e obter registros objetivos. No caso das ameaças à validade teórica, que é determinada por fatores de confusão como viés na seleção dos estudos (Petersen et al., 2015), adotamos um protocolo e estratégias de revisão após cada filtro. Os estudos não acessíveis pela Capes não foram incluídos. Assim, pesquisas relevantes podem ter sido desconsideradas. Sobre a validade interpretativa — credibilidade das interpretações (Petersen et al., 2015) —, somente o autor discute as conclusões podendo ter um impacto relevante. Entretanto, procurei registrar e esclarecer como derivei as conclusões, para aprimorar tanto minha percepção quanto capacidade quanto pesquisador.

### 3.4 CONSIDERAÇÕES FINAIS

Esse capítulo apresentou uma síntese dos resultados do mapeamento sistemático da literatura da Ciência da Computação que investiga a desinformação digital. A revisão mostra uma tendência em segmentar a análise do fenômeno, focando em estágios específicos do ciclo de vida da desinformação e pouco relacionados. Algumas pesquisas adotam modelos de comunicação restritivos, com um foco genérico da difusão como um processo, e que abordam a perspectiva de partes interessadas específicas (geralmente o consumidor ou o criador da desinformação). Ainda, abordar o fenômeno como sociotécnicos se mostra uma tarefa complexa, com uma gama de

<sup>22</sup>Uma discussão detalhada das diferenças entre os modelos de comunicação está no Apêndice A.1.3

<sup>23</sup>Com esse termo o autor referencia a capacidade de transferência do conhecimento gerado de (Oulasvirta e Hornbæk, 2016): "Número de usuários, tarefas e contextos para os quais a solução pode ser aplicada; contextos qualitativamente novos em que a solução pode ser aplicada."

elementos relevantes em diferentes níveis de abstração que requerem metodologias que integram diversos artefatos, métodos e teorias, e exigem conhecimento e experiência em técnicas de análise quantitativa e qualitativa. Ademais, a consideração de aspectos sociotécnicos inter-relacionados se limita a grupos pré-determinados de aspectos usados em outros estudos e parece existir dificuldades no relacionamento de aspectos em diferentes níveis de abstração. Essa tendência pouco estimula a identificação e estudo de novos tipos de aspectos e suas inter-relações. Além disso, poucos estudos propõe métodos para analisar um caso de desinformação de modo situado, buscando aspectos sociotécnicos ligados a possíveis impactos ou fatores de influência para um contexto bem definido.

Existe espaço para investigação de abordagens sociotécnicas e sistêmica que auxiliem a análise da desinformação, sobretudo considerando as partes interessadas como peça central fenômeno. Devido a complexidade do fenômeno é necessário apoiar os projetistas de intervenções a entenderem a natureza sociotécnica da desinformação digital. Por isso, esta tese propõe uma abordagem sociotécnica e sistêmica para análise de desinformação, instrumentalizando os projetistas interessados no fenômeno à analisarem casos de desinformação.

A Tabela 3.8 apresenta a comparação das principais formas de abordar o fenômeno e os fatores relevantes identificados e discutidos nos resultados do mapeamento.

Tabela 3.8: Comparação entre diferentes abordagens sociotécnicas do fenômeno da desinformação.

| Abordagens<br>Pontos de Comparação                     | Abordagens<br>sistêmicas<br>(Dhar et al., 2016) | Comunicação<br>como processo<br>(Madhusudan, 2003) | Ecossistema<br>(Zannettou et al., 2019) | Ontologia<br>(Zhou e Zhang, 2007) | Análise<br>sociotécnica<br>(Starbird et al., 2019) | Dismiss |
|--------------------------------------------------------|-------------------------------------------------|----------------------------------------------------|-----------------------------------------|-----------------------------------|----------------------------------------------------|---------|
| Apoia a análise de aspectos sociotécnicos              | abrangência limitada                            | abrangência limitada                               | abrangência limitada                    | abrangência limitada              | abrangência limitada                               | sim     |
| Apoia a identificação de inter-relações nos aspectos   | Foco em aspectos técnicos                       | Foco em aspectos técnicos                          | sim                                     | Foco em aspectos técnicos         | sim                                                | sim     |
| Aborda a estrutura de um caso com estágios integrados  | pouco representativa do contexto                | pouco representativa do contexto                   | sim                                     | pouco representativa do contexto  | sim                                                | sim     |
| Apoia a identificação de partes interessadas           | não                                             | não                                                | não                                     | pre determinada                   | não                                                | sim     |
| Aborda a subjetividade das partes interessadas         | não                                             | não                                                | não                                     | não                               | não                                                | sim     |
| Apoia a análise em profundidade e largura              | não                                             | não                                                | não                                     | não                               | não                                                | sim     |
| Apoia o estudo de desinformação deliberada e acidental | não                                             | não                                                | não                                     | não                               | não                                                | sim     |

## 4 DISMISS - UMA ABORDAGEM PARA ANÁLISE SOCIOTÉCNICA DA DESINFORMAÇÃO

Para apoiar o entendimento da natureza sociotécnica do fenômeno da desinformação é preciso abordar a estrutura de um caso de desinformação, identificar as partes interessadas, e analisar os aspectos sociotécnicos que caracterizam o contexto do caso. A Dismiss é uma abordagem inspirada na Semiótica Organizacional (SO) proposta para apoiar a análise sociotécnica de desinformação que utiliza os artefatos da Diagrama de Identificação de Partes Interessadas (DPI), a Escada Semiótica (ES), e o Modelo Conceitual do Ciclo de Vida da Desinformação Digital incorporados no método de Análise de Enquadramento. A abordagem consiste de duas estratégias que podem ser empregadas para realizar a análise de casos de desinformação: 1) a análise sociotécnica a nível de postagem/profundidade, e 2) a análise sociotécnica a nível de caso/largura. A Dismiss representa uma ferramenta epistêmica que possibilita diagnosticar casos de desinformação para identificar as partes interessadas e os aspectos sociotécnicos que caracterizam problemas de modo situado, sendo projetada para analisar a desinformação acidental e deliberada, além de permitir o estudo em profundidade e em largura de um caso de desinformação.

### 4.1 VISÃO GERAL

O propósito da Dismiss é apoiar o entendimento da desinformação como um fenômeno sociotécnico, auxiliando na compreensão da estrutura de um caso de desinformação e construindo um entendimento abrangente de aspectos que caracterizam o seu contexto para apoiar a identificação e entendimento de problemas relevantes de serem mitigados. Para isso, a Dismiss ajuda a identificar os elementos básicos que são as **partes interessadas**, os **canais de comunicação** e os **aspectos sociotécnicos**, e como podem estar relacionados em cada caso. A Dismiss foi projetada para ser aplicada na fase de entendimento do problema que antecede o desenvolvimento de uma estratégia de mitigação e solução.

A Dismiss possui um arquitetura composta por três componentes principais e um secundário (Figura 4.1). O primeiro componente principal é constituído pelos artefatos da SO, Diagrama de Identificação de Partes Interessadas (DPI) e Escada Semiótica (ES), sendo usados para caracterizar o contexto de um caso de desinformação digital. O segundo é o método de Análise de Enquadramento (AE) que inspirou a técnica de exploração do ciclo de vida da desinformação e as questões que guiam a análise de objetos desinformativos na Dismiss. O terceiro componente é o Modelo Conceitual do Ciclo de Vida da Desinformação que possibilita uma visão sistêmica do caso de desinformação e auxilia na exploração e caracterização abrangente do fenômeno da desinformação. O quarto componente (secundário) corresponde aos eixos complementares que sustentam a análise sociotécnica na forma de materiais de apoio opcionais (Princípios de Influência, falácias lógicas e vieses cognitivos). Cada um desses componentes busca lidar com um dos desafios da análise sociotécnica identificados no mapeamento sistemático da literatura: 1) análise fragmentada de aspectos sociotécnicos, 2) partes interessadas genéricas e 3) segmentação da estrutura de um caso de desinformação.

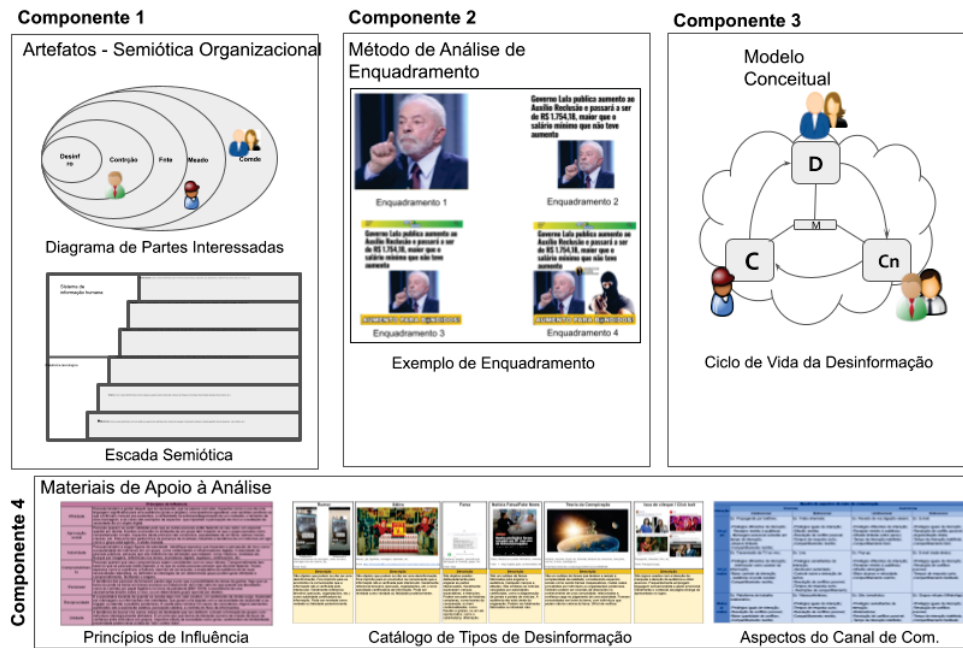


Figura 4.1: Componentes da Dismiss. Fonte: Autor.

A Dismiss foi desenvolvida para ser aplicada de forma flexível, considerando as duas estratégias de análise de casos de desinformação: Estratégia de Análise em profundidade e Estratégia de Análise em largura. A primeira é inspirada no que a literatura aborda como "análise de postagem" (Zhang et al., 2019a). Ela contempla o uso dos artefatos da SO para análise do contexto do caso e os passos da AE para análise do objeto desinformativo no contexto onde aparece, com objetivo de analisar um único objeto desinformativo em um único canal de comunicação. Trata-se de uma Análise de Enquadramento situada em uma única mídia de comunicação, estruturada e informada pelos artefatos da SO.

A segunda estratégia consiste da exploração do caso de desinformação pelas lentes do Ciclo de Vida da Desinformação Digital, expandindo a caracterização do contexto com os artefatos da SO e possibilitando a Análise de Enquadramento em vários objetos desinformativos e suas transformações que aparecem em diferentes meios de comunicação. Ela permite identificar padrões no conteúdo de diferentes objetos desinformativos, estudar as transformações da desinformação, identificar a abrangência da difusão de uma desinformação em um caso, obter uma perspectiva da estrutura organizacional do caso e seu ecossistema, e as extensões dos impactos da desinformação. Ambas as estratégias permitem construir um entendimento do caso de desinformação, quais as circunstância da criação, difusão e consumo, e quais as consequências da desinformação digital no caso analisado.

A primeira estratégia foi elaborada para simplificar a análise sociotécnica da desinformação digital, ao custo da qualidade da caracterização do contexto de um caso de desinformação. A segunda estratégia tem um maior custo de esforço do analista para aplicação (em relação a primeira estratégia), com a vantagem de obter uma visão abrangente do caso, possibilitando identificar problemas relacionados a criação, difusão e consumo de desinformação. A seguir são descritos as etapas da Dismiss e as duas estratégias são explicadas e exemplificadas.

#### 4.1.1 Estrutura da Dismiss

A Dismiss é dividida em três etapas inspiradas nos Métodos de Articulação de Problemas — Um conjunto de métodos que podem ser aplicados na fase inicial de um projeto quando se

depara com um problema vago e complexo para apoiar o entendimento de problemas (Kolkman, 1993)<sup>1</sup>. A Figura 4.2 apresenta uma representação dessas três etapas:

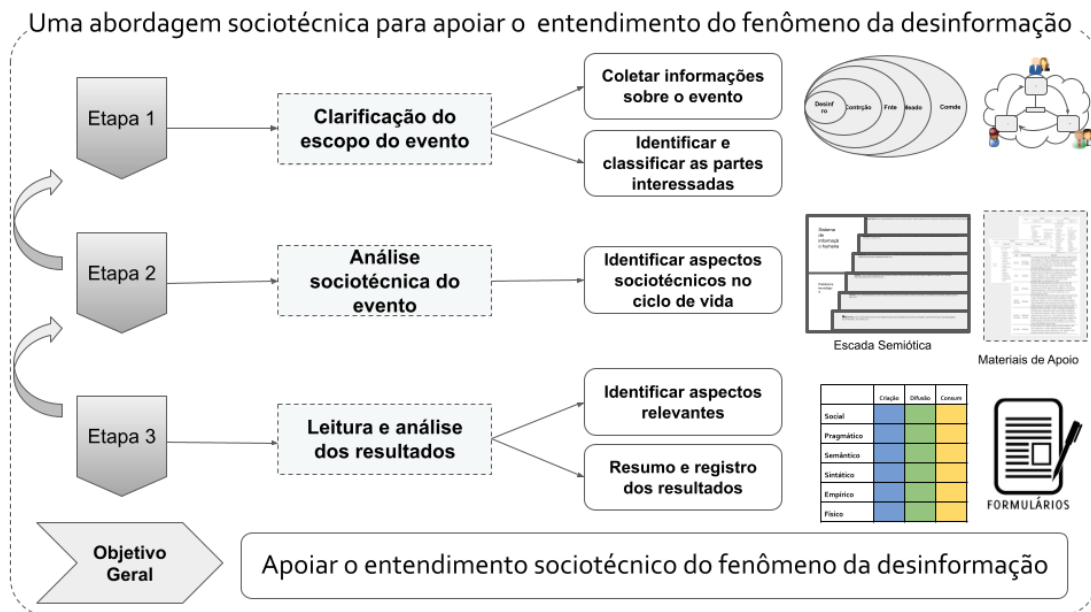


Figura 4.2: Etapas da Dismiss. Fonte: Autor.

1. Familiarização: essa etapa aborda a familiarização do analista com o contexto do caso de desinformação. O analista explora o contexto para compreender a estrutura do caso e identificar os elementos básicos que o constituem: as partes interessadas, os canais de comunicação e os objetos desinformativos. Para isso, o analista utiliza o Diagrama de Partes Interessadas (DPI) e perguntas-guia<sup>2</sup>. Essa etapa aborda os desafios revelados no mapeamento relacionados ao entendimento da estrutura de um caso, a identificação das partes interessadas de modo abrangente e suas atividades, e como elas se relacionam entre si, com os canais de comunicação e com o objeto desinformativo.
2. Análise sociotécnica: essa etapa aprofunda o entendimento do analista sobre o caso, apoiando a identificação de aspectos sociotécnicos considerando os elementos básicos identificados na etapa anterior. Para isso, o analista usa a Escada Semiótica e materiais de apoio<sup>3</sup>. Essa etapa aborda os desafios revelados no mapeamento relacionados ao levantamento de aspectos sociais, humanos e técnicos em diferentes níveis de abstração, considerando os estágios da vida de um objeto desinformativo que estruturam o caso.
3. Identificação e análise de problemas: essa etapa contempla a revisão dos dados identificados nas Etapas 1 e 2, onde são analisadas inter-relações de aspectos sociotécnicos e sua relevância no caso. Para isso, o analista utiliza uma adaptação da Escada Semiótica, perguntas-guia e os formulários propostos para registro dos dados. Essa etapa aborda os desafios relacionados a identificação de inter-relações entre aspectos sociotécnicos em diferentes níveis de abstração e em cada estágio do ciclo de vida da desinformação.

<sup>1</sup>O termo etapas usado nesse capítulo refere-se sempre ao passo a passo de aplicação da Dismiss. O termo estágio refere-se sempre ao Ciclo de Vida da Desinformação (criação, difusão e consumo)

<sup>2</sup>Esse termo se refere a perguntas desenvolvidas com base nos resultados do mapeamento e inspiradas na AE para guiar a identificação dos elementos básicos de um caso de desinformação.

<sup>3</sup>Materiais desenvolvidos para apoiar a identificação de aspectos sociotécnicos relevantes ao fenômeno da desinformação digital.



#### 4.1.2 Etapa 1 - Familiarização

A Etapa 1 envolve a exploração do contexto do caso de desinformação, as partes interessadas e as suas características. As atividades dessa etapa envolvem a coleta de informações para delimitar o escopo temporal de análise, identificar e categorização das partes interessadas, o(s) canal(ais) de comunicação, e o(s) objeto(s) desinformativo(s). Nessa etapa, o analista observa o contexto do caso para coletar informações sobre os elementos básicos do caso com a ajuda de artefatos e técnicas de análise. O analista usa o artefato DPI da SO adaptado para o estudo do fenômeno da desinformação digital e os materiais de apoio opcionais: Formulários de Registro (Formulários de Registro de Casos - para o registro das informações gerais do caso e o Formulários de Registro de Partes Interessadas (FRPI) - para armazenar os dados sobre as partes interessadas). Ao final dessa etapa, o analista terá um registro das partes interessadas, quais objetos desinformativos e canais fazem parte do caso, e quais serão analisados na Etapa 2.

##### 4.1.2.1 Coleta de informações

A coleta de informações visa familiarizar o analista com o contexto do caso e delimitar o escopo da análise. Primeiramente, o analista deve abordar o caso a partir do canal de comunicação em que o objeto desinformativo aparece. A partir do objeto, o analista deve identificar o tipo de objeto, a temática, onde o objeto aparece (o complexo de elementos do canal de comunicação: origem, destino, canal de comunicação, plataformas digitais de comunicação e dispositivos), qual o escopo temporal do caso (início e fim do caso) e as partes interessadas, registrando os dados no Formulário de Registro de Caso (FRC) (Tabela 4.1).

Tabela 4.1: Formulário de apoio: registro de caso de desinformação. Fonte: Autor.

| Formulário de Registro de Caso                 |              |                   |             |
|------------------------------------------------|--------------|-------------------|-------------|
| Título do caso:                                |              | Temática do caso: |             |
| Data do caso:                                  |              | Tipo de objeto:   |             |
| Complexo de elementos do canal de comunicação: | Dispositivos | Tecnologias       | Plataformas |
| Descrição do caso:                             |              |                   |             |
| Fontes:                                        |              |                   |             |

Para auxiliar na exploração do caso, o analista deve responder às seguintes perguntas-guia:

- Qual o tipo do objeto desinformativo?
  - O tipo do objeto implica em algumas características específicas para interação. Cada tipo de objeto está ligado a um propósito e aspectos que induzem um certo tipo de interação (*affordances*). Por exemplo, notícias falsas são farsas que geralmente tem objetivo de ganho social ou ideológico. Por isso, elas exploram a similaridade da diagramação de notícias reais, com elementos que buscam imitar um objeto digital de uma fonte habitual de consumo (portais jornalísticos de reputação), para atrair a atenção e diminuir a percepção de risco da audiência. Para essa classificação a Dismiss oferece a Catálogo dos Tipos de Desinformação <sup>4</sup> (Figura 4.3).

<sup>4</sup>O catálogo foi construído com base nos resultados do mapeamento da literatura e de análises sociotécnicas de diferentes tipos de objetos feitas pelo autor presentes no Capítulo 5.



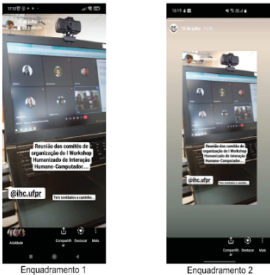
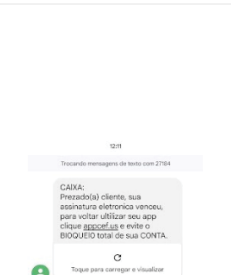
|                                                                                                                                                                                                                                                                                                                                                                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                                                                                                                                |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Rumor</b></p>  <p>Reencaminhamento de mensagem, cortes de vídeos, mensagens fora de contexto, etc;</p> <p>Fonte: Autor</p>                                                                                                                                               | <p><b>Sátira</b></p>  <p>Memes, gifs, figurinhas, montagens, hiperboles, etc;</p> <p>Fonte: <a href="https://www.poder360.com.br/eleicoes/ciro-mostra-lula-e-bolsonaro-como-o">https://www.poder360.com.br/eleicoes/ciro-mostra-lula-e-bolsonaro-como-o</a></p>                                                                                                                                                  | <p><b>Farsa</b></p>  <p>Esquemas, engodos, personificação, deturpação de conteúdo, golpes, etc;</p> <p>Fonte: Autor</p>                                                                                                                                     |
| <p><b>Descrição</b></p> <p>São objetos que podem ou não ser uma desinformação. Fica implícito para os envolvidos na comunicação que a informação não é verificada pelo interlocutor. Geralmente referencia terceiros (pessoas, organizações, etc.) como autoridade certificadora da informação. Pode ser revelada como verdade ou falsidade posteriormente.</p> | <p><b>Descrição</b></p> <p>São objetos que podem ou não ser uma desinformação. Fica implícito para os envolvidos na comunicação que a informação não é verificada pelo interlocutor. Geralmente referencia terceiros (pessoas, organizações, etc.) como autoridade certificadora da informação. Pode ser revelada como verdade ou falsidade posteriormente.</p>                                                                                                                                    | <p><b>Descrição</b></p> <p>São objetos usados deliberadamente para enganar as partes interessadas. Geralmente exploraram crenças, expectativas, e intenções. Podem ser parte de histórias complexas, como teorias da conspiração, ou bem contextualizadas, como fraudes e golpes, ou um ato oportunístico, como o cyberbullying difamação.</p> |
| <p><b>Notícia Falsa/Fake News</b></p>  <p>Vídeos, Matérias em portais informativos, entrevistas, etc;</p> <p>Fonte: 1 - <a href="https://globo.globo.com/brasil/idoso-di-pes">https://globo.globo.com/brasil/idoso-di-pes</a></p>                                             | <p><b>Teoria da Conspiração</b></p>  <p>Histórias, absurdos, ficção, etc; Iluminati, fantasmas do comunismo, terra plana, anti-vax, Q-anon, astrologia etc;</p> <p>Fonte: <a href="https://twitter.com/choquei/status/1635793607373254656">https://twitter.com/choquei/status/1635793607373254656</a></p>                                                                                                       | <p><b>Isca de clique / Click bait</b></p>  <p>Propaganda, chamidas, links, etc;</p> <p>Fonte: Propagandapaga;</p>                                                                                                                                          |
| <p><b>Descrição</b></p> <p>São um subtipo de farsas fabricadas para enganar a audiência, manipular crenças e atitudes. São similares às notícias veiculadas por autoridades certificadas, como a diagramação de jornais e portais de notícias. A audiência não está ciente da enganação. Podem ser totalmente fabricadas ou introduzir viés.</p>                | <p><b>Descrição</b></p> <p>São um subtipo de farsas que tendem a reduzir a complexidade da realidade, considerando aspectos sociais como sendo tramas maquiavélicas, muitas vezes concebidas por indivíduos ou organizações poderosas. Também podem ser oriundas de distorções no conhecimento de uma comunidade, relacionadas a confiança cega no julgamento de uma autoridade. Formam comunidades em torno do tema, com indivíduos que podem não ter ciência da farsa. Difícil de verificar.</p> | <p><b>Descrição</b></p> <p>São signos usados com a intenção de manipular a atenção da audiência e obter acessos. Frequentemente empregam linguagem sensacionalista e apelo emocional. Geralmente o conteúdo da página diverge do apresentado no signo.</p>                                                                                     |

Figura 4.3: Catálogo dos tipos de desinformação. Fonte: Autor.

- Onde o objeto aparece?
  - Informações sobre o complexo de elementos que constituem o canal de comunicação: dispositivos, tecnologias e plataformas digitais. Informações como origem, destino, tipos de dispositivos, plataformas digitais de comunicação, etc. Estas informações são importantes para compreender as capacidades do canal de comunicação, se existem e como se dão as relações entre as partes interessadas, a localização geográfica, *affordances* tecnológicas, etc. Para apoiar esta tarefa a Dismiss oferece questões-guia.
- Qual o escopo temporal do caso?
  - Toda análise de casos de desinformação deve ser situada, compreendendo o espaço e o tempo em que acontecem/aconteceram. O início e fim do caso são importantes

para evitar viés na identificação e análise de aspectos sociotécnicos, e para entender como podem estar inter-relacionados entre os elementos básicos que compõem aquele contexto do caso.

- Quais as partes interessadas?
  - São os participantes do caso, tendo papéis e interesses específicos de cada um. Identificá-los amplia o entendimento da estrutura do caso e como cada parte interessada afeta e é afetada pela desinformação. Por exemplo, controladores de redes sociais como o Facebook, possuem interesses relacionados ao ganho monetário oriundo da interação dos seus usuários com o conteúdo que circula por sua plataforma. Isso pode implicar em ações pouco efetivas para mitigar conteúdo viral com desinformações.
- Qual a temática? Qual sua relação com o ambiente?
  - A temática é um dos atributos usados para estudar casos com mais de um objeto desinformativo, identificar como o tema está relacionado ao ambiente das partes interessadas, sua ligação com acontecimentos reais, e rastrear as transformações da desinformação na linha do tempo do caso. Por exemplo, em casos de crise e catástrofes naturais é comum surgir desinformação sobre o que está acontecendo (Zhang et al., 2019a). O FRC apresenta uma lista de temas para facilitar a identificação e campos para identificar a relação com o ambiente.

#### 4.1.2.2 *Como coletar as informações?*

A coleta de informações deve ser situada, ou seja, a partir do canal no qual o objeto desinformativo surge, considerando o escopo temporal do caso e o contexto da comunicação. Nesse sentido, a Dismiss sugere duas abordagens para coleta de informações: em profundidade e em largura. No primeiro caso, o analista identifica um objeto e inicia a exploração do contexto onde ele ocorreu, buscando por elementos que respondam às perguntas-guia. A Dismiss sugere iniciar a coleta a partir do objeto, identificando o meio de comunicação imediatamente ligado ao objeto, as partes interessadas, à temática e o contexto da comunicação neste meio. Essa abordagem é utilizada na Análise Sociotécnica de Objetos Desinformativos (Seção 4.3.1).

Na estratégia de Análise Sociotécnica do Ciclo de Vida, o analista expande a busca para identificar relatos do aparecimento de objetos com um mesmo assunto em diferentes meios de comunicação em um mesmo período de tempo. Por exemplo, matérias em portais informativos, checadores de fatos e postagens em redes sociais podem revelar como as partes interessadas lidam ou lidaram com a desinformação, e como partes interessadas que não estão diretamente envolvidas podem ser impactadas. Além disso, estas fontes podem conter detalhes sobre o caso e o objeto, como quando ele surgiu, como ele se propagou, por quais tecnologias e plataformas de comunicação ele se difundiu, e as estratégias de mitigação empregadas (Seção 4.3.2).

#### 4.1.2.3 *Identificação das partes interessadas*

Além da seleção dos objetos, o escopo da análise é delimitado pelas partes interessadas que constituem o caso. Enquanto se explora o contexto do caso de desinformação, o analista deve realizar a identificação das partes interessadas considerando os níveis do DPI. Para guiar o levantamento das partes interessadas, deve-se explorar o contexto de acordo com a estratégia da abordagem adotada pelo analista. As estratégias de exploração e identificação de partes interessadas são apresentadas nas Seções 4.3.1 e 4.3.2.

A Figura 4.4 apresenta o DPI e as descrições de cada nível adaptadas para a Dismiss. No centro do DPI fica o caso de desinformação. Cada camada foca na relação que uma parte

interessada pode ter com o caso. O analista deve fazer um exercício mental visando perceber cada estágio do ciclo de vida da desinformação dentro das camadas do DPI. O DPI auxilia a identificar partes interessadas além das óbvias de locutor e audiência, ou que não estão explícitas nas fontes de informações na exploração do contexto do caso.

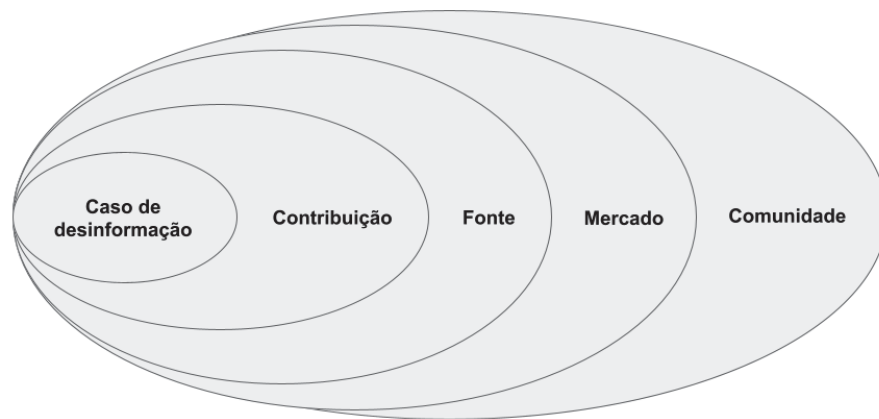


Figura 4.4: Diagrama de Partes Interessadas. Adaptado de (Kolkman, 1993).

O DPI é dividido em quatro camadas:

- **Contribuição:** Esta camada considera as partes interessadas envolvidas diretamente com a concepção e difusão da desinformação. Por exemplo, no estágio de criação, podemos identificar as partes interessadas que construíram o objeto desinformativo, partes interessadas que planejaram uma campanha de difamação, partes interessadas que o financiaram a divulgação, partes interessadas que assinam/assumem a responsabilidade diretamente pelo conteúdo, etc. Na difusão podemos identificar partes contratadas para divulgar o objeto em diferentes canais, como bots ou influenciadores patrocinados.
- **Fonte:** Nesta camada estão aqueles que fornecem informações e/ou são fonte de dados para criação do objeto, ou que ainda fazem o uso deles. Esta camada corresponde a atores que interagem com a desinformação indiretamente, disseminando e consumindo, como a audiência pretendida da desinformação, os atores que fornecem informações para concepção e difusão da mensagem em uma rede social, e simpatizantes. Por exemplo, na difusão podemos identificar partes interessadas que compartilham uma desinformação no Whatsapp no grupo da família, ou nas redes sociais pessoais, etc. No consumo, podemos identificar partes como o público-alvo, público não entendido, etc.
- **Mercado:** Na camada Mercado estão as partes interessadas que possuem alguma relação de parceria ou concorrência com o caso. Esta camada está associada às partes interessadas parceiras e concorrentes da desinformação, como redes sociais que se beneficiam com o engajamento gerado por objetos desinformativos ou que precisam mitigar seus efeitos, provedores de Internet com planos de acesso que restringem a comunicação na camada de serviço/aplicação e checadores de fatos validam mensagens.

- **Comunidade:** A camada comunidade contempla as partes interessadas que podem influenciar ou ainda sofrer influência do caso indiretamente. Ela se refere às partes interessadas que são impactadas indiretamente pela desinformação, como partes interessadas que se beneficiam do conteúdo do objeto desinformativo em disputas políticas, legisladores que precisam lidar com a prevenção e dissuasão de casos de desinformação deliberados, empresas que sofrem difamação, etc.

O analista registra os dados no Formulário para registro de partes interessadas (Figura 4.2). O formulário apresenta dicas sobre como preencher cada campo, com descrições e exemplos.

Tabela 4.2: Formulário de apoio: Registro de Partes Interessadas - Adaptação do DPI.

| Formulário de Registro de Partes Interessadas |       |         |            |
|-----------------------------------------------|-------|---------|------------|
| Contribuição                                  | Fonte | Mercado | Comunidade |
|                                               |       |         |            |
|                                               |       |         |            |
|                                               |       |         |            |

Para apoiar a aplicação do DPI, a Dismiss sugere as seguintes questões-guia<sup>5</sup>:

- Quem são os participantes do caso? Quais atividades são exercidas por eles?
  - Quem cria, quem difunde, o público-alvo, quem consome, quem financia, divulga, corrige, verifica, censura, compartilha, etc.
- Quais partes interessadas podem se beneficiar ou sofrer consequências no caso?
  - Quem investiga, estuda, analisa, legisla, advoga, controla (por exemplo, o meio de comunicação), perde/ganha dinheiro/status social, etc.
- Como as partes interessadas estão envolvidas indiretamente no caso de desinformação?
  - Pensar nos possíveis interesses e como as partes interessadas já identificadas podem estar relacionadas pode auxiliar a identificar outras partes interessadas ou papéis específicos. Por exemplo, marcas de produtos que aparecem no fundo de imagens ou fotos, fotos de pessoas famosas ou referências a entidades do estado, outros usuários de uma determinada plataforma digital que não são o público-alvo, etc.

A identificação das partes interessadas pode ocorrer em outras etapas da abordagem, não sendo uma atividade restrita da Etapa 1. As questões propostas para guiar a identificação podem ser expandidas ou modificadas, dependendo das necessidades. Para um entendimento abrangente do caso, a Dismiss recomenda identificar pelo menos uma parte interessada em cada camada do DPI.

#### 4.1.3 Etapa 2 - Análise sociotécnica do caso

Nessa etapa, o analista identifica aspectos sociotécnicos que constituem o caso de desinformação. Para isso, o analista usa a ES, as partes interessadas identificadas na etapa anterior, os canais de comunicação, o objeto desinformativo e questões-guia propostas para apoiar o levantamento de aspectos sociotécnicos de interesse para o fenômeno da desinformação digital. A ES auxilia a considerar diversos aspectos sociotécnicos em diferentes níveis de

<sup>5</sup>A literatura indica uma série de passos práticos, que vão desde a descrição do problema até o uso de questões para guiar a identificação, como em Pereira e Baranauskas (2015)

abstração. As partes interessadas trazem diferentes perspectivas em ação no caso, dando um suporte para identificar aspectos sociotécnicos do contexto sociocultural. Os canais de comunicação possibilitam identificar aspectos sociotécnicos da plataforma tecnológica que impactam a interação com o objeto digital de desinformação. Ao final dessa etapa o analista terá uma caracterização do caso e sua estrutura. Esses dados seguem para a Etapa 3 de leitura e análise de resultados.

#### 4.1.3.1 A Escada Semiótica

A Figura 4.5 apresenta a ES. A ES apoia a análise sociotécnica do caso de desinformação, estruturando a análise de aspectos relevantes em diferentes níveis de abstração. A figura mostra a ES exemplos de aspectos em cada nível.

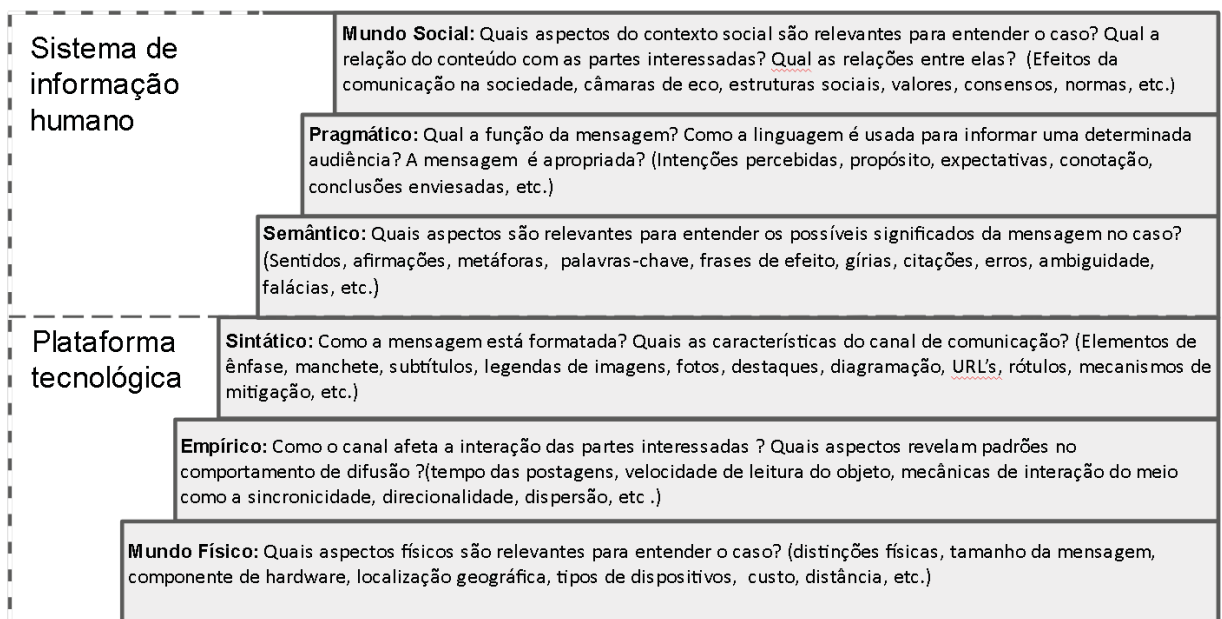


Figura 4.5: Adaptação da Escada Semiótica. Fonte: Inspirado em (Guimarães et al., 2008).

A análise deve ser feita da seguinte forma:

- Adotar a perspectiva do caso e observar os aspectos sociotécnicos relacionados às partes interessadas, ao canal de comunicação e ao objeto desinformativo;
- O analista utiliza a ES para apoiar a identificação de aspectos sociotécnicos, e formulários de registro de dados para cada estágio do ciclo de vida (criação, difusão, e consumo);
- Identificar aspectos sociotécnicos nos níveis da escada consultando as questões-guia e os materiais de apoio, e refletir sobre os aspectos sociotécnicos que sejam relevantes para as partes interessadas.

#### 4.1.3.2 Aspectos Sociotécnicos e o Fenômeno da Desinformação

Antes de explicar os passos para o levantamento de aspectos sociotécnicos é preciso compreender os aspectos socioculturais e técnicos são de interesse para análises do fenômeno da desinformação. A Tabela 4.3 reúne exemplos de aspectos relevantes encontrados na literatura para serem observados nas análises. Esses aspectos servem de base para auxiliar o analista e informar a análise, constituindo as questões-guia.



Tabela 4.3: Exemplos de aspectos relevantes em casos de desinformação.

| Nível      | Aspectos relevantes para análise sociotécnica de casos de desinformação                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Social     | <ul style="list-style-type: none"> <li>Autoridades impactam no resultado de discussões em redes sociais e na adoção de pontos de vista pelo público-alvo (Wilder e Vorobeychik, 2018);</li> <li>Correções feitas por estranhos são recebidas como ataques (Hannak et al., 2014);</li> <li>Agrupamentos permitem a disseminação mais rápida de informações (Safieddine et al., 2017);</li> </ul>                                                                                                                                                                                                                                                                                                               |
| Pragmático | <ul style="list-style-type: none"> <li>Usuários compartilham falsidades para obter as opiniões de outras pessoas, expressar as próprias opiniões e interagir com outras pessoas (Chen e Sin, 2013);</li> <li>O valor do conteúdo e o valor social de compartilhamento impactam nas decisões de propagar uma mentira deliberadamente (Wang e Fussell, 2020);</li> <li>A fonte, relevância pessoal e interesse pessoal são motivações para consumo e compartilhamento (Wang e Fussell, 2020);</li> </ul>                                                                                                                                                                                                        |
| Semântico  | <ul style="list-style-type: none"> <li>Aspectos que impactam na credibilidade: fonte oficial, linguagem formal, referências para informações. (Hassan, 2018; Ghaisani et al., 2017);</li> <li>URL conhecido, conta oficial, fotos de perfil, informações profissionais de perfil, e-mail válido, familiaridade com a fonte, etc. (Zhang e Ghorbani, 2020; Pacheco et al., 2019);</li> <li>Aspectos de preconceito/viés: linguagem sentimental, linguagem agressiva, discurso político, etc. (Patankar et al., 2019);</li> <li>A novidade/popularidade do tema atrai a atenção dos leitores que podem se impressionar facilmente (Feng e Hirst, 2013).</li> </ul>                                              |
| Sintático  | <ul style="list-style-type: none"> <li>Características visuais que impactam positivamente na percepção de credibilidade: diagramação similar a objetos digitais confiáveis, com cores, figuras, logos, acrônimos, e estilo (Flanagin e Metzger, 2007);</li> <li>Características que impactam negativamente: erros sintáticos, URLs desconhecidas, erros nas URLs, signos não usuais na URL (Zhang e Ghorbani, 2020);</li> <li>Aspectos que impactam na percepção de relevância: popularidade, quantidade curtidas (reações), compartilhamentos, hashtags, links, referências, etc. (Yasser et al., 2018; Shao et al., 2018);</li> <li>Desinformação se origina em poucas fontes. (Galitsky, 2015);</li> </ul> |
| Empírico   | <ul style="list-style-type: none"> <li>Criptografia dificulta medidas de monitoramento e detecção (Tyagi et al., 2019; Melo et al., 2019);</li> <li>A velocidade de difusão de informações falsas durante crises públicas é alta (Zhu et al., 2018);</li> <li>A sincronicidade afeta a percepção de credibilidade (Zhou e Zhang, 2007);</li> </ul>                                                                                                                                                                                                                                                                                                                                                            |
| Físico     | <p>Localização da origem e data da divulgação podem revelar contas falsas (Nilforoshan e Shah, 2019).</p> <p>Assinatura do dispositivo, formato de captura e produção de dados, (e.g. imagem gerada sinteticamente ou capturada por uma câmera, deepfakes, etc.).</p>                                                                                                                                                                                                                                                                                                                                                                                                                                         |

Com base no que a literatura indica como aspectos relevantes, a Dismiss oferece outros dois materiais de apoio ao diagnóstico de casos de desinformação que constituem eixos complementares para sustentar as análises e a identificação de aspectos sociotécnicos relevantes. Os Quadros de Aspectos Sociotécnicos do Meio de Comunicação 4.4 e Princípios de Influência 4.5 são elementos de suporte opcionais que podem ser utilizados pelo analista.

#### 4.1.4 Etapa 3 - Identificação, Análise e Entendimento de Problemas

Nessa etapa, o analista revê os aspectos sociotécnicos identificados e busca por suas inter-relações. Para isso deve-se reler os aspectos identificados e registrados na Escada Semiótica e buscar entender como podem estar inter-relacionados. Com estas informações o analista reflete sobre os aspectos mais relevantes do caso para identificar características do objeto desinformativo, do canal de comunicação que devem ser levadas em consideração para o desenvolvimento de intervenções no fenômeno. O analista considera a inter-relação entre os aspectos de duas formas:

- Nos degraus da Escada Semiótica: deve-se analisar como os aspectos identificados em cada degrau se relacionam entre si. A Dimsis sugere o seguinte questionário para apoiar a identificação de problemas:
  - De que modo aspectos da plataforma tecnológica impactam a interação da audiência e entendimento da mensagem? Como o conteúdo da mensagem se relaciona com



Tabela 4.4: Quadro de Aspectos do Meio de Comunicação. Fonte: Autor.

| Interação                   | Síncrona                                                                                                                                                                                                                 |                                                                                                                                                                                                                                        | Assíncrona                                                                                                                                                                                                   |                                                                                                                                                                                  |
|-----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                             | Unidirecional                                                                                                                                                                                                            | Bidirecional                                                                                                                                                                                                                           | Unidirecional                                                                                                                                                                                                | Bidirecional                                                                                                                                                                     |
| <b>Um p/<br/>um</b>         | Ex: Propaganda por telefone;<br><br>-Privilégios diferentes de interação;<br>-Receptor restrito à audiência;<br>- Mensagem acessível somente em tempo de interação;<br>-Alcance limitado;<br>-Compartilhamento restrito; | Ex: Vídeo-chamada;<br><br>-Privilégios iguais de interação;<br>-Difusão restrita;<br>-Resolução de conflito possível;<br>-Tempos de resposta curto;<br>-Compartilhamento restrito;                                                     | Ex: Recado de voz (ligação celular);<br><br>-Privilégios diferentes de interação;<br>-Receptor restrito à audiência;<br>-Difusão limitada (entre pares);<br>-Duração indefinida;<br>-Compartilhamento fácil; | Ex: E-mail;<br><br>-Privilégios iguais de interação;<br>-Resolução de conflito possível;<br>-Argumentação direta;<br>-Tempo de interação indefinido;<br>-Compartilhamento fácil; |
| <b>Um p/<br/>muitos</b>     | Ex: Transmissão de TV ao vivo;<br><br>-Privilégios diferentes de interação;<br>-Interlocutor como curador da informação;<br>-Maior controle da interação;<br>- Audiência só pode receber;<br>-Compartilhamento restrito; | Ex: Live;<br><br>-Privilégios semelhantes de interação;<br>-Interlocutor autoridade;<br>-Controle sobre a interação de outros;<br>-Resolução de conflitos possível;<br>-Tempos de resposta curto;<br>- Restrições de compartilhamento; | Ex: Pop-up;<br><br>-Privilégios diferentes de interação;<br>-Receptor restrito à audiência;<br>-Difusão abrangente;<br>-Maior alcance e velocidade;<br>-Compartilhamento restrito;                           | Ex: E-mail (mala-direta);<br><br>-Privilégios iguais de interação;<br>-Resolução de conflitos possível;<br>-Tempos de resposta curto;<br>-Compartilhamento fácil;                |
| <b>Muitos p/<br/>muitos</b> | Ex: Plataforma de trabalho colaborativo;<br><br>-Privilégios iguais de interação;<br>-Resolução de conflitos possível;<br>-Maior quantidade de conflitos;<br>-Compartilhamento restrito;                                 | Ex: Videoconferência;<br><br>-Privilégios iguais de interação;<br>-Tempos de resposta curto;<br>-Resolução de conflitos possível;<br>-Compartilhamento restrito;                                                                       | Ex: Site Jornalístico;<br><br>-Privilégios semelhantes de interação;<br>-Moderadores;<br>-Resolução de conflitos possível;<br>-Tempo de interação indefinido;                                                | Ex: Grupos virtuais (WhatsApp);<br><br>-Privilégios iguais de interação;<br>-Resolução de conflitos possível;<br>-Tempo de interação indefinido;<br>-Compartilhamento fácil;     |

Tabela 4.5: Quadro de Princípios de Influência. Fonte: Autor.

|                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|-------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Afinidade</b>        | Pessoas tendem a gostar daquilo que as representa, que se parece com elas. Aspectos como o uso de uma linguagem significativa para uma audiência (gírias e jargões), uma aparência agradável, usar opiniões positivas ou que confirmam crenças pre-existentes, a similaridade na estrutura/diagramação de um conteúdo, o tamanho de uma mensagem, e as cores são exemplos de aspectos que impactam a percepção de risco e a avaliação de veracidade de um objeto digital.      |
| <b>Aprovação social</b> | Pessoas querem se sentir validadas pelo que as outras pessoas estão fazendo ao seu redor, em especial quando em dúvida. Eventos ocorrendo no ambiente de um grupo têm impacto no que o grupo percebe como comportamento correto. Aspectos deste princípio são consensos, popularidade de um tema, valores morais, crenças, etc. Este princípio ganha força na presença de incerteza, induzindo a tendência de um indivíduo em agir como o grupo está agindo - o efeito manada. |
| <b>Autoridade</b>       | Pessoas tendem a seguir figuras de autoridade. Identifique aspectos que impactam a percepção de risco como a popularidade de indivíduos em um grupo, como celebridades e influenciadores digitais; A autoridade de pessoas públicas, pessoas que são referência nas atividades que realizam, como médicos, cientistas etc. Busque por imagens, referências nos textos, assinaturas digitais, logotipos, certificações, etc.                                                    |
| <b>Comprometimento</b>  | Pessoas querem que seus compromissos sejam consistentes com os seus valores. Comprometimento tem base no que as pessoas estão fazendo, e no que as outras pessoas pensam que ela está fazendo. Assim, aspectos como a coerência, a fluência, domínio de um assunto, a expectativa da audiência de ter suas necessidades atendidas, ou defender as ideologias de um determinado grupo podem gerar afinidade e comprometimento, facilitando o engano.                            |
| <b>Escassez</b>         | A tendência das pessoas de temerem perder algo é pior que a possibilidade de deixar de ganhar. Algo que se torna (ou pode se tornar em breve) escasso ou indisponível, tem mais valor do que quando era abundante - como as corridas por papel higiênico no início da pandemia [BBC 2020], a inexistência de uma vacina/conhecimento sobre o vírus, ou um determinado grupo que luta por direitos.                                                                             |
| <b>Reciprocidade</b>    | É a expectativa inerente de quando se recebe algo com valor positivo, um sentimento de dívida surge. Explorado por mensagens ou informações não solicitadas, que geram uma ligação com a necessidade de responder e se engajar, ou atender a pedidos posteriores. É usada para construir um senso de consciência. Alguns elementos explorados são a exposição seletiva, percepção seletiva, e controle do fluxo de informações.                                                |
| <b>Unidade</b>          | A tendência de buscar nos outros, traços de identidade que nos definem, e levam a formação de grupos com indivíduos ideologicamente alinhados. É um princípio que deriva da interação social e da criação de laços de confiança entre indivíduos em grupos. Aspectos tribais da sociedade como gírias, sentimentos de familiaridade, proximidade podem levar a ideia de "nós" contra "eles".                                                                                   |

as diferentes partes interessadas? Como o objeto desinformativo foi construído? Como o objeto guia a leitura? Qual a mensagem que busca concluir? Quais referências ele aponta? Quais as possíveis consequências do caso?

- Entre partes interessadas: deve-se analisar as partes interessadas de cada grupo do DPI e como se posicionam no caso. A Dimsis sugere o seguinte questionário para apoiar a identificação de problemas:
  - Quais interesses, expectativas, propósito das diferentes partes interessadas que constituem o caso? Como as partes interessadas estão organizadas no caso? Quais delas são autoridades, celebridades, influenciadores digitais e como se relacionam com o objeto desinformativo? Como elas se relacionam entre si? Como elas se relacionam com o canal de comunicação? Qual delas exerce maior controle na comunicação ao utilizar aquele canal? Como o canal de comunicação impacta a interação de cada parte?

Depois de identificar as inter-relações deve-se sugerir problemas com base nessas inter-relações.

O que podemos entender com os aspectos sociotécnicos identificados?

- Motivações das partes interessadas;
- Problemas com a regulamentação, rastreamento e responsabilização;
- Aspectos socioculturais de influência no consumo;
- Contexto social e cultural das partes interessadas;
- Consequências e a Repercussão: comportamentos e atitudes prejudiciais, etc;
- Responsabilidades dos controladores dos canais;
- Relacionamento de diferentes partes interessadas com os canais;
- Limitações técnicas para mitigação;
- Regulação e controle de usuários e conteúdos;
- Padrões de conteúdo em diferentes objetos de um mesmo caso;
- Estrutura e comportamento de difusão;

## 4.2 UM MODELO CONCEITUAL DO CICLO DE VIDA DA DESINFORMAÇÃO

Ao se abordar o fenômeno da desinformação a partir das partes interessadas, existe a necessidade de uma estrutura para conduzir a investigação e delimitar o escopo de análise. Quando se considera as partes interessadas, o aspecto "tempo" é fundamental para construir um entendimento abrangente dos problemas relacionados à uma determinada desinformação digital e identificar aspectos de modo situado em caso de desinformação analisado. Isso é importante pois aspectos sociotécnicos como crenças, valores, autoridades, consensos, e mesmo fatores tecnológicos como os mecanismos de interação ofertados por uma determinada tecnologia de comunicação tem efeitos válidos dentro de um tempo específico, para partes interessadas específicas, e em um espaço específico<sup>6</sup>. Por isso, esta tese propõe um modelo conceitual do Ciclo de Vida da Desinformação Digital que estrutura o estudo do fenômeno da desinformação digital, considerando os estágios da vida de uma desinformação para auxiliar os projetistas a lidar com a complexidade do fenômeno. A caracterização dos estágios do ciclo de vida pode ser encontrada no Apêndice A.6.

<sup>6</sup>O aspecto *tempo* é relevante para entender a relação entre aspectos sociotécnicos e a desinformação para um determinado contexto em um determinado momento, e estudar os potenciais impactos para as partes interessadas.

Uma desinformação digital existe na forma de um objeto digital, como notícias falsificadas, isca-clique (Zhou e Zafarani, 2018), mensagens do WhatsApp (Melo et al., 2019), ou e-mails (Bazan, 2017). Esse objeto digital é composto por diferentes signos que podem representar uma desinformação. Nessa tese, o termo objeto desinformativo é usado para referenciar esse tipo de objeto digital. Cada objeto desinformativo está atrelado a um ciclo de vida.

O ciclo de vida da desinformação engloba os elementos: partes interessadas, objetos desinformativos, e canais de comunicação. As partes interessadas são as participantes do caso, como quem cria um objeto desinformativo, quem ganha/perde com a desinformação, quem controla as plataformas de comunicação, quem distribui a desinformação, quem consome, quem fornece serviços para acesso às redes sociais, quem investiga/corrigi, etc. Os objetos desinformativos são as desinformações no formato digital, com diferentes tipos de objetos que possuem diferentes características e impactos. Os canais de comunicação são o meio pelo qual as partes interessadas e os objetos desinformativos interagem.

Conforme apresentado no Capítulo 3, existem quatro estágios no ciclo de vida um objeto desinformativo<sup>7</sup>:

1. **Criação:** Esse estágio foca no surgimento de um objeto desinformativo, quando a desinformação é construída e divulgada pela primeira vez.
2. **Difusão:** Esse estágio foca nos aspectos da propagação do objeto desinformativo, como as capacidades da mídia de comunicação e o comportamento de difusão;
3. **Consumo:** Nesse estágio o foco está na interação da audiência com o objeto desinformativo, relacionado a ações como, ler, compartilhar, ignorar e reagir;
4. **Reiteração:** Esse estágio foca no reaparecimento de um objeto desinformativo e das transformações que ocorrem durante sua vida.

Estratégias de **Mitigação** podem ser aplicadas para combater a desinformação em qualquer estágio do ciclo. A Figura 4.6 apresenta uma representação do ciclo de vida da desinformação e seus estágios (criação, difusão e consumo).

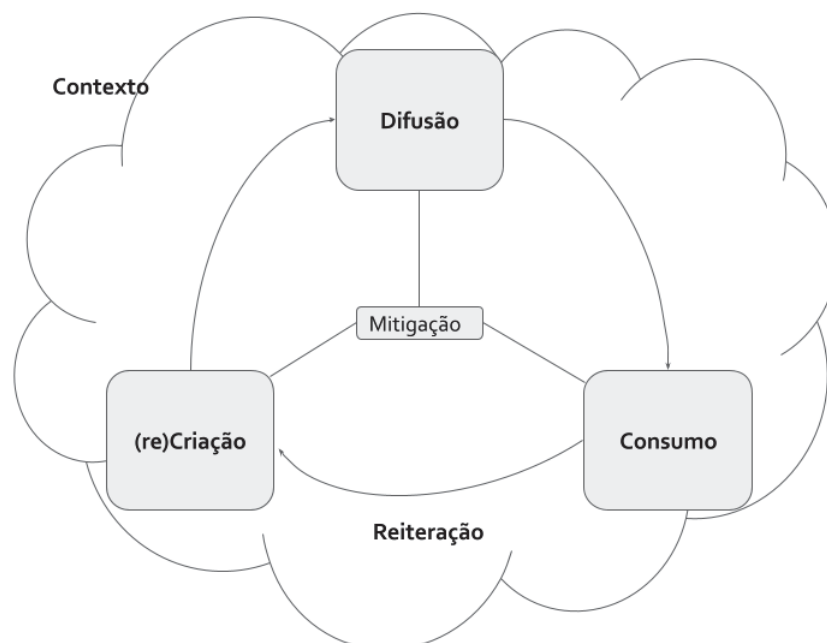


Figura 4.6: Estrutura do ciclo da desinformação digital.

<sup>7</sup>A caracterização dos estágios é apresentada em detalhes no Apêndice A.6.

Não é uma tarefa trivial prever a criação de um objeto desinformativo, e não se sabe como um objeto se transformará ou quando desvanecerá (Qin et al., 2018). Conforme a interação do meio, do objeto e da audiência ocorrem, o objeto desinformativo pode migrar para outros canais de comunicação, sofrendo alteração de formato e/ou conteúdo. A Figura 4.7 apresenta um exemplo de um caso de desinformação no ciclo de vida, da criação às suas transformações enquanto é difundido. No primeiro quadro um objeto desinformativo é criado, enfatizando alguns aspectos de interesse e excluindo outros. No segundo quadro, o objeto é divulgado por um canal de comunicação, usando de elementos de ênfase (títulos e subtítulos) para conduzir a atenção da audiência e promover uma determinada perspectiva. No terceiro quadro, o objeto desinformativo é propagado em uma outra mídia de transmissão, onde aspectos sociotécnicos como a popularidade de quem o compartilhou, são incorporados no contexto do caso. No quarto quadro, a mensagem é observada em um ambiente de interação profissional com pressões sociais (relações e interesses) que agregam mais complexidade para análise do consumo. Considerar a transformação de um objeto é importante para a **rastreabilidade** (origem, disseminação, modificações, etc.), para a responsabilização das partes interessadas envolvidas na criação e difusão, na identificação de narrativas e no estudo de campanhas desinformativas orquestradas (Starbird et al., 2019).



Figura 4.7: Exemplo de um caso de desinformação pela perspectiva do ciclo de vida da desinformação.

Como a Figura 4.7 demonstra, um caso de desinformação apresenta uma complexidade de aspectos e partes interessadas envolvidas em diferentes níveis de abstração em diferentes contextos. Nesse sentido, o Ciclo de Vida da Desinformação auxilia a abordar o fenômeno de modo sistêmico, estruturando a investigação de casos, auxiliando a relacionar partes interessadas, canais de comunicação e objetos desinformativos em cada estágio da vida de uma desinformação, e apoiando a análise dos diferentes enquadramentos, construindo uma visão sociotécnica do fenômeno.

### 4.3 ESTRATEGIAS DE APLICAÇÃO

A Dismiss consiste de duas estratégias para análise de casos de desinformação: 1) a análise sociotécnica de objetos desinformativos e 2) a análise sociotécnica do ciclo de vida da

desinformação digital. A estratégia para análise de objetos desinformativos foca na análise de um objeto digital de desinformação em somente um canal de comunicação, possibilitando um estudo direto e simplificado (em relação a análise do ciclo) para auxiliar no entendimento de problemas em um contexto mais imediato de análise. A estratégia de análise do ciclo de vida expande a análise de objetos para diferentes estágios do seu ciclo de vida, possibilitando um estudo abrangente que engloba as transformações do objeto e o entendimento de aspectos inter-relacionados nos estágios do ciclo de um modo sistêmico.

#### 4.3.1 Análise sociotécnica de objetos desinformativos (Profundidade)

A estratégia de análise sociotécnica de objetos desinformativos forma a base da abordagem de análise sociotécnica de casos de desinformação. Ela foca na análise de um objeto em um canal de comunicação, permitindo caracterizar o contexto em que a desinformação aparece e melhor compreender os problemas.

Para exemplificar a aplicação da Estratégia 1 utilizaremos um caso real. Em março de 2020 circulou pelo WhatsApp no Brasil uma mensagem que relacionava o combate do SARS-CoV-2 com o nível de Ph do corpo humano (Uol, 2020b; G1, 2021b). A mensagem apresentava um tratamento onde o indivíduo deveria comer certos alimentos para atingir um nível de Ph que supostamente aniquila a SARS-CoV-2. A mensagem trazia em sua estrutura referência a autoridades científicas, informações sobre alimentos com pH errado, além da orientação de que é possível alterar o Ph do corpo (Johnson, 2020). A Figura 4.8 apresenta o objeto desinformativo original, a primeira aparição dessa desinformação no Brasil em março de 2020. Ainda, essa desinformação foi corrigida no final de março por agências brasileiras de checagem de (Lupa e G1) (Uol, 2020b; G1, 2021b), apesar das correções, o objeto continuou circulando.

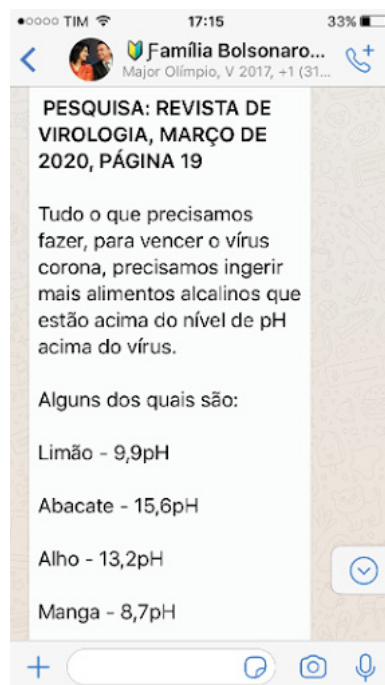


Figura 4.8: Mensagem desinformativa. Fonte: Uol (Uol, 2020b).

##### 4.3.1.1 Etapa 1 - Familiarização

Nessa etapa, o análise observa o contexto da mensagem e identifica os elementos básicos para análise. O passo-a-passo a seguir representa a exploração do contexto em profundidade. O

analista identifica um objeto (Figura 4.8) e inicia a exploração do contexto onde ele ocorreu. A Dismiss sugere iniciar a exploração do canal de comunicação diretamente onde o objeto aparece, identificando que constituem esse complexo. As Figuras 4.9 e 4.10 apresentam um exemplo desta exploração e da identificação de objetos e dos elementos do canal de comunicação.

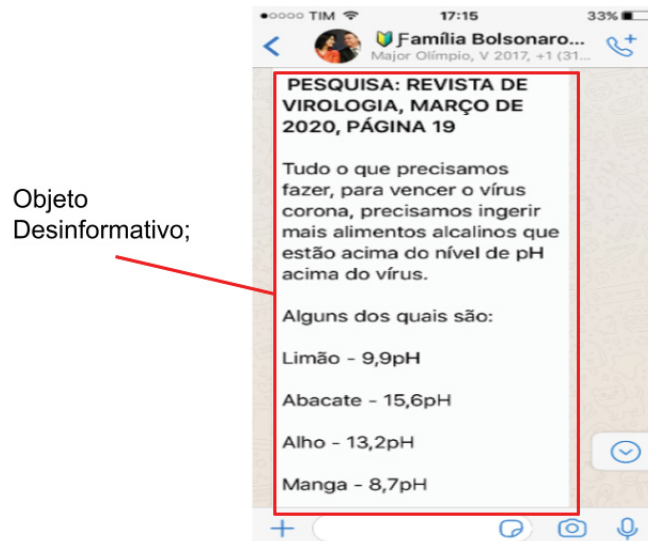


Figura 4.9: Identificação de Objetos Desinformativo.



Figura 4.10: Exploração do complexo de elementos do canal de comunicação (Dispositivos, Tecnologias e Plataformas).

O analista preenche as informações no FRED como no exemplo a seguir (Tabela 4.6).



Tabela 4.6: Formulário de apoio: registro de caso de desinformação. Fonte: Autor.

| Formulário de Registro de Caso                                                                            |                     |                                               |                    |
|-----------------------------------------------------------------------------------------------------------|---------------------|-----------------------------------------------|--------------------|
| <b>Título do caso:</b> <i>Farsa Covid-19 e pH</i>                                                         |                     | <b>Temática do caso:</b> <i>Saúde</i>         |                    |
| <b>Data do caso:</b> <i>03/2020</i>                                                                       |                     | <b>Tipo de objeto:</b> <i>Farsa</i>           |                    |
| <b>Complexo de elementos do canal de comunicação:</b>                                                     | <b>Dispositivos</b> | <b>Tecnologias</b>                            | <b>Plataformas</b> |
|                                                                                                           | -Smartphones;       | - xDSL;TLLx;<br>-Wifi;Ethernet;<br>-3G/4G/5G; | -WhatsApp;         |
| <b>Descrição do caso:</b> Desinformação sobre tratamento alternativo para a Covid-19 - pH do corpo humano |                     |                                               |                    |
| <b>Fontes:</b> <a href="#">Uol</a>                                                                        |                     |                                               |                    |

#### 4.3.1.2 Familiarização - Identificando Partes Interessadas

Nessa etapa, o analista observa os dados do contexto do objeto desinformativo para identificar as partes interessadas. O DPI auxilia o analista nesta tarefa, proporcionando uma arcabouço para pensar em atores em diferentes níveis. As Figuras 4.11 e 4.12 representam a exploração do caso para identificação das partes interessadas.

| Formulário de Registro de Caso                        |                                                                                                                                                                                                                                                                       |                                                 |                    |
|-------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------|--------------------|
| <b>Título do caso:</b>                                | <i>Farsa Covid-19 e ph</i>                                                                                                                                                                                                                                            | <b>Temática do caso:</b>                        | <i>Saúde</i>       |
| <b>Data do caso:</b>                                  | <i>03/2020</i>                                                                                                                                                                                                                                                        | <b>Tipo de objeto:</b>                          | <i>Farsa</i>       |
| <b>Complexo de Elementos do Canal de comunicação:</b> | <b>Dispositivos</b>                                                                                                                                                                                                                                                   | <b>Tecnologias</b>                              | <b>Plataformas</b> |
|                                                       | -Smartphones;                                                                                                                                                                                                                                                         | - xDSL;TLLx;<br>- Wifi;Ethernet;<br>- 3G/4G/5G; | - WhatsApp;        |
| <b>Descrição do caso:</b>                             | <i>Desinformação sobre tratamento alternativo para Covid-19 - ph do corpo humano</i>                                                                                                                                                                                  |                                                 |                    |
| <b>Fontes:</b>                                        | <a href="https://noticias.uol.com.br/politica/ultimas-noticias/2020/03/31/atos-anti-isolamento-de-bolsonaro-ativam-rede-de-fake-news">https://noticias.uol.com.br/politica/ultimas-noticias/2020/03/31/atos-anti-isolamento-de-bolsonaro-ativam-rede-de-fake-news</a> |                                                 |                    |

Figura 4.11: Identificação de Partes Interessadas.

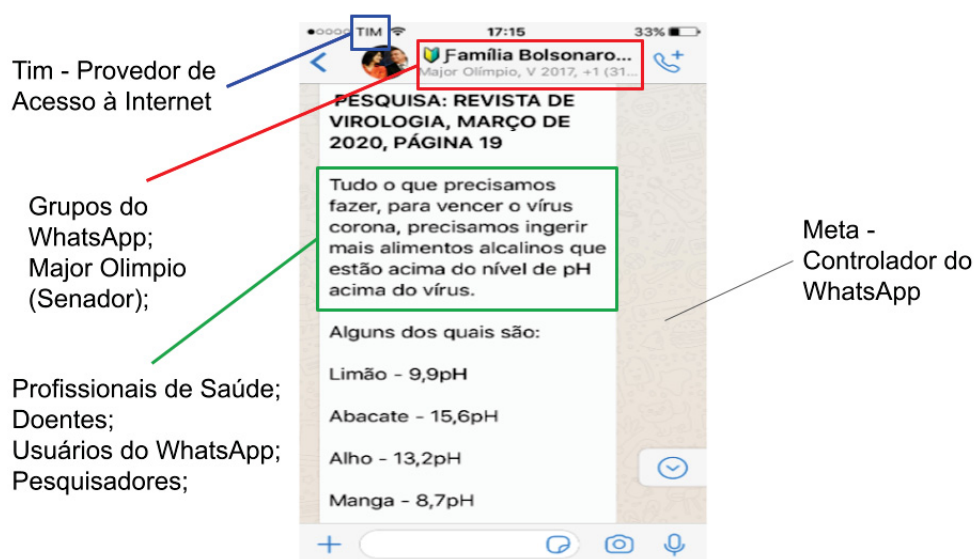


Figura 4.12: Identificação de Partes Interessadas.

A Figura 4.13 apresenta uma representação do artefato DPI preenchida com algumas partes interessadas identificadas no caso.

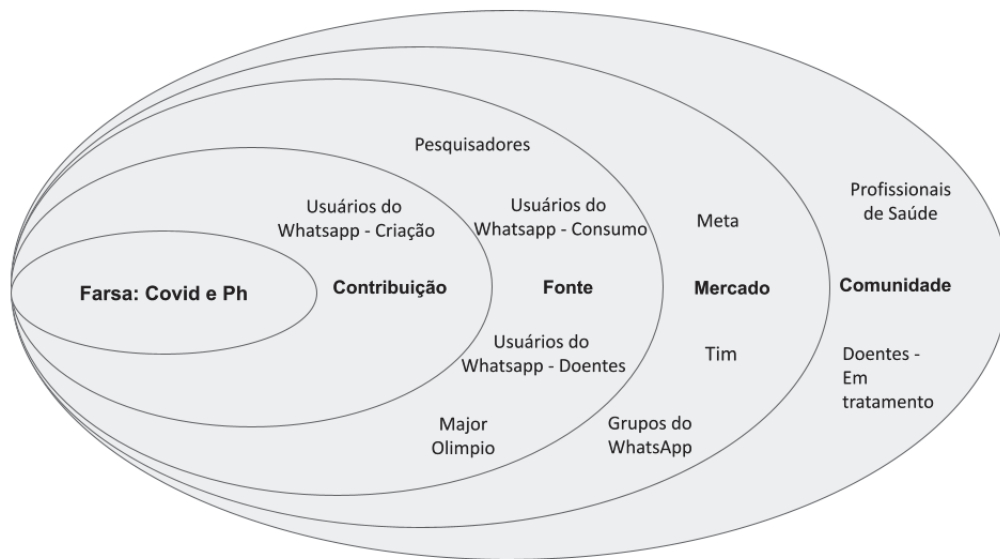


Figura 4.13: Mapeamento das Partes Interessadas no DPI.

As partes interessadas identificadas devem ser registradas no Formulário de Registro de Partes Interessadas (Figura 4.7).

Tabela 4.7: Mapeamento das Partes Interessadas no DPI.

| Formulário de Registro de Partes Interessadas |                             |                    |                        |
|-----------------------------------------------|-----------------------------|--------------------|------------------------|
| Contribuição                                  | Fonte                       | Mercado            | Comunidade             |
| Usuários Whatsapp (criação)                   | Pesquisadores               | Meta               | Profissionais de Saúde |
|                                               | Usuários Whatsapp (consumo) | Tim                | Doentes                |
|                                               | Usuários Whatsapp (doentes) | Grupos do WhatsApp |                        |
|                                               | Major Olimpio               |                    |                        |

Ao final dessa etapa, o analista terá identificado os elementos básicos para se conduzir a análise sociotécnica da desinformação digital:

- Ecossistema do fenômeno no caso: as partes interessadas em diferentes níveis de abstração e como estão dispostas em relação ao canal de comunicação e o objeto desinformativo;
- Estrutura da difusão: os elementos que constituem o complexo do canal de comunicação nos diferentes níveis técnico/tecnológicos;

#### 4.3.1.3 Etapa 2 - Análise Sociotécnica do Caso

Nessa etapa, o analista emprega o artefato da ES para investigar o objeto desinformativo, considerando os elementos identificados na Etapa 1. O analista identificara aspectos sociotécnicos relevantes, amparado por materiais de apoio e questões que guiam a análise. Ao final desta etapa o analista terá um conjunto de aspectos sociotécnicos que caracterizam o contexto em que o objeto desinformativo aparece. Esses dados seguem para a Etapa 3 de leitura, análise e discussão de resultados. A Figura 4.14 mostra uma representação visual de como a análise pode ser feita.



Figura 4.14: Mapeamento das Partes Interessadas no DPI.

Nessa etapa, o análise observa o contexto da mensagem e identifica os elementos básicos para análise. O passo-a-passo a seguir representa a exploração do contexto em profundidade.

Tabela 4.8: Formulário de Registro de Aspectos Sociotécnicos (Análise do Objeto) - Preenchido.

| Formulário de Registro de Partes Interessadas |                                                                                                                                       |
|-----------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|
| Mundo Social                                  | Explora o tema popular para atrair a audiência;<br>Explora a falta de informações sobre o vírus e oferece solução simples;            |
| Pragmático                                    | Deturpa conclusões<br>Tenta induzir o consumo de tratamento;                                                                          |
| Semântico                                     | Linguagem informal induz afinidade<br>Erro na escala do pH (vai de 0-15);                                                             |
| Sintático                                     | Dificulta a verificação (sem links para a fonte do conteúdo);<br>Estrutura guia a atenção com elementos de ênfase (título, listagem); |
| Empírico                                      | Facilidade de compartilhamento;<br>Facilidade de alterar a mensagem;                                                                  |
| Mundo Físico                                  | Tamanho do objeto (qntd de elementos na mensagem);<br>Quantidade de smartphones no Brasil;                                            |

#### 4.3.1.4 Etapa 3 - Identificar Problemas

Nessa etapa, o analista revê os aspectos sociotécnicos identificados e busca por suas inter-relações nos estágios do ciclo de vida da desinformação. Com estas informações o analista reflete sobre os aspectos mais relevantes do caso para identificar características do objeto desinformativo, do canal de comunicação e de cada estágio do ciclo que devem ser levadas em consideração para o desenvolvimento de intervenções no fenômeno. O analista considera a inter-relação entre os aspectos de duas formas:

- **Entre partes interessadas:** O analista deve observar como aspectos em diferentes níveis de abstração para cada parte interessada podem estar inter-relacionados, e como podem impactar as partes interessadas. Por exemplo, quais interesses, expectativas, e propósito do caso? Como impactam a criação, difusão e consumo? Quais partes

interessadas representam autoridades (Influenciadores, políticos, especialistas, etc.)? É possível identificar relações entre grupos de partes interessadas? Relações afetivas?

- **Entre dimensões semióticas:** O analista deve observar como aspectos em diferentes degraus da Escada Semiótica podem estar inter-relacionados em um estágio, e como podem impactar o caso de desinformação. Por exemplo, quais aspectos da plataforma tecnológica afetam a representação ou interpretação? Como a temática/diagramação/tipo de objeto desinformativo se relaciona com o ambiente das partes interessadas? Como as referências e fontes de informação se relacionam com o conteúdo do objeto? São populares? Qual a relação entre o conteúdo e o posicionamento ideológico das partes interessadas?

Ainda, os formulários de registro servem como um banco de dados de casos analisados. O banco de dados auxilia no estudo em largura do fenômeno, possibilitando identificar as mudanças na criação de objetos desinformativos em torno de uma temática ou fatos conhecidos, as transformações no contexto e partes interessadas, a identificar narrativas e monitorar sua evolução (como teorias da conspiração), e no estudo de campanhas orquestradas<sup>8</sup>.

O que podemos entender com os aspectos sociotécnicos identificados:

- Entendimento sobre as diferentes partes interessadas envolvidas em um caso de desinformação;
- Motivações das partes interessadas;
- Contexto social e cultural das partes interessadas;
- Entendimento sobre os aspectos sociotécnicos e suas inter-relações no ciclo de vida da desinformação;
- Problemas com a regulamentação, rastreio e responsabilização;
- Aspectos socioculturais de influência no consumo;
- Limitações técnicas para mitigação;
- Regulação e controle de usuários e conteúdos;
- Banco de dados de casos e objetos analisados;

#### 4.3.2 Análise sociotécnica do Ciclo de Vida da Desinformação

A estratégia de análise sociotécnica do ciclo de vida emprega as mesmas técnicas e passos da estratégia de análise de objetos, porém a estende para analisar o ciclo de vida da desinformação. Nessa estratégia o analista busca perceber o caso pela ótica de cada estágio do ciclo, servindo de arcabouço para guiar a análise, trazendo uma perspectiva que ressalta os aspectos de interesse específicos de cada estágio (criação, difusão e consumo). A estratégia de análise sociotécnica do ciclo de vida apoia a construção de uma visão sistêmica do fenômeno, permitindo compreender a linha do tempo do caso, as transformações da desinformação em diferentes objetos e padrões de conteúdo, as diferentes interações das partes interessadas e suas perspectivas para cada estágio, e como os elementos básicos de um caso podem estar inter-relacionados, integrando problemas complexos no fenômeno da desinformação digital.

Para exemplificar a aplicação da Estratégia 2, utilizaremos um caso real<sup>9</sup>. Em janeiro de 2023 circulou por diversos meios de comunicação no Brasil uma mensagem que afirmava sobre um reajuste específico para auxílio reclusão que passaria ser maior que o salário mínimo (Uol, 2020b; G1, 2021b). A mensagem apresentava uma foto do Presidente Lula ao lado de uma outra foto de um ator encapuzado e armado, apontando uma arma para o espectador, e um texto que

<sup>8</sup>Um tutorial de aplicação e exemplos podem ser encontrados no endereço: <https://docs.google.com/spreadsheets/d/1lZORTRirkr-kKYzZXC9gV8lip3d3lfsVG7X8nY7RLbg/>.

<sup>9</sup>Análise sociotécnica completa do caso em <https://www.youtube.com/watch?v=LfMrysNp9p8>.

destacava "aumento para b@ndido". A Figura 4.15 apresenta o objeto desinformativo. Ainda, essa desinformação foi corrigida no final de março por agências brasileiras de checagem de (Lupa e G1) (Fagundes, 2023; G1, 2023b). Apesar das correções, a desinformação continuou circulando.



Figura 4.15: Objeto desinformativo para exemplificação da Estratégia 2. Fonte: WhatsApp do autor.

#### 4.3.2.1 Etapa 1 - Identificar Elementos Básicos

Nessa etapa, o análise observa o contexto da mensagem e identifica os elementos básicos para análise. O passo-a-passo a seguir representa a exploração do contexto em profundidade. O analista identifica um objeto (Figura 4.8) e inicia a exploração do contexto onde ele ocorreu. A Dismiss sugere iniciar a exploração do canal de comunicação diretamente onde o objeto aparece, identificando os elementos que constituem esse complexo. As Figuras 4.9 e 4.10 apresentam um exemplo dessa exploração.





Figura 4.16: Exploração do contexto com base no objeto.

Observando o conteúdo do objeto, nota-se ao menos três canais de comunicação diferentes. O analista deve "seguir" estas fontes para explorar o contexto buscando entender a linha do tempo do caos. No exemplo, a desinformação referencia o site "Agora Notícias" como fonte do conteúdo. Além disso, a assinatura no meio do objeto desinformativo apresenta outro canal de comunicação que pode ter originado o objeto desinformativo. Ainda, o objeto foi recebido no Whatsapp. Nessa análise inicial, busca-se pela origem da desinformação, rastreando os diferentes meios por onde passou.

As Figuras 4.17, 4.18, 4.19, 4.20 e 4.21 exemplificam a exploração desses canais encontrados no conteúdo do objeto desinformativo, em busca novos meios e elementos para análise.

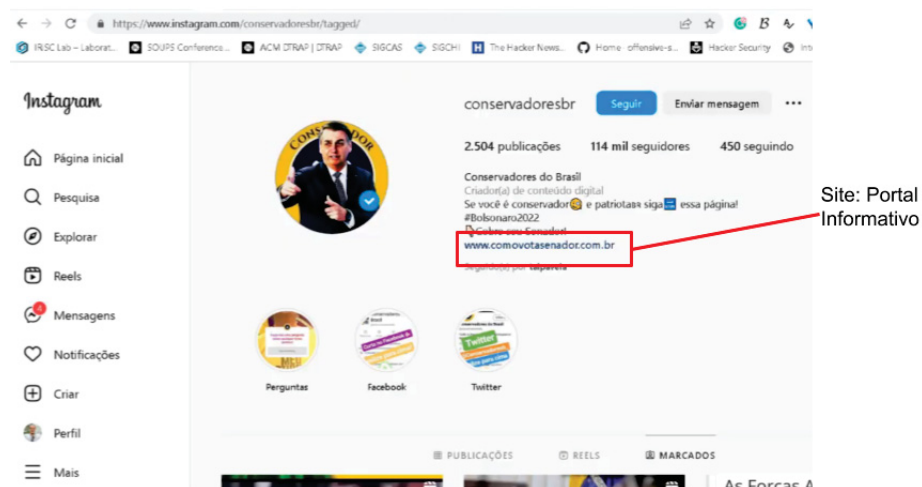


Figura 4.17: Exploração do contexto com base no ciclo de vida - Busca no Instagram.

A Figura 4.17 mostra o canal do Instagram referenciado no objeto desinformativo. Na imagem é possível perceber links para outros canais de comunicação que podem ser explorados.





Figura 4.18: Exploração do contexto com base no ciclo de vida - Busca no Site Agora Notícia.

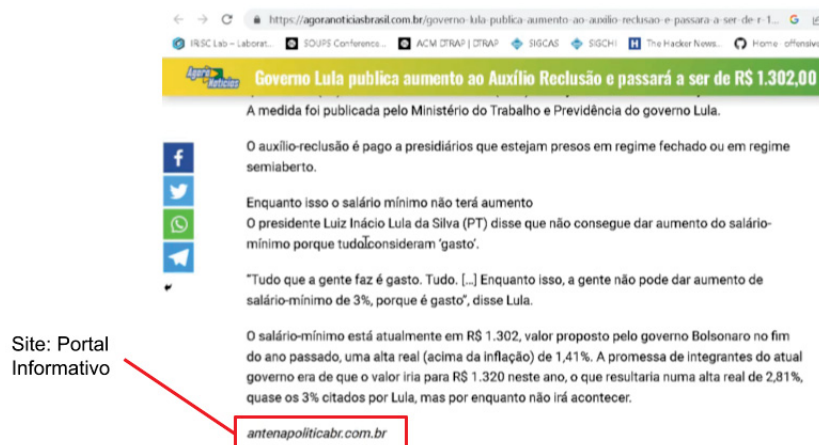


Figura 4.19: Exploração do contexto com base no ciclo de vida - Busca no Site Agora Notícia.

As Figuras 4.18 e 4.19, mostram o canal de comunicação da suposta fonte das informações. A exploração do conteúdo revela uma outra fonte para a desinformação: o site antenapolitica.com.br.



Figura 4.20: Exploração do contexto com base no ciclo de vida - Busca no Site Antena Política.

A Figura 4.20 mostra o site "Antena Política" e a desinformação quase idêntica (mesmo título, mesmas figuras; texto um pouco diferente) ao site "Agora Notícias".

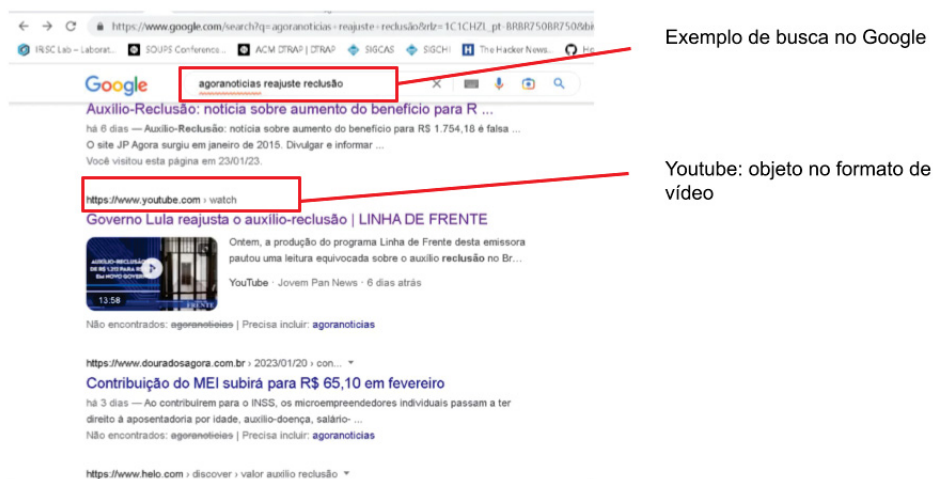


Figura 4.21: Exploração do contexto com base no ciclo de vida - Busca no Google.

A Figura 4.21 mostra outros meios de explorar o contexto do caso, como pesquisas no Google/YouTube ou outros buscadores.

O analista deve considerar quais são os elementos que constituem o complexo dos diferentes canais de comunicação e quais os dispositivos e as tecnologias em cada plataforma tecnológica para preencher o FRC. A Tabela 4.9 apresenta um exemplo de preenchimento do FRC considerando a Estratégia 2.

Tabela 4.9: Formulário de Registro de Caso - Preenchido.

| Formulário de Registro de Caso                                                                        |                                     |                                           |                                                 |
|-------------------------------------------------------------------------------------------------------|-------------------------------------|-------------------------------------------|-------------------------------------------------|
| <b>Título do caso:</b> Farsa - aumento auxílio reclusão                                               |                                     | <b>Temática do caso:</b> Economia         |                                                 |
| <b>Data do caso:</b> 01/2023                                                                          |                                     | <b>Tipo de objeto:</b> Farsa              |                                                 |
| <b>Complexo de elementos do canal de comunicação:</b>                                                 | <b>Dispositivos</b>                 | <b>Tecnologias</b>                        | <b>Plataformas</b>                              |
|                                                                                                       | Smartphones;<br>Notebooks;<br>PC's; | xDSL;TLLx;<br>Wifi;Ethernet;<br>3G/4G/5G; | WhatsApp;<br>Instagram;<br>Portais;<br>YouTube; |
| <b>Descrição do caso:</b> Desinformação sobre aumento do auxílio reclusão                             |                                     |                                           |                                                 |
| <b>Fontes:</b> <a href="#">Instagram</a> ; <a href="#">Agora Notícias</a> ; <a href="#">TopNews</a> ; |                                     |                                           |                                                 |

#### 4.3.2.2 Etapa 1 - Identificar Partes Interessadas

Quase da mesma forma que na análise sociotécnica de objeto desinformativo, na estratégia de análise sociotécnica do ciclo o analista deve observar os diferentes objetos e canais de comunicação identificados para fazer o levantamento das partes interessadas. A Figura 4.22 exemplifica a exploração do objeto para identificar partes interessadas relacionadas ao seu conteúdo, e a Figura 4.23 apresenta um exemplo onde o analista considera alguns elementos dos canais de comunicação identificados no FRC para observar as partes interessadas.



Figura 4.22: Exploração do conteúdo do objeto para identificar Partes Interessadas.

| Formulário de Registro de Caso                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                                             |                                                                      |
|-------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------|----------------------------------------------------------------------|
| <b>Título do caso:</b>                                | Farsa - aumento auxílio reclusão                                                                                                                                                                                                                                                                                                                                                                                                                              | <b>Temática do caso:</b>                    | Saúde                                                                |
| <b>Data do caso:</b>                                  | 01/2023                                                                                                                                                                                                                                                                                                                                                                                                                                                       | <b>Tipo de objeto:</b>                      | Farsa                                                                |
| <b>Complexo de Elementos do Canal de comunicação:</b> | <b>Dispositivos</b>                                                                                                                                                                                                                                                                                                                                                                                                                                           | <b>Tecnologias</b>                          | <b>Plataformas</b>                                                   |
|                                                       | -Smartphones;<br>-Notebooks;<br>-PC's;                                                                                                                                                                                                                                                                                                                                                                                                                        | -xDSL;TLLx;<br>-Wifi;Ethernet;<br>3G/4G/5G; | - WhatsApp;<br>- Instagram;<br>- Portais Informativos;<br>- YouTube; |
| <b>Descrição do caso:</b>                             | Desinformação sobre suposto aumento no auxílio reclusão;                                                                                                                                                                                                                                                                                                                                                                                                      |                                             |                                                                      |
| <b>Fontes:</b>                                        | <a href="https://www.instagram.com/conservadoresbr/">https://www.instagram.com/conservadoresbr/</a><br><a href="https://agoranoticiasbrasil.com.br/">https://agoranoticiasbrasil.com.br/</a><br><a href="https://www.topnews.com.br/noticia/63539/governo-lula-publica-aumento-ao-auxilio-reclusao-e-passara-a-ser-de-r-15418-ma">https://www.topnews.com.br/noticia/63539/governo-lula-publica-aumento-ao-auxilio-reclusao-e-passara-a-ser-de-r-15418-ma</a> |                                             |                                                                      |

Figura 4.23: Possíveis fontes para observar partes interessadas no FRC preenchido.

A Figura 4.24 mostra um exemplo de exploração do contexto para identificar partes interessadas ocultas, como os controladores das plataformas digitais, e da infra-estrutura que da suporte a esta plataformas. Nessa figura, o site registro.br é consultado para identificar os responsáveis pelo domínio "agoranoticiasbrasil.com.br" e os controladores da infra-estrutura que hospeda o site<sup>10</sup>.

Whois

agoranoticiasbrasil.com.br

Exibir resultado completo

Copyright © NIC.br  
A utilização dos dados abaixo é permitida somente conforme descrito na Política de Privacidade, sendo proibida a sua distribuição, comercialização ou reprodução, em particular para fins publicitários ou propósitos similares.  
2023-01-23 11:56:21 -0300 - IP: 2001:1284f02e610edc15a026e4f07b2c

**Domínio agoranoticiasbrasil.com.br**

|                    |                               |
|--------------------|-------------------------------|
| TITULAR            | Rosemeire Domingues de Barros |
| DOCUMENTO          | 247.756.328-92                |
| PAÍS               | BR                            |
| CONTATO DO TITULAR | RODBA78                       |

Figura 4.24: Exploração do contexto para identificar Partes Interessadas- Ferramentas úteis.

A Figura 4.25 apresenta uma representação de algumas partes interessadas do caso em análise mapeadas no DPI, e a Figura 4.10 mostra o Formulário de Registro de Partes Interessadas.

<sup>10</sup>O site Registro.br é um exemplo de fonte de informação que oferece várias ferramentas para coleta de dados abertos sobre os domínios registrados com .br

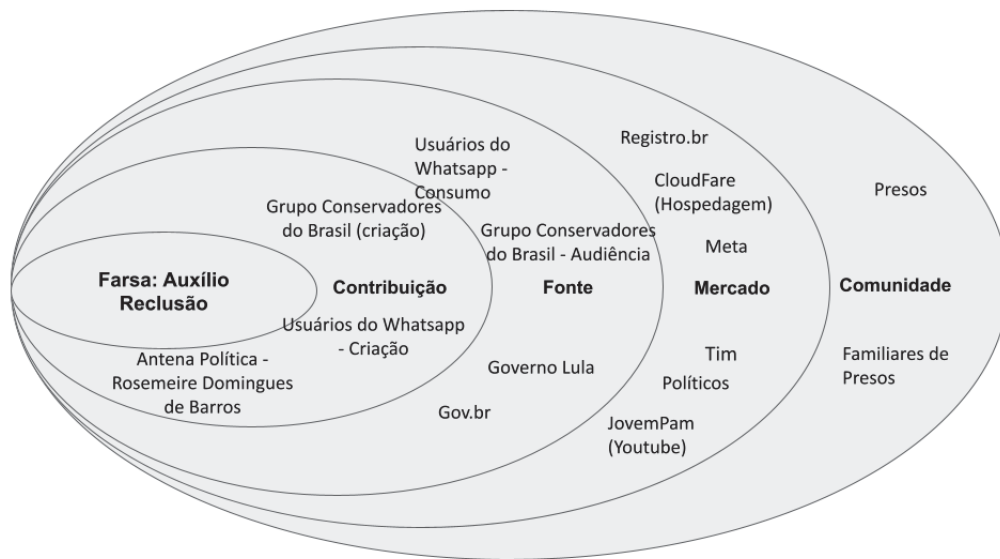


Figura 4.25: Partes Interessadas mapeadas no DPI.

Tabela 4.10: Formulário de Registro de Partes Interessadas.

| Formulário de Registro de Partes Interessadas |                                           |                         |                              |
|-----------------------------------------------|-------------------------------------------|-------------------------|------------------------------|
| Contribuição                                  | Fonte                                     | Mercado                 | Comunidade                   |
| Usuários WhatsApp (criação)                   | Grupo Conservadores do Brasil (audiência) | Meta                    | Pessoas presas               |
| Grupo Conservadores do Brasil (criação)       | Usuários WhatsApp (consumo)               | Tim (outras operadoras) | Famíliares de pessoas presas |
| Antena Política - Rosemaride Barros           | Governo Lula                              | Registro.br             |                              |
|                                               | gov.br                                    | Cloudfare (Hospedagem)  |                              |
|                                               |                                           |                         | JovemPam (Youtube)           |

#### 4.3.2.3 Final da Etapa 1

Ao final dessa etapa o analista terá identificado os elementos básicos para se conduzir a análise sociotécnica da desinformação digital:

- Ecossistema do fenômeno no caso: as partes interessadas em diferentes níveis de abstração e como estão dispostas em relação ao canal de comunicação e o objeto desinformativo;
- Estrutura da difusão: os elementos que constituem o complexo do canal de comunicação nos diferentes níveis técnico/tecnológicos;
- Linha do tempo do caso: os diferentes objetos desinformativos encontrados durante a exploração do contexto do caso.

Esses elementos seguem para Etapa 2 onde serão analisados em profundidade.

#### 4.3.2.4 Etapa 2 - Análise Sociotécnica do Ciclo

Nesta etapa, o analista emprega o artefato adaptado do AS para investigar o ciclo de vida do objeto desinformativo. O analista identificara aspectos técnicos, humanos, sociais e organizacionais que impactam nas mecânicas da comunicação de desinformação, amparado

por materiais de apoio e questões que guiam a análise. Ao final desta etapa o analista terá um conjunto de aspectos sociotécnicos agrupados por parte interessada e por estágio do ciclo de vida (criação, difusão e consumo). Esses dados seguem para a Etapa 3 de leitura, análise e discussão de resultados. A Figura 4.26 mostra uma representação visual de como a análise será feita.

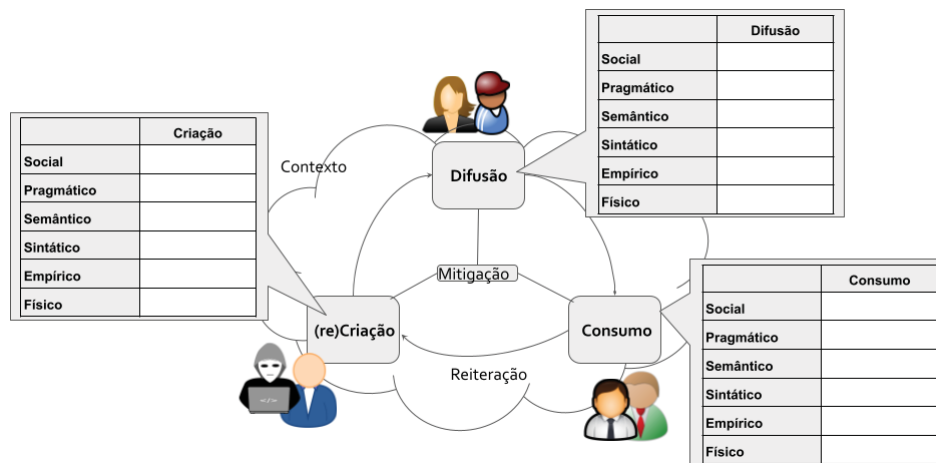


Figura 4.26: Representação da Análise sociotécnica do Ciclo de Vida da Desinformação usando a Escada Semiótica.

#### 4.3.2.5 Análise Sociotécnica do Estágio de Criação

A análise do estágio de criação visa identificar aspectos sociotécnicos relevantes na criação e divulgação inicial do objeto desinformativo. O analista deve considerar a perspectiva das partes interessadas envolvidas nesse estágio (criação), aspectos do canal de comunicação e do objeto desinformativo, buscando entender como eles podem estar relacionados. Nesse sentido, é importante adotar a perspectiva do criado ao realizar a análise de enquadramento e entender como o conteúdo foi criado, para quem foi criado, por que foi criado, o que o meio de difusão em que aparece impacta na criação e divulgação inicial, etc. Para guiar a identificação de aspectos relevantes, a Dismiss propõe as seguintes questões-guia:

- Quais aspectos sociotécnicos do objeto são relevantes na criação?
  - O analista observa o objeto, as partes interessadas e o contexto em busca de aspectos relevantes para cada parte interessada identificada. A Dismiss sugere buscar por indicadores de influência, como autoridades, proximidade afetiva, homofilia, relevância do tema/interesse pessoal.
  - Por exemplo, a relevância do conteúdo, a relação do conteúdo com o ambiente (temática, aparência, cores, etc.) e a popularidade, uso palavras ambíguas, frases que criam o senso de urgência ou induzem um senso de pertencimento, citações, uso gírias significativas para o público, conclusão e posicionamento, introdução de viés, a aparência do objeto, etc.
- Quais aspectos sociotécnicos do canal de comunicação são relevantes na criação?
  - O analista observa o canal de comunicação e busca por aspectos relevantes para cada parte interessada identificada.
  - Por exemplo, a popularidade do canal de comunicação (Facebook, Twitter, etc.), a quantidade de usuários de uma plataforma tecnológica em um contexto social



(Whatsapp no Brasil), aspectos de dispersão de mensagens nesse canal (multicast, unicast, etc;), aspectos de sincronicidade nesse canal (assíncrono ou síncrono), etc.

- Quais aspectos sociotécnicos do contexto estão relacionados com a intenção das partes interessadas na criação?
  - O analista observa o contexto e busca por aspectos relevantes para cada parte interessada identificada, como aspectos relacionados a relevância para cada partes, aspectos que indicam a intencionalidade de quem cria, aspectos que buscam passar credibilidade, etc.

Por exemplo, a escolha do público-alvo e de uma temática específica para um determinado caso, aspectos como a autoridade de quem cria/assina/endossa o conteúdo, o posicionamento ideológico de quem cria ou consome (quem e como ganha com o conteúdo?), pressupostos e conclusões no texto, assinaturas, carimbos.

As Figuras 4.27 e 4.28 apresentam um exemplo de identificação de aspectos sociotécnicos no estágio de criação.



Figura 4.27: Análise sociotécnica pela perspectiva da criação da desinformação - Sistema de Informação Humano.

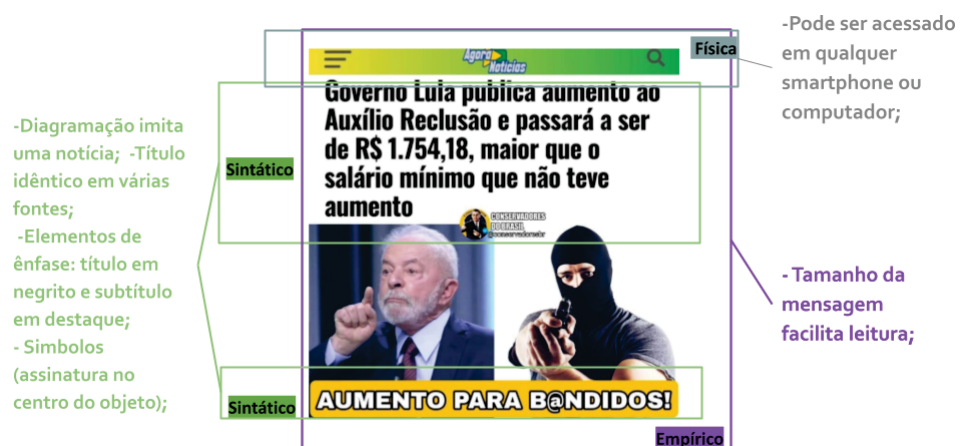


Figura 4.28: Análise sociotécnica pela perspectiva da criação da desinformação - Plataforma Tecnológica.

#### 4.3.2.6 *Análise Sociotécnica do Estágio de Difusão*

A análise do estágio de difusão visa identificar aspectos sociotécnicos relevantes na difusão objeto desinformativo. A análise investiga os recursos da mídia de comunicação, o conteúdo do objeto e a perspectiva das partes interessadas envolvidas na difusão. Nesse sentido, é importante adotar a perspectiva das partes interessadas envolvidas na difusão para realizar a análise de enquadramento e entender como o objeto se difunde, para onde se difunde, o que o meio de difusão restringe e facilita na sua propagação, os interesses de cada parte interessada, etc. Para guiar a identificação de aspectos relevantes nesse estágio (difusão), a Dismiss propõe as seguintes questões-guia:

- Quais aspectos sociotécnicos do objeto são relevantes na difusão?
  - O analista observa o objeto e busca por aspectos relevantes para cada parte interessada identificada. Por exemplo, se contém *hashtags*, menções e links para difusão rápida, se induz a divulgação com frases diretas, quantidade de caracteres, conteúdo puramente textual, se tem imagens/vídeos, etc.
- Quais aspectos sociotécnicos do canal de comunicação são relevantes na difusão?
  - O analista observa o contexto e busca por aspectos relevantes para cada parte interessada identificada, como aspectos relacionados a relevância do caso para cada parte, aspectos que indicam a intencionalidade de quem compartilha ou endossa a divulgação, aspectos que impactam a percepção de credibilidade, etc. Por exemplo, as políticas de uso da aplicação e dos serviços prestados, se existem mecanismos ativos de mitigação de difusão, se existem alertas para a audiência, tipos de planos de acesso a dados (tráfego gratuito para redes sociais, etc.), entre outros.
- Quais aspectos sociotécnicos do contexto de difusão estão relacionados com a intenção das partes interessadas?
  - O analista observa o canal de comunicação e busca por aspectos relevantes para cada parte interessada identificada. Por exemplo, as políticas de uso da aplicação e dos serviços prestados, se a legislação está sendo cumprida, se existem mecanismos ativos de mitigação da desinformação, a posição dos controladores do meio de comunicação quanto a mitigação da difusão, etc.

As Figuras 4.29 e 4.30 apresentam um exemplo de identificação de aspectos sociotécnicos no estágio de difusão.



Figura 4.29: Análise sociotécnica pela perspectiva da difusão da desinformação - Sistema de Informação Humana.

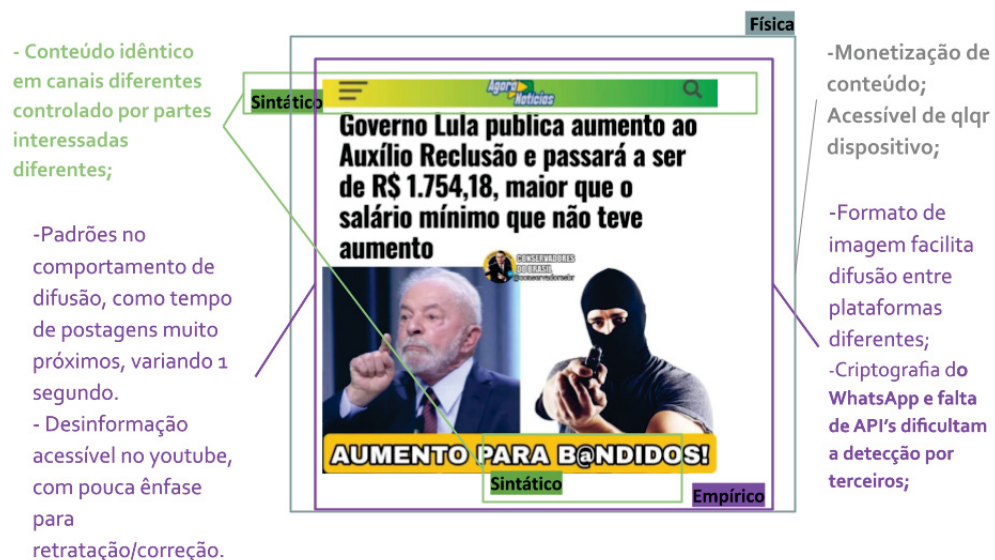


Figura 4.30: Análise sociotécnica pela perspectiva da difusão da desinformação - Sistema de Informação Humano.

#### 4.3.2.7 Análise Sociotécnica do Estágio de Consumo

A análise do consumo visa identificar aspectos sociotécnicos relacionados ao posicionamento de partes interessadas que consomem a desinformação, ao conteúdo do objeto e as capacidades do meio de comunicação. Nesse sentido, é importante adotar a perspectiva de quem consome a desinformação para realizar a análise de enquadramento e entender quais aspectos do objeto facilitam ou restringem o consumo, por quem é consumido, como é consumido, por que foi consumido, o que o meio de difusão impacta no consumo, etc. Para guiar a identificação de aspectos relevantes nesse estágio (consumo), a Dismiss propõe as seguintes questões-guia:

- Quais aspectos sociotécnicos do objeto são relevantes no consumo?
  - Por exemplo, a relação do conteúdo com o ambiente da audiência, os valores do grupo, as expectativas da audiência, os relacionamentos afetivos, como a audiência interpreta a autoridade da fonte de informação, etc.
- Quais aspectos sociotécnicos do canal de comunicação são relevantes no consumo?

- O analista observa o canal de comunicação e busca por aspectos relevantes, considerando as partes interessadas e a perspectiva do estágio de consumo. Por exemplo, a credibilidade do portal, a autoridade da fonte da mensagem, se o meio permite ver mensagens antigas (sincronicidade), se o meio permite a conversação, facilidade de compartilhamento e verificação, e custo de acesso.
- Quais aspectos sociotécnicos do consumo estão relacionados com as partes interessadas?
  - O analista observa o contexto e busca por aspectos relevantes para cada parte interessada identificada, como aspectos relacionados a relevância do caso para cada parte, aspectos que indicam a intencionalidade percebida, aspectos que impactam a percepção de credibilidade, etc. Por exemplo, alinhamento de posicionamento ideológico, se ocorreram discussões, compartilhamentos, alterações, curtidas, etc.

As Figuras 4.31 e 4.32 apresentam um exemplo de identificação de aspectos sociotécnicos no estágio de consumo.

-Assunto polêmico no contexto social gera engajamento com a desinformação e evoca comportamentos e atitudes prejudiciais (agressões, discurso de ódio, difamação, etc.);  
- Hábito de consumir informações sem verificar;

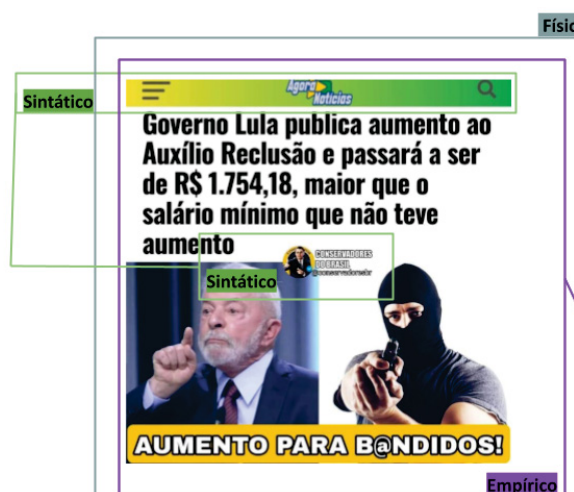
Significados diferentes para audiências diferentes: evocam sentimentos diferentes; Raciocínio motivado;



Expectativa da audiência de incompetência e corrupção;

Figura 4.31: Análise sociotécnica pela perspectiva do consumo da desinformação - Sistema de Informação Humano.

-Padrão de diagramação de notícia diminui a percepção de risco;  
-Consumo em diferentes meios de comunicação com restrições e controles;



-Pode ser acessado em qualquer smartphone ou computador;

-Dependendo do meio de consumo (site, twitter...), Poucos recursos de interação para verificação do conteúdo

Figura 4.32: Análise sociotécnica pela perspectiva do consumo da desinformação - Sistema de Informação Humana.

Para guiar a análise dos aspectos relevantes do estágio de criação, difusão e consumo o analista deve usar o Formulário de Análise Sociotécnica do Ciclo de Vida (FRASC - Figura 4.11). O FASC foi projetado para facilitar o registro dos aspectos sociotécnicos, mapeando a Escada

Semiótica juntamente com os Estágios do Ciclo de Vida da Desinformação em um mesmo artefato.

Tabela 4.11: Exploração do contexto com base no ciclo de vida - Busca em outros canais.

|                     | Criação                                                                                                                              | Difusão                                                                                                          | Consumo                                                                                                                        |
|---------------------|--------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|
| <b>Mundo Social</b> | Assunto de relevância no contexto social que é polarizado;<br>Lula é autoridade popular;                                             | Portal de conhecido por viés político e desinformação;<br>Hospedagem tem responsabilidade legal;                 | Assunto polêmico gera engajamento com a desinformação e evoca atitudes prejudiciais (agressões, discurso de ódio, difamação.); |
| <b>Pragmática</b>   | Conclusão enviesada com base em inverdade;<br>Objetivo parece ser difamar o Lula e seu Governo;                                      | Difusão de conteúdo viral é Interesse dos controladores dos canais;                                              | Expectativa da audiência de incompetência e corrupção;                                                                         |
| <b>Semântico</b>    | Foto contrasta Lula e um bandido armado;<br>Frase de efeito com Palavra-chave “B@NDIDOS” (estigmatizada)                             | Significados diferentes para audiências diferentes: evocam sentimentos diferentes; Raciocínio motivado;          | Grupos do WhatsApp com credibilidade para algumas audiência;                                                                   |
| <b>Sintático</b>    | Diagramação imita notícias; Título idêntico em várias fontes (origens);                                                              | Palavra bandidos aparece com alterações em canais diferentes; Imagens diferentes (ou nenhuma) em outros objetos; | Padrão de diagramação de notícia diminui a percepção de risco; consumo sem verificação;                                        |
| <b>Empírico</b>     | Formato de imagem facilita difusão entre plataformas; Criptografia do WhatsApp e falta de API's dificultam a detecção por terceiros; | Dependendo do meio de consumo (site, twitter) Poucos recursos de interação para verificação do conteúdo          | Formato de imagem facilita difusão;                                                                                            |
| <b>Mundo Físico</b> | Pode ser acessado em qualquer smartphone ou computador                                                                               | Monetização de conteúdo; Acessível de qlqr dispositivo;                                                          | Pode ser acessado em qualquer smartphone ou computador;                                                                        |

#### 4.3.2.8 Final da Etapa 2

Ao final da Etapa 2 teremos um caracterização dos estágios do ciclo:

- Criação
  - Contexto sociotécnico para criação/surgimento da desinformação;
  - Aspectos de posicionamento, oportunismo, polarização política, desconfiança, incerteza;
  - Origem do objeto e divulgação inicial;
  - Aspectos estéticos dos objetos parecem ter um padrão (possivelmente atrelado às partes interessadas que criam);
- Difusão
  - Características sociotécnicas da difusão;
  - Autoridades, alterações, caracterização dos canais de difusão;
  - Políticos, jornalistas, influenciadores digitais, pessoas físicas, etc;
  - Comportamento da difusão: tempo das postagens, transformações no formato e conteúdo;
  - Responsabilidades, mecanismos de mitigação e controle;
- Consumo
  - Características sociotécnicas do consumo;
  - Aspectos da vida diária da audiência que afetam o consumo;
  - Relações interpessoais, polarização política, desconfiança em fontes de informação tradicionais, etc;
  - Reações e repercussão: Agressões, discursos de ódio, difamação;
  - Significados para diferentes partes interessadas;



#### 4.3.2.9 Etapa 3 - Identificar Problemas

Nesta etapa o analista revê os aspectos sociotécnicos identificados e busca por suas inter-relações nos estágios do ciclo de vida da desinformação. Com estas informações o analista reflete sobre os aspectos mais relevantes do caso para identificar características do objeto desinformativo, do canal de comunicação e de cada estágio do ciclo que devem ser levadas em consideração para o desenvolvimento de intervenções no fenômeno. O analista considera a inter-relação entre os aspectos de três formas:

- **Entre partes interessadas:** O analista deve observar como aspectos em diferentes níveis de abstração para cada parte interessada podem estar inter-relacionados, e como podem impactar as partes interessadas. Por exemplo, quais interesses, expectativas, e propósito do caso? Como impactam a criação, difusão e consumo? Quais partes interessadas representam autoridades (Influenciadores, políticos, especialistas, etc.)? É possível identificar relações entre grupos de partes interessadas? Relações afetivas?
- **Entre estágios do ciclo:** O analista deve observar como aspectos em diferentes níveis de abstração podem estar inter-relacionados no ciclo de vida, e como podem impactar o ciclo de vida. Por exemplo, quais atividades de um estágio têm impacto no outro? Como a criação e o consumo estão relacionados? Como a difusão se relaciona com o consumo? Como o objeto foi divulgado, onde ele surgiu? O objeto foi re-compartilhado? Quais as transformações do objeto?
- **Entre dimensões semióticas:** O analista deve observar como aspectos em diferentes níveis de abstração podem estar inter-relacionados em um estágio, e como podem impactar o caso de desinformação. Por exemplo, quais aspectos da plataforma tecnológica afetam a representação ou interpretação? Como a temática/diagramação/tipo de objeto desinformativo se relaciona com o ambiente das partes interessadas? Como as referências e fontes de informação se relacionam com o conteúdo do objeto? São populares? Qual a relação entre o conteúdo e o posicionamento ideológico das partes interessadas?

Tabela 4.12: Formulário de Análise Sociotécnica preenchido.

|                     | Criação                                                                                                                              | Difusão                                                                                                          | Consumo                                                                                                                        |
|---------------------|--------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|
| <b>Mundo Social</b> | Assunto de relevância no contexto social que é polarizado; Lula é autoridade popular;                                                | Portal de conhecido por viés político e desinformação; Hospedagem tem responsabilidade legal;                    | Assunto polêmico gera engajamento com a desinformação e evoca atitudes prejudiciais (agressões, discurso de ódio, difamação.); |
| <b>Pragmática</b>   | Conclusão enviesada com base em inverdade; Objetivo parece ser difamar o Lula e seu Governo;                                         | Difusão de conteúdo viral é Interesse dos controladores dos canais;                                              | Expectativa da audiência de incompetência e corrupção;                                                                         |
| <b>Semântico</b>    | Foto contrasta Lula e um bandido armado; Frase de efeito com Palavra-chave “B@NDIDOS” (estigmatizada)                                | Significados diferentes para audiências diferentes; evocam sentimentos diferentes; Raciocínio motivado;          | Grupos do WhatsApp com credibilidade para algumas audiência;                                                                   |
| <b>Sintático</b>    | Diagramação imita notícias; Título idêntico em várias fontes (origens);                                                              | Palavra bandidos aparece com alterações em canais diferentes; Imagens diferentes (ou nenhuma) em outros objetos; | Padrão de diagramação de notícia diminui a percepção de risco; consumo sem verificação;                                        |
| <b>Empírico</b>     | Formato de imagem facilita difusão entre plataformas; Criptografia do WhatsApp e falta de API's dificultam a detecção por terceiros; | Dependendo do meio de consumo (site, twitter) Poucos recursos de interação para verificação do conteúdo          | Formato de imagem facilita difusão;                                                                                            |
| <b>Mundo Físico</b> | Pode ser acessado em qualquer smartphone ou computador                                                                               | Monetização de conteúdo; Acessível de qlqr dispositivo;                                                          | Pode ser acessado em qualquer smartphone ou computador;                                                                        |



Alguns problemas identificados na análise:

- Legislação e fiscalização são ineficazes para prevenir o surgimento de desinformação;
- Carência de mecanismos legais efetivos de ampla responsabilização;
- Permissão de monetização de portais com tradição em espalhar desinformação;
- A desinformação do caso pode ter surgido de modo acidental e se propagado sem controle;
- Jornalismo pobre, deficiência intelectual ou negligência na interpretação do ato publicado no diário oficial da união (originou o burburinho);
- Vieses podem ter sido introduzidos de propósito na notícia falsa com objetivo de monetizar acessos aos diferentes portais (des)informativos;
- Popularidade do assunto, ideologia da audiência, financiamento do portal, etc;
- A polarização política e o extremismo construídos pela cultura de “unidade” (nós contra eles) influenciam a repercussão em atitudes prejudiciais, como a agressão, o discurso do ódio e a difamação.

#### 4.4 CONSIDERAÇÕES SOBRE A DISMISS

A Dismiss foi proposta pensando em atender os requisitos identificados no mapeamento da literatura para abordar o fenômeno de modo abrangente por uma perspectiva sociotécnica. A Tabela 4.13 apresenta uma visão geral de como a Dismiss atende aos desafios de abordar a desinformação digital como um fenômeno sociotécnico.

Tabela 4.13: Requisitos para abordar a desinformação digital como um fenômeno sociotécnico.

| Requisitos                                             | Dismiss                                                                                                                                                                                                                                                                                                                                                                                                                   |
|--------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Apoia a análise de aspectos sociotécnicos              | A Dismiss estende o método de Análise de Enquadramento incorporando os artefatos da Escada Semiótica e Materiais de Apoio para estruturar e guiar o levantamento de aspectos sociotécnicos com a granularidade de seis níveis de abstração que vão do nível do Mundo Social ao Mundo Físico.                                                                                                                              |
| Apoia a identificação de inter-relações nos aspectos   | O Modelo Conceitual do Ciclo de Vida, juntamente com a Escada Semiótica e o Diagrama de Partes Interessadas proporcionam observar as inter-relações de aspectos sociotécnicos nos seis níveis da Escada, em cada estágio do ciclo de vida e para cada parte interessada identificada.                                                                                                                                     |
| Aborda a subjetividade das partes interessadas         | A Análise de Enquadramento permite abordar as diferentes perspectivas dos participantes de um caso de desinformação, estruturando a análise e a construção dos possíveis enquadramentos com que uma desinformação pode ser interpretada/entendida/observada.                                                                                                                                                              |
| Aborda a estrutura de um caso com estágios integrados  | O Modelo Conceitual do Ciclo de Vida representa a estrutura de um caso de desinformação, permitindo observar especificidades de cada estágio da vida de um desinformação, bem como os elementos em cada estágio do ciclo de vida se relacionam no caso como um todo.                                                                                                                                                      |
| Apoia a identificação de partes interessadas           | O Modelo Conceitual do Ciclo de Vida e o Diagrama de Partes Interessadas apoiam a exploração do contexto de um caso e a identificação de diferentes atores e suas atividades.                                                                                                                                                                                                                                             |
| Apoia a análise em profundidade e largura              | A estrutura modular da Dismiss torna flexível, apresentando duas formas diferentes de analisar um caso de desinformação. Fica a critério do analista empregar considera uma única parte interessada, em um único canal de comunicação interagindo com um único objeto desinformativo (Profundidade), ou considerando várias partes interessadas, vários canais de comunicação e vários objetos desinformativos (largura). |
| Apoia o estudo de desinformação deliberada e acidental | A Dismiss foi proposta com base em um conceito de desinformação que considera a perspectiva subjetiva de como diferentes partes interessadas entendem um mesmo objeto desinformativo, permitindo ao analista identificar aspectos sociotécnicos em cada caso para classificar o tipo de desinformação.                                                                                                                    |

## 5 AVALIAÇÃO E RESULTADOS ALCANÇADOS

A Dismiss foi avaliada de forma construtiva durante seu desenvolvimento usando o método de grupo focal, estudos em pequena escala, oficinas de análise sociotécnica de desinformação, e questionários inspirados no Modelo Estendido de Aceitação de Tecnologia - ETAM (*Extended Technology Acceptance Model*) (Fathema et al., 2015). O grupo focal avaliou o desenvolvimento da abordagem em diferentes momentos, contribuindo com melhorias para o método da pesquisa e para a solução proposta. Para os estudos de pequena escala, o autor empregou a Dismiss em análises de diferentes tipos de desinformação e contextos. Nesses estudos foram avaliados diferentes métodos de análise de conteúdo e artefatos em relação a sua capacidade de apoiar o estudo sociotécnico do fenômeno da desinformação. Os resultados informaram as escolhas dos componentes da Dismiss, suas etapas e melhorias para o Modelo do Ciclo de Vida. As oficinas serviram para avaliar a utilidade e facilidade de uso percebidas da Dismiss com o público-alvo. Os resultados das oficinas apontaram a capacidade da Dismiss para apoiar o entendimento na natureza sociotécnica do fenômeno da desinformação e melhorias para simplificar o uso da Dismiss.

### 5.1 AVALIAÇÃO QUALITATIVA POR GRUPO FOCAL

Como parte de uma metodologia de pesquisa construtiva, o método de grupo focal foi aplicado para avaliar o desenvolvimento da pesquisa em diferentes momentos. O objetivo foi buscar por opiniões críticas a respeito da capacidade de resolução de problemas como explicadas por Oulasvirta e Hornbæk (2016), considerando a significância, a efetividade e eficiência da solução. A significância da solução representa a importância da Dismiss e sua pertinência em resolver o problema de pesquisa (Oulasvirta e Hornbæk, 2016), sendo avaliada de modo qualitativo pelos estudos em pequena escala e pela percepção dos participantes das oficinas. A efetividade mede a abrangência de características do problema que a solução lida e sua capacidade (Oulasvirta e Hornbæk, 2016), sendo avaliada pelo grupo focal, estudos em pequena escala e pela percepção dos participantes das oficinas registradas pelos questionários. A eficiência considera os recursos necessários para sua aplicação e uso (Oulasvirta e Hornbæk, 2016), sendo avaliada pela percepção dos participantes das oficinas registradas pelos questionários e observações no uso dos artefatos da Dismiss<sup>1</sup>.

Foi utilizado a técnica grupo focal com um formato semiestruturado como referência para as discussões, onde o autor ou um colaborador serviam de moderadores. Os encontros ocorreram entre agosto de 2019 e maio de 2022. Os participantes do grupo focal são pesquisadores em treinamento no laboratório de pesquisa em Interação Humano-Computador (IHC) da Universidade Federal do Paraná. O tamanho do grupo variou entre 5-11 participantes por sessão, sendo composto por mestrandos e doutorandos. Os participantes foram codificados (P1, P2, ..., P11) e a coleta de dados foi realizada em 11 encontros, nos quais participaram os pesquisadores como avaliadores da pesquisa, o autor e um moderador. Os dados foram coletados por anotações e em alguns encontros por gravação multimídia.

As sessões foram planejadas de modo semiestruturado, onde a proposta do estudo era avaliação da pesquisa do autor. Os participantes foram instruídos a avaliar criticamente o

<sup>1</sup>O termos significância, eficiência e eficácia usados nesse capítulo se referem a perspectiva de Oulasvirta e Hornbæk sobre a capacidade de resolução de problemas (Oulasvirta e Hornbæk, 2016)

trabalho apresentado em relação a metodologia de pesquisa, ao problema e solução, e a forma. As atividades feitas em cada sessão:

- Os participantes assistiram a apresentações feitas pelo autor relatando o estado do seu trabalho em diferentes momentos do desenvolvimento da pesquisa.
- Os encontros duraram aproximadamente 120 minutos, sendo previstos 40 50 minutos de apresentação e o restante para avaliação e discussão da proposta pelos participantes.
- Após a apresentação, o moderador mediou o debate do grupo indagando as opiniões.

As anotações foram coletadas para cada participante e representam os apontamentos críticos. As gravações foram analisadas diretamente, utilizando uma abordagem informal inspirada na técnica de análise de conteúdo (Lazar et al., 2017). Os dados foram agrupados pelo tema: problema de pesquisa e solução proposta, o método de pesquisa, e a apresentação. Foram desconsiderados comentários redundantes e opiniões não relacionadas aos objetivos de cada estudo.

### 5.1.1 Resultados e discussões

As falas e anotações foram avaliadas observando as críticas construtivas para refinar melhorias. O Quadro 5.1 apresenta um resumo deste 11 encontros, apresentando a data e hora, tipo do encontro (presencial ou video-conferência), o conteúdo, o objetivo, e os principais apontamentos.

Tabela 5.1: Detalhamento dos encontros com o grupo focal.

| Data e Duração                                                  | Objetivo                                                                                                                                                                                                                  | Participantes                              | Principais apontamentos                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-----------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 30/08/2019 -<br>13:30hs<br>15:30hs<br>Presencial -<br>Anotações | Apresentar um ensaio sobre a desinformação como um fenômeno sociotécnico, buscando opiniões críticas para aprimorar o desenvolvimento de conceitos e fundamentos da pesquisa.                                             | 5<br>pesquisadores,<br>autor e<br>mediador | <ul style="list-style-type: none"> <li>• Definir um conceito claro para o termo "sociotécnico";</li> <li>• Resumir vantagens de abordar o fenômeno como sociotécnico;</li> <li>• Resumir como os indícios se relacionam com a perspectiva Semiótica Organizacional;</li> </ul>                                                                                                                                                                                                                                                                                   |
| 27/09/2019 -<br>10:00hs<br>12:00hs<br>Presencial -<br>Anotações | Apresentar aprimoramentos da perspectiva sociotécnica e a apresentação do " <i>position paper</i> " para o Interact 2020.                                                                                                 | 5<br>pesquisadores,<br>autor e<br>mediador | <ul style="list-style-type: none"> <li>• Explicar a motivação de usar a Semiótica Organizacional;</li> <li>• Padronizar os termos da fundamentação teórica;</li> <li>• Alguns slides tem muita informação;</li> <li>• Erros gramaticais e slides sem número;</li> </ul>                                                                                                                                                                                                                                                                                          |
| 25/03/2020 -<br>09:45 12:00hs<br>Presencial -<br>Anotações      | Apresentar a proposta de mapeamento da literatura, como os dados estão estruturados e como são sintetizados, buscando opiniões críticas para aprimorar o método de análise, a estruturação e apresentação dos resultados. | 9<br>pesquisadores,<br>autor e<br>mediador | <ul style="list-style-type: none"> <li>• Detalhes do protocolo, distraem o foco nos resultados;</li> <li>• Caracterizar os participantes;</li> <li>• Quais dados coletados para cada questão de pesquisa;</li> <li>• Explicar o método de análise;</li> <li>• Fluxograma não informa as atividades;</li> <li>• Falta resumo e síntese dos resultados (quais implicações?);</li> <li>• Gráficos inadequados para a qntd. de artigos retornados na busca ;</li> <li>• Conclusões repetitivas;</li> <li>• Erros gramaticais e alguns slides sem números;</li> </ul> |
| 29/04/2020 -<br>14:hs 17:00hs<br>Virtual -<br>Anotações         | Apresentar os materiais e métodos do estudo de caso da análise do WhatsApp, os dados coletados e resultados. O objetivo é obter opiniões críticas sobre metodologia do estudo e os resultados apresentados.               | 9<br>pesquisadores,<br>autor e<br>mediador | <ul style="list-style-type: none"> <li>• Caracterizar os participantes;</li> <li>• Mais ilustrações de problemas na interface do WhatsApp;</li> <li>• Sintetizar o contexto do estudo e focar nos resultados;</li> <li>• Explicar utilidade do estudo para a pesquisa.</li> <li>• Erros gramaticais e alguns slides sem numeração;</li> </ul>                                                                                                                                                                                                                    |
| 27/08/2020 -<br>16:hs 18:00hs<br>Virtual -<br>Anotações         | Apresentar os resultados do mapeamento sistemático e discutir oportunidades de pesquisa identificadas. O Objetivo é obter opiniões críticas sobre os dados e apresentação dos resultados do mapeamento.                   | 7<br>pesquisadores,<br>e autor;            | <ul style="list-style-type: none"> <li>• Apresentação extensa e sobrecarregada;</li> <li>• Detalhamento dos dados tira o foco que das discussões;</li> <li>• Melhorar a síntese dos resultados;</li> <li>• Fazer quadros para comparação dos estudos, pontos fortes e fracos;</li> <li>• Explicar as "poucas pesquisas sociotécnicas";</li> <li>• Erros gramaticais, alguns slides sem números;</li> </ul>                                                                                                                                                       |

Continua na próxima página

| Tabela 5.1 – continuação da página anterior                           |                                                                                                                                                                                                                                              |                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-----------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data e Duração                                                        | Objetivo                                                                                                                                                                                                                                     | Participantes                                | Principais apontamentos                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 19/11/2020 17hs<br>18:30hs<br>Virtual -<br>Anotações                  | Apresentar o estudo em pequena escala "Spearphishing RSA", os artefatos, métodos, dados coletados e resultados. O objetivo é obter opiniões críticas para refinar a construção abordagem para análise de desinformação.                      | 7<br>pesquisadores,<br>autor o<br>mediador   | <ul style="list-style-type: none"> <li>Foco em apresentar o todo antes de descer em profundidade;</li> <li>O método de análise Análise Semântica não parece ajudar;</li> <li>Verificar a necessidade de usar a Cebola Organizacional;</li> <li>Expandir o escopo de análise;</li> <li>A abordagem sobrecarregada, muitos conceitos, artefatos e formulários;</li> <li>Rever a pertinência dos Métodos de Análise Semântica e Modelagem Ontológica;</li> </ul>                                                                                                                                                                                                                                                                                   |
| 08/07/2021<br>15:hs 18:00hs<br>Virtual -<br>Anotações                 | Apresentar a formalização do modelo conceitual do ciclo de vida da desinformação, a caracterização dos estágios, os conceitos e formas de uso em busca de críticas para refiná-lo.                                                           | 8<br>pesquisadores,<br>o autor e<br>mediador | <ul style="list-style-type: none"> <li>Figura do ciclo confusa;</li> <li>O que é o Meta-ciclo, no ciclo de vida. Confunde a audiência;</li> <li>Não está clara a utilidade do modelo.</li> <li>Qual o significado de evento desinformativo e demais termos?</li> <li>Por que do evento, o que ele traz para análise?</li> <li>Menos foco na Semiótica Organizacional;</li> <li>Como aplicar a abordagem? Materiais de apoio?</li> <li>Por que o problema é um problema? Não está devidamente explicado;</li> <li>O que quer dizer sistêmico?</li> <li>Não está claro como o ciclo de vida ajuda a entender o fenômeno;</li> <li>O que quer dizer com "modelo para o fenômeno"? O que vai ser modelado? Qual a extensão desse modelo?</li> </ul> |
| 30/09/2021<br>15:00 17:00hs<br>Virtual -<br>Anotações                 | Apresentar o estudo em pequena escala "Farsa do pH", os artefatos, métodos, dados coletados e resultados. O objetivo é obter opiniões críticas sobre o estudo e os resultados apresentados.                                                  | 7<br>pesquisadores,<br>autor e<br>mediador   | <ul style="list-style-type: none"> <li>Muita informação, carga teórica é exaustiva;</li> <li>Conceituar desinformação deliberada e acidental pela perspectiva da Semiótica organizacional;</li> <li>A discussão da desinformação usando a escada semiótica é cansativa e muito detalhada;</li> <li>Não está claro como aspectos da dimensão física e empírica são desinformações;</li> <li>Exemplos não convenceram, parecem não relacionados;</li> <li>Conceito de interpretante não está claro;</li> </ul>                                                                                                                                                                                                                                    |
| 28/10/2011<br>15:00hs<br>17:18hs<br>Virtual -<br>Gravação de<br>vídeo | Apresentar a proposta de qualificação: motivação, problema, metodologia de pesquisa, proposta de solução, avaliação e cronograma. O objetivo é obter opiniões críticas sobre o trabalho desenvolvido, a metodologia, o problema e a solução. | 6<br>pesquisadores,<br>autor e<br>mediador   | <ul style="list-style-type: none"> <li>Muito foco em como a SO funciona;</li> <li>Peso da carga teórica da Semiótica Organizacional;</li> <li>Explicar o todo antes de descer em profundidade;</li> <li>Problema não está claro, objetivos não parecem relacionados;</li> <li>Ser mais objetivo;</li> <li>Muitos slides, slides sobrecarregados;</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                     |
| 07/04/2022<br>15:00hs<br>17:00hs<br>Virtual -<br>Anotações            | Apresentar a proposta de qualificação: motivação, problema, metodologia de pesquisa, proposta de solução, avaliação e cronograma. O objetivo é obter opiniões críticas sobre o trabalho desenvolvido, a metodologia, o problema e a solução. | 6<br>pesquisadores,<br>autor e<br>mediador   | <ul style="list-style-type: none"> <li>A descrição do problema muito genérica;</li> <li>Não da pra ligar objetivos com o problema;</li> <li>O público alvo da pesquisa precisa ser caracterizado;</li> <li>Não está claro como a abordagem pode apoiar o entendimento do fenômeno;</li> <li>A avaliação não está clara, o que avaliar, como e com quem (limitações na caracterização do público alvo);</li> </ul>                                                                                                                                                                                                                                                                                                                               |
| 05/05/2022<br>15:00hs<br>17:00hs<br>Virtual -<br>Gravação de<br>vídeo | Apresentar a proposta de qualificação: motivação, problema, metodologia de pesquisa, proposta de solução, avaliação e cronograma. O objetivo é obter opiniões críticas sobre o trabalho desenvolvido, a metodologia, o problema e a solução. | 5<br>pesquisadores,<br>e autor;              | <ul style="list-style-type: none"> <li>Estudo com grupo focal está mal caracterizado;</li> <li>Foi Estruturado/semiestruturado/livre?</li> <li>Quem são os participantes?</li> <li>Sintetizar as contribuições.</li> <li>Clarificar o objetivo, questões, e a coleta de dados;</li> <li>Alguns slides estão sobrecarregados de informações;</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                          |

Nos primeiros estágios da pesquisa, as críticas e opiniões do grupo focal auxiliaram na clarificação do problema de pesquisa e no desenvolvimento da fundamentação teórica. Por exemplo, na sessão de 30/08/2019 foi apresentada a motivação para se adotar uma perspectiva sociotécnica ao investigar o fenômeno. Nesta sessão, o participante P1 critica a falta de esclarecimento para o termo "sociotécnico" que fundamenta o estudo. P1 comenta que diferentes autores apresentam diferentes entendimentos para sociotécnico e que é preciso esclarecer a perspectiva adotada pela pesquisa do autor. No dia 27/09/2022, ao apresentar o refinamento da perspectiva sociotécnica, P5 indagou a motivação para adotar a Semiótica Organizacional (SO).

Estes e outros apontamentos auxiliaram a clarificar o problema e a motivação de se adotar a SO, resultando em um artigo que detalha esse posicionamento<sup>2</sup>.

As sessões dos dias 25/03/2019 e 27/08/2020 trataram do mapeamento sistemático da literatura. Nas primeiras sessões o foco das críticas foi direcionado para as motivações e necessidades para realizar o mapeamento, limitações do protocolo proposto, e a apresentação dos dados e resultados preliminares. As críticas retrataram insuficiências do protocolo, e da estruturação e análise dos dados que informaram melhorias para o desenvolvimento do mapeamento. Por exemplo, no dia 25/03/2019, P1 apontou a necessidade de exemplos dos tipos de dados para cada campo dos formulários de extração de dados, e P5 ressaltou a necessidade caracterizar os participantes e explicar o processo de filtragem a ser adotado.

Na sessão de 27/08/2020, o foco das discussões foi a estruturação dos dados e apresentação de discussões e resultados finais do mapeamento. Nesta sessão, P2 sugeriu melhorias para facilitar a apresentação dos dados e o entendimento da audiência. Por exemplo, construir quadros para sintetizar os resultados mais relevantes da comparação das pesquisas facilitaria para a audiência visualizar as limitações da literatura. Além disso, P3 sugeriu resumir os principais aspectos de cada linha de pesquisa em um mapa mental, P7 falou para apontar vantagens e desvantagens, e P5 pediu para relacionar como a perspectiva sociotécnica pode auxiliar a lidar com as desvantagens. Estes e outros apontamentos permitiram aprimorar o protocolo de mapeamento, a apresentação dos dados e posicionar a pesquisa do autor na literatura.

A sessão de 29/04/2020 tratou da apresentação de um estudo de caso onde os recursos de interação do WhatsApp foram analisados por especialistas em IHC aplicando um método de inspeção orientado a valores<sup>3</sup>. Os apontamentos auxiliaram a clarificar a relação dos resultados do estudo com a pesquisa do autor e focar nos aspectos mais relevantes. Por exemplo, P5 sugeriu sintetizar a descrição do contexto do estudo e dedicar mais tempo apresentando os resultados e discussões. P2 indagou sobre como o estudo de caso se relaciona com a pesquisa do autor, e comentou sobre a importância dessa discussão. Os apontamentos levaram a discussões de como o estudo revelou partes interessadas e aspectos sociotécnicos relevantes. No estudo, os participantes relataram limitações do WhatsApp, como o rótulo pouco significativo para "mensagem encaminhada", a falta de mecanismos ativos para verificar conteúdo, e indícios do impacto de valores na percepção de credibilidade dos participantes. Nesse sentido, o estudo auxiliou no desenvolvimento da Dismiss, melhorando a compreensão da relação entre aspectos sociais e técnicos, da complexidade dos elementos que compõem o canal de comunicação, e a relevância em considerar a perspectiva das partes interessadas de um caso de desinformação.

As sessão de 08/07/2021 tratou do modelo conceitual do ciclo de vida da desinformação. Nessa sessão, os participantes discutiram aspectos relacionados à motivação e necessidade de uma abordagem sistêmica e estruturada, aspectos sobre a fundamentação teórica e os conceitos do modelo, além de apontamentos e considerações que resultaram no refinamento da estrutura e da figura do ciclo de vida. O autor apresentou a motivação para criação do modelo conceitual, a caracterização de cada estágio do ciclo de vida usando o método de Análise Semântica e os conceitos de objeto desinformativo, desinformação acidental e deliberada, e a estrutura de um caso de desinformação. P1 fez apontamentos sobre a necessidade de redefinir o problema de pesquisa para esclarecer a necessidade da estrutura do modelo. P5 e P6 comentaram sobre a dificuldade de entender como o modelo conceitual auxilia no entendimento do fenômeno da desinformação, especialmente a carga teórica necessária para entender o modelo. Esses e outros apontamentos resultaram em melhorias para simplificar o modelo e refinar a descrição do problema de pesquisa, em especial na síntese das características de cada estágio e dos conceitos

<sup>2</sup>"Dismiss false information: A value matter- pode ser encontrado nos Apêndices A.5

<sup>3</sup>O artigo pode ser encontrado nos Apêndices A.5.



definidos no modelo. Os apontamentos auxiliaram a esclarecer elementos relevantes do modelo, como a sugestão de P1 de rever o meta-ciclo da desinformação. A Figura 5.1 apresenta a evolução do modelo conceitual do ciclo de vida da desinformação.

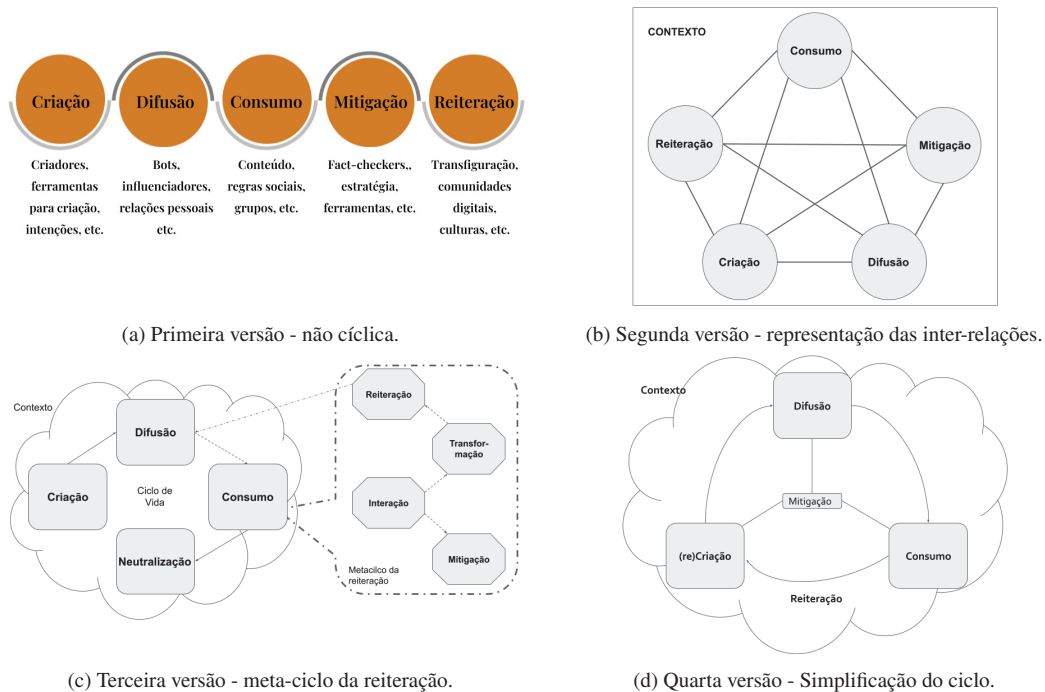


Figura 5.1: Evolução da representação visual do ciclo de vida da desinformação.

As sessões de 08/07/2021, 28/10/2021, 07/04/2022 e 05/05/2022 trataram da Dismiss. Nas sessões de 2021, os primeiros apontamentos criticavam a carga teórica necessária para se entender a abordagem, a estruturação dos passos para aplicação da abordagem e a significância da abordagem para o problema de pesquisa. Por exemplo, na sessão de 08/07/2021, P1 explica que o foco dado aos fundamentos da SO desviam a atenção da proposta e sugere simplificar a carga teórica da abordagem. Na reunião de 28/10/2021 P5 sugeriu melhorias para o esquema de aplicação geral da abordagem, indagando a utilidade e necessidade de todos os formulários de registro de dados da análise. A sessão de 05/05/2022 tratou do estado da tese e da apresentação para a qualificação. Nessa sessão, foram sugeridas melhorias para sintetizar a apresentação e clarificar as relações dos resultados de cada etapa da pesquisa. Estes e outros apontamentos foram fundamentais para lidar com as limitações de eficiência (Oulasvirta e Hornbæk (2016)) da abordagem, e levaram a simplificação nos métodos de análise da Dismiss.

As sessões de 19/11/2020 e 30/09/2021 trataram dos estudos em pequena escala que analisaram, respectivamente, o caso de "spearphishing" na RSA e uma Farsa sobre pH no WhatsApp. Na sessão de 19/11/2020, P1 e P6 criticaram o uso do método de Análise Semântica e Modelagem Ontológica, o fato destes trazerem mais complexidade do que já existia para o entendimento e uso da abordagem, além de que a abordagem não deixava claro o que fazer com os dados que a Análise Semântica revela. Na sessão de 19/11/2020, os participantes fizeram questionamentos e apontamentos relevantes sobre os resultados dos estudos, identificando inter-relações de aspectos sociotécnicos e discutindo os seus impactos. Foram discutidas melhorias para o Modelo do Ciclo de Vida, além do potencial de se expandir o escopo da análise para considerar o estágio de reiteração da desinformação, rastreando suas transformações em um estudo longitudinal. O autor optou por expandir o escopo da análise e acompanhar o caso por mais tempo — aproximadamente 2 anos.



Na primeira reunião de 2022, percebeu-se que o autor estava trabalhando em dois problemas simultâneos, abordar a desinformação como fenômeno sociotécnico e a análise do impacto de aspectos sociotécnicos. As discussões focaram nas dificuldades encontradas, tanto do autor para explicar quanto dos participantes em entender a análise de aspectos sociotécnicos de nível Pragmático na Escada Semiótica (ES), e como relacioná-los a aspectos Sociais de modo simplificado sem perder o rigor<sup>4</sup>. Isso levou a delimitação do escopo de operação da Dismiss, e a prorrogação da pesquisa sobre análise dos impactos de aspectos sociotécnicos para trabalhos futuros. Após adaptações, as últimas seções com o grupo focal revelaram os aspectos positivos da mudança de perspectiva quanto a abordagem dos estágios do ciclo de vida, como o entendimento claro do problema e solução, além do foco dos apontamentos da seção ter mudado para outros aspectos da pesquisa, como a avaliação da abordagem, e a criação do manual e do tutorial.

## 5.2 ESTUDOS EM PEQUENA ESCALA

Essa seção apresenta um resumo dos resultados da aplicação da abordagem em estudos de pequena escala realizados pelo autor<sup>5</sup>. Foram avaliados sete casos de desinformação, incluindo tipos de objetos desinformativos como notícias falsas, teorias da conspiração, e farsas (Figura 5.2). As avaliações foram realizadas com dois objetivos: primeiramente avaliar a viabilidade de métodos e artefatos para análise sociotécnica da desinformação digital, e depois para refinamento dos componentes selecionados para o desenvolvimento da abordagem Dismiss<sup>6</sup>.

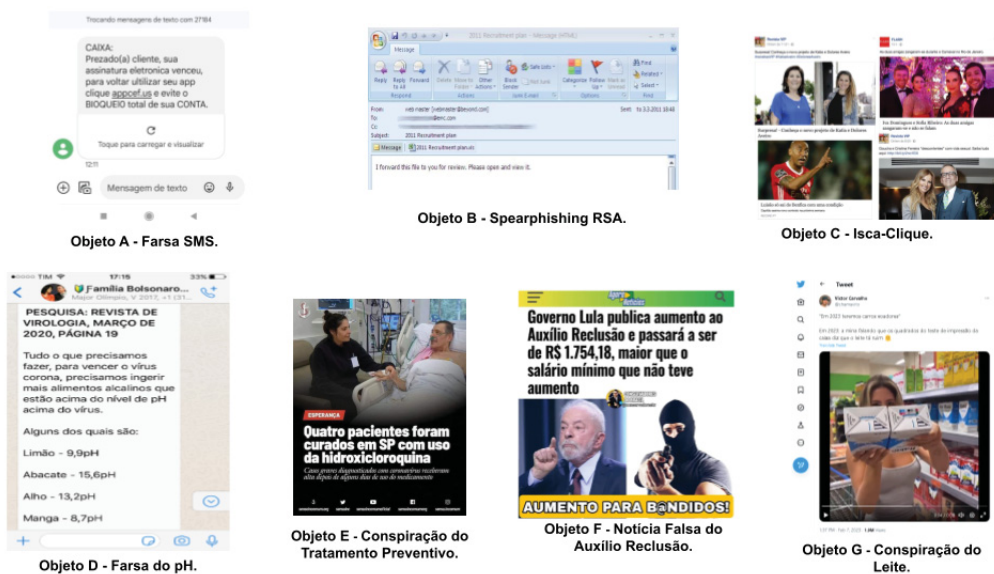


Figura 5.2: Coleção de objetos desinformativos analisados pelo autor.

<sup>4</sup>Mais sobre esse desdobramento no Apêndice A.2.

<sup>5</sup>Segundo (Robson, 2017), estudos em pequena escala são adequados para prova de conceito e avaliação de viabilidade, sendo geralmente conduzidos por um único pesquisador.

<sup>6</sup>Detalhes das análises podem ser encontradas no Apêndice A.7.1.1

Tabela 5.2: Resumo dos Estudos em Pequena Escala

| Objeto                                | Data                                                              | Objetivos                                                                                                                                                                                                                    | Resultados                                                                                                                                                                                                                                                                                                                                                                      | Apontamentos (Grupo Focal)                                                                                                                                                                                                                                                                                           |
|---------------------------------------|-------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Objeto A<br>Farsa no SMS              | Início:<br>12/04/20<br>Final: 25/04/20<br>Grupo Focal: s/         | - Investigar uma abordagem para análise sociotécnica de desinformação;<br>- Explorar os artefatos Cebola Organizacional, Escada Semiótica; e Diagrama de Identificação de Partes Interessadas;                               | - Todos os artefatos testados foram selecionados;<br>- Necessidade de estruturar a análise;<br>- Determinação dos elementos básicos de um caso;<br>- Características das Partes interessadas;<br>- Canal de comunicação é um complexo de elementos;<br>- Caracterização do tipo de desinformação "Farsa";                                                                       | -                                                                                                                                                                                                                                                                                                                    |
| Objeto B<br>Spearphishing RSA         | Início:<br>03/06/20<br>Final: 20/06/20<br>Grupo Focal: 19/11/2020 | - Investigar uma abordagem para análise sociotécnica de desinformação;<br>- Refinar os artefatos da SO;<br>- Explorar os métodos de Análise Semântica e Modelagem Ontológica;<br>- Discutir melhoramentos com o grupo focal; | - Refinamento no uso dos artefatos da SO;<br>- Refinamento dos conceitos de desinformação.<br>- Definidas etapas de exploração do contexto e análise semiótica;<br>- Criação de "diretrizes" para identificação de partes interessadas;<br>- Criação de "diretrizes" para análise do canal de comunicação;<br>- Refinamento da caracterização da desinformação do tipo "Farsa"; | - Necessidade de simplificação da abordagem;<br>- Descarte do artefato Cebola Organizacional;<br>- Descarte dos métodos de análise selecionados;<br>- Necessidade de simplificação dos conceitos;<br>- Refinar etapas para guiar a análise.                                                                          |
| Objeto C<br>Isca-clique               | Início:<br>24/04/20<br>Final: 02/05/20<br>Grupo Focal: s/         | - Refinar os artefatos da SO;<br>- Explorar o método de Análise de Enquadramento;<br>- Refinar os passos para operacionalizar a Dismiss;                                                                                     | - Refinamento das etapas da Dismiss;<br>- Modelo Conceitual do Ciclo de Vida (estrutura do caso);<br>- Refinamento de diretrizes de análise sociotécnica;<br>- Refinar a caracterização dos estágios do Ciclo de Vida;<br>- Caracterização do tipo de desinformação "Isca-Click";                                                                                               | -                                                                                                                                                                                                                                                                                                                    |
| Objeto D<br>Farsa do pH.              | Início:<br>07/05/20<br>Final: 13/09/22<br>Grupo focal: 30/09/2021 | - Refinar a aplicação da Dismiss;<br>- Formalizar as etapas da Dismiss;<br>- Discutir melhoramentos com o grupo focal;                                                                                                       | - Refinamento no uso dos artefatos da SO;<br>- Refinar as etapas da Dismiss;<br>- Caracterização dos estágios do ciclo;<br>- Diretrizes viraram questões-guia e exemplos;<br>- Refinamento da caracterização do tipo de desinformação "Farsa";                                                                                                                                  | - Simplificar a aplicação da Dismiss;<br>- Melhorias no Modelo Conceitual;<br>- Criar materiais de apoio para aplicação da Dismiss;<br>- Discussões de inter-relações de aspectos sociotécnicos;<br>- Ampliação do escopo de análise: Análise da Reiteração;<br>- Discussões dos impactos de aspectos sociotécnicos; |
| Objeto E<br>Conspiração da Cloroquina | Início:<br>11/06/21<br>Final: 20/06/21<br>Grupo Focal: s/         | - Refinar a aplicação da Dismiss;<br>- Explorar eixos complementares para sustentar a análise;                                                                                                                               | - Simplificação das etapas e formulários da Dismiss;<br>- Questões-guia para guiar a análise de inter-relações;<br>- Criação do manual de uso e de um passo-a-passo;<br>- Caracterização da desinformação do tipo "Teoria da Conspiração";                                                                                                                                      | -                                                                                                                                                                                                                                                                                                                    |
| Objeto F<br>Notícia Falsa Reclusão*   | Início:<br>20/01/23<br>Final: 25/01/23<br>Grupo Focal: s/         | - Refinar a aplicação da Dismiss;<br>- Refinar materiais de apoio;<br>- Refinar o manual da Dismiss;                                                                                                                         | - Simplificação do passo-a-passo e o manual da Dismiss;<br>- Criação do vídeo-tutorial de aplicação da Dismiss;<br>- Caracterização da desinformação do tipo "Notícia Falsa";                                                                                                                                                                                                   | -                                                                                                                                                                                                                                                                                                                    |
| Objeto G<br>Conspiração do Leite**    | Início:<br>13/02/23<br>Final: 16/01/23<br>Grupo Focal: s/         | - Refinar a aplicação da Dismiss;<br>- Refinar materiais de apoio;<br>- Criar vídeo-tutorial;                                                                                                                                | - Simplificação do passo-a-passo e o manual da Dismiss;<br>- Criação do vídeo-tutorial de aplicação da Dismiss;<br>- Caracterização da desinformação do tipo "Teoria da Conspiração";                                                                                                                                                                                           | -                                                                                                                                                                                                                                                                                                                    |

\* Um vídeo-tutorial demonstrando à análise completa pode ser encontrado em: <https://youtu.be/hWyjB9-KIE>.\*\* Um vídeo-tutorial demonstrando à análise completa pode ser encontrado em: <https://youtu.be/NzWYrbrLjfc>.

Os primeiros estudos foram prospectivos para identificar as capacidades dos artefatos da Escada Semiótica (ES), Cebola Organizacional (CO)<sup>7</sup>, e o Diagrama de Identificação de Partes Interessadas (DPI) da SO em apoiar a análise sociotécnica do fenômeno. Como na literatura revisada (Capítulo 3), foi identificada uma lacuna referente a metodologia para análise sociotécnica da desinformação, capaz de caracterizar o contexto de um caso de desinformação de modo abrangente (auxiliar na identificação dos elementos básicos e na caracterização do contexto do caso), especialmente no que diz respeito a identificação de partes interessadas e aspectos sociotécnicos em diferentes níveis de abstração, esses artefatos foram escolhidos pelo seu propósito de uso.

O autor investigou a viabilidade de uso dos métodos de Análise Semântica e Modelagem Ontológica (Liu, 2000)<sup>8</sup>, oriundos do Método de Articulação de Problemas (Stamper, 1994). Os primeiros estudos (Casos: Farsa do SMS e Spearphishing na RSA) foram planejados de modo semi-estruturado, onde o autor testou em um mesmo objeto desinformativo diferentes artefatos e técnicas de análise, coletou os dados do estudo em formulários e anotações, e avaliou a granularidade de aspectos sociotécnicos e a complexidade conceitual e de uso dos artefatos. Os resultados foram levados para discussão com o grupo focal. As análises apontaram para a utilidade do DPI em apoiar o levantamento de partes interessadas, e a efetividade da Escada Semiótica e da CO em identificar aspectos sociotécnicos no contexto do caso analisado. Os seis níveis de análise da Escada Semiótica possibilitaram maior granularidade de aspectos sociotécnicos identificados do que os 3 níveis da CO, ao custo de maior complexidade conceitual e de trabalho manual. Essa avaliação informou a decisão do autor em escolher o DPI e a Escada Semiótica para integrarem a abordagem Dismiss, além de levar ao desenvolvimento de formulários para apoiar análise e armazenar os dados coletados.

No segundo estudo, foi analisado um objeto desinformativo do tipo farsa no WhatsApp que levou à análise detalhada do Modelo Conceitual do Ciclo de Vida da Desinformação. O autor investigou a viabilidade de uso da Análise de Enquadramento (Hamborg et al., 2017) para análise de narrativas. Os dados das análises foram coletados em formulários, e os resultados foram apresentados e discutidos com o grupo focal. Os achados nas análises indicaram a utilidade e facilidade de uso da Análise de Enquadramento (AE). Essa avaliação informou a decisão do autor em escolher a Análise de Enquadramento para compor a Dismiss.

Em relação às outras análises, o autor empregou os artefatos e métodos selecionados para analisar outros objetos e contextos em busca de desenvolver os passos e estruturar a Dismiss, com relação aos artefatos ES, DPI, dos materiais de apoio, do Modelo Conceitual do Ciclo de Vida da Desinformação e do método de Análise de Enquadramento. Os dados das análises foram coletados em formulários. Os resultados informaram aprimoramentos para o Modelo Conceitual do Ciclo de Vida da Desinformação, detalhando atividades, partes interessadas e aspectos sociotécnicos em cada caso. Ainda, com base nas avaliações dos resultados pelo grupo focal feitos após os estudos de "Spearphishing" na RSA e Farsa do pH, foram simplificados os conceitos de caso de desinformação e a caracterização dos estágios do ciclo de vida. Além disso, os resultados da aplicação levaram ao desenvolvimento de um tutorial para utilização da DISMISS, das perguntas para guiar o levantamento de aspectos sociotécnicos, dos materiais de apoio para análise sociotécnica do caso de desinformação e dos formulários de síntese e registro dos dados da análise.

São apresentados a seguir os dois estudos em pequena escala que foram discutidos com o grupo focal, e que tiveram maior influência nas decisões do autor sobre os componentes para a

<sup>7</sup>A fundamentação teórica sobre a CO pode ser encontrada no Apêndice A.1.5.

<sup>8</sup>Mais sobre os métodos e os resultados da aplicação estão no Apêndice A.6

Dismiss. Além disso, o caso de phishing na RSA também foi escolhido por mostrar o potencial da Dismiss para ser aplicado em outros contextos, como na análise forense de ataques cibernéticos.

### 5.2.1 Análise do esquema de *phishing* - Farsa na RSA

Na análise de Fraude na RSA, o autor avaliou o uso dos artefatos Escada Semiótica, Cebola Organizacional e Diagrama de Partes Interessadas, considerando diferentes técnicas para abordar o contexto: o método de Análise Semântica e a Modelagem Ontológica<sup>9</sup>. Os métodos são usados para identificar e modelar elementos, procedimentos e aspectos inter-dependentes em atividades organizadas (Liu, 2000), por exemplo, como uma atitude ou atividade depende de propriedades específicas em um meio, ou do encadeamento de atividades. A Cebola Organizacional é um artefato da SO empregado na análise de organizações (Liu, 2000) e foi usada na Etapa de Clarificação do Escopo da DISMISS para apoiar a coleta de informações do contexto do caso e das partes interessadas<sup>10</sup>. Ainda, a análise sociotécnica foi feita colocando em evidência a perspectiva de cada parte interessada em cada estágio do ciclo de vida. Os resultados da aplicação dos artefatos permitiram compreender aspectos sociotécnicos diretamente relacionados às partes interessadas para cada estágio do ciclo de vida. Os resultados da aplicação da Cebola Organizacional ampliaram o entendimento do contexto e de aspectos relevantes das partes interessadas, complementando a análise. Os resultados da aplicação da Análise Semântica e Modelagem Ontológica permitiram maior clarificação da inter-relação de aspectos sociotécnicos e como impactam no ciclo de vida da desinformação. No entanto, a avaliação por grupo focal da Dismiss revelou limitações de eficácia e eficiência em relação à carga teórica para compreensão da DISMISS e a complexidade de aplicação desses métodos. Os resultados levaram ao abandono destes métodos de análise e da Cebola Organizacional, levando a adoção do método de Análise de Enquadramento e a simplificação da DISMISS em sua estrutura apresentada no Capítulo 4.

#### 5.2.1.1 A Farsa na RSA

Em 2011, a RSA Security era a divisão de segurança da corporação EMC, criando produtos de rede, criptografia e segurança computacional (Westervelt, 2011). Em março de 2011 a empresa sofreu um ataque de *spear phishing* (Zetter, 2011). Este tipo de ataque usa a desinformação para persuadir indivíduos especialmente selecionados a executar uma ação aparentemente inofensiva, mas que causará prejuízos (Caputo et al., 2013). Nesse caso, um grupo de hackers selecionou quatro empregados sem importância na hierarquia da empresa como alvos. A mensagem falsa foi construída usando o e-mail como mídia digital, contendo o assunto "Plano de Recrutamento 2011" e um arquivo do Excel com um *malware* oculto. O e-mail foi detectado como spam pela infraestrutura de segurança da empresa e direcionado para o lixo eletrônico. Um dos quatro empregados que receberam o e-mail abriu o arquivo do Excel, infectando o seu terminal. Com o acesso, os hackers usaram o terminal para escalar os privilégios na infraestrutura e roubar dados dos clientes (Mikko, 2021). A Figura A.16 mostra mensagem.

<sup>9</sup>Explicados em detalhes no Apêndice A.7.1.1

<sup>10</sup>Este artefato é explicado nos Apêndices A.1.5. Além de ser usado na avaliação, a CO foi usada para caracterização dos estágios do ciclo de vida com base nos dados levantados pelo mapeamento da literatura.

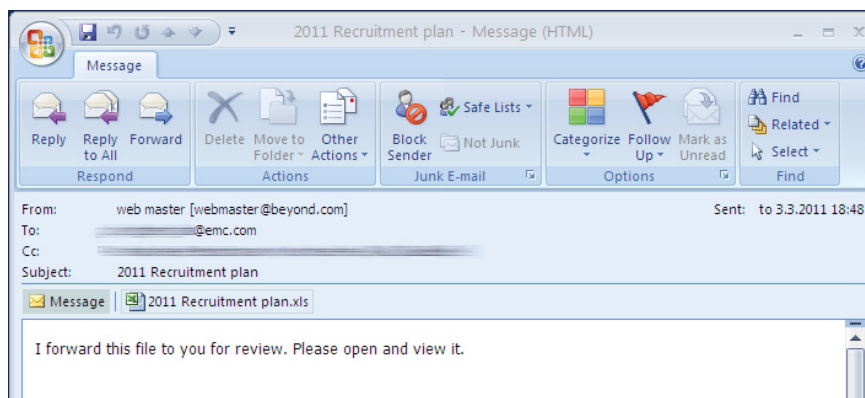


Figura 5.3: Representação do objeto desinformativo(phishing). Fonte: F-Secure (Mikko, 2021).

A familiarização (Etapa 1 da DISMISS) foi feita com base em documentos oficiais que detalharam o ataque (Rivner, 2011; Coviello, 2021; RSA, 2021), notícias em portais informativos tradicionais (Westervelt, 2011; Zetter, 2011) e relatórios de empresas de segurança (Mikko, 2021). Identificamos 13 partes interessadas para o contexto do caso de desinformação. As partes interessadas foram divididas em dois grupos principais, os interessados em manipular (atacantes), e os relacionados à organização (RSA). Figura A.17 mostra o mapa de partes interessadas organizadas no Diagrama de Identificação de Partes Interessada (DPI).

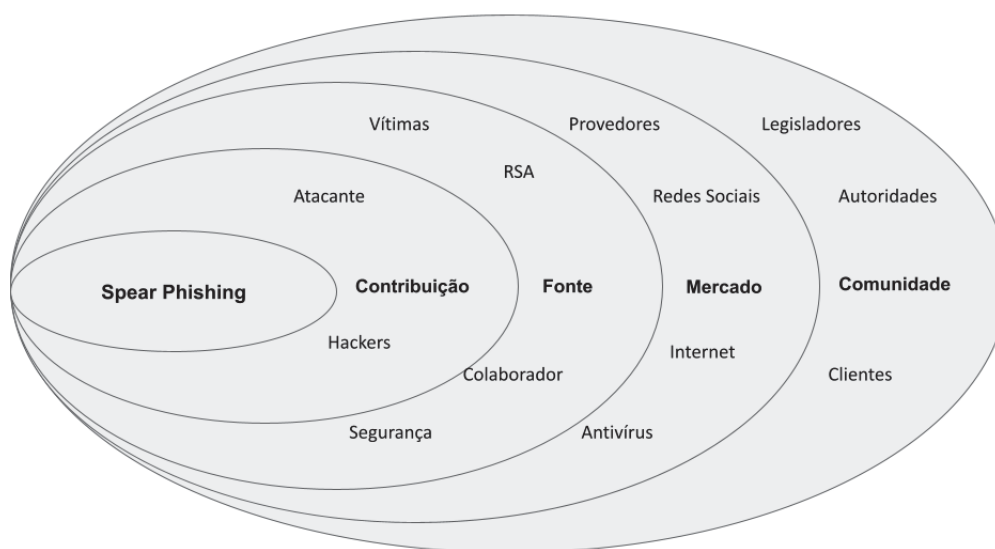


Figura 5.4: Partes interessadas (RSA Phishing)

Foram identificados 46 aspectos sociotécnicos, 29 do sistema de informação humana e 17 da plataforma tecnológica. Os dados da análise foram registrados nos formulários propostos para este fim<sup>11</sup>.

#### 5.2.1.2 Análise Semiótica do objeto desinformativo

O objeto desinformativo não era sofisticado, com uma estrutura simples e um texto impessoal, que apelava para a curiosidade do colaborador, levando-o a abrir um arquivo infectado. Os mecanismos técnicos de segurança detectaram o e-mail como possível spam e encaminharam

<sup>11</sup>[https://docs.google.com/spreadsheets/analise\\_spear](https://docs.google.com/spreadsheets/analise_spear).



a mensagem para caixa de spam. Um dos quatro colaboradores que recebeu o e-mail retirou a mensagem da caixa de spam e a abriu. A Figura 5.5 apresenta alguns dos aspectos sociotécnicos identificados no objeto desinformativo.

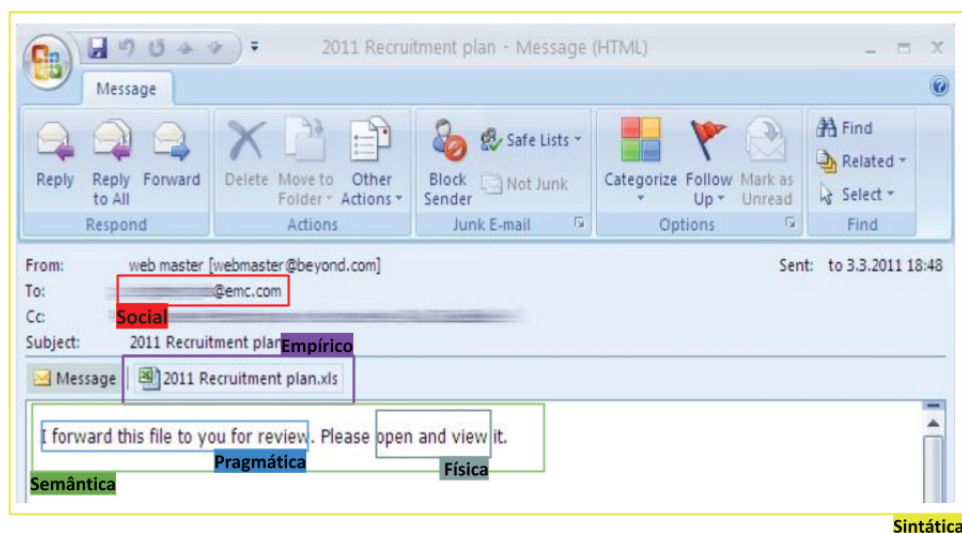


Figura 5.5: Propriedades semióticas no objeto desinformativo.

Foram identificados 46 aspectos sociotécnicos, 29 do sistema de informação humana e 17 da plataforma tecnológica. Os dados da análise foram registrados em tabelas do Google Sheets<sup>12</sup>. O Quadro A.22 apresenta um resumo dos aspectos identificados.

<sup>12</sup>[https://docs.google.com/spreadsheets/analise\\_spear](https://docs.google.com/spreadsheets/analise_spear)



Tabela 5.3: Resumo dos aspectos sociotécnicos identificados na análise.

| Dimensão Semiótica | Aspecto                                                                                                  | Exemplos                                                                                                                                                                                                                                                                                                           |
|--------------------|----------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Social             | Atividades profissionais;<br>Impacto social;<br>Normas de segurança;                                     | A desinformação explora normas do campo de informação da empresa;<br>Causa impactos negativos na imagem da empresa;<br>Prejudica o relacionamento com os 'clientes';                                                                                                                                               |
| Pragmática         | Exploração de expectativas;<br>Exploração de intenções;                                                  | A desinformação tem o objetivo de influenciar os alvos a abrir o arquivo com o <i>malware</i> anexado no e-mail;<br>Empregados tem funções e responsabilidades dentro da organização;<br>Exploração da curiosidade do alvo;                                                                                        |
| Semântica          | Percepções da audiência;<br>Vieses cognitivos;<br>Inexperiência/Ignorância;<br>Linguagem representativa; | O pedido de ajuda para revisar o documento configura o viés cognitivo Quid-pro-quo - a tendência do ser humano em querer ajudar;<br>O aspecto persuasivo da temática, com título e nomes representativos de tarefas da organização;<br>Usuário é convencido a abrir e-mail marcado como spam;                      |
| Sintática          | Linguagem;<br>Diagramação;<br>Perfis de usuários;                                                        | Informações dos perfis dos colaboradores em redes sociais, para identificar vulnerabilidades passíveis de exploração, como seu cargo, idade, escolaridade;<br>E-mail na caixa de spam;<br>E-mail explora a similaridade com tarefas diárias;                                                                       |
| Empírica           | Tempo de leitura;<br>Interação da Comunicação;<br>Codificação;                                           | Quantidade de mensagens com o mesmo remetente (4 e-mails);<br>Arquivo contém um <i>malware</i> que explora uma falha <i>Zero-day</i> ;<br>Arquivo codificado do Excel com um tamanho adequado para transmissão e armazenagem;<br>Desinformação requer a abertura de um outro arquivo, fora da aplicação de e-mail; |
| Física             | Localização;<br>Infraestrutura de rede;<br>Normas de segurança;                                          | Atacantes precisam que o arquivo seja aberto dentro do domínio da empresa;<br>A Internet favorece o anonimato dos criadores;                                                                                                                                                                                       |

### 5.2.1.3 Considerações sobre a análise

A análise revelou aspectos sociotécnicos que embasaram discussões sobre fatores de influência relevantes para o caso. Por exemplo, a análise das partes interessadas evidenciou a complexidade e diversidade dos atores envolvidos em um caso de **desinformação deliberada**, possibilitando uma visão sistêmica dos aspectos que impactam nas suas atividades no contexto do caso. Por exemplo, a criação da desinformação envolveu atividades de planejamento, coleta de dados do público-alvo e o desenvolvimento de um *malware*, indicando a especialização de papéis para diferentes partes interessadas. Outro exemplo de aspecto relevante em relação às partes interessadas foi a existência de participantes indiretos afetados no caso, como os clientes da RSA que tiveram seus dados vazados. Isso ampliou o entendimento das consequências de

uma desinformação deliberada para além do caso propriamente dito, como os possíveis impactos na reputação da RSA e sua relação com clientes.

Em relação a identificação de aspectos sociotécnicos, a análise com a Cebola Organizacional revelou indícios de inter-relações, como a popularidade e situação econômica da empresa (motivação dos atacantes), a disponibilidade de informações relevantes sobre a organização e as partes interessadas (com perfis em redes sociais que informavam o cargo e a especialização dos funcionários selecionados), e o anonimato proporcionado pela natureza livre da Internet, sendo características que facilitaram a elaboração do ataque usando a informação como arma. Nesse sentido, a análise auxiliou a identificar fatores que representam vulnerabilidades ou um risco para a segurança da organização.

A análise do complexo de elementos do canal de comunicação revelou indícios sobre a organização do caso. Por exemplo, a difusão foi facilitada pelas vulnerabilidades dos mecanismos técnicos e tecnológicos de segurança da empresa, como o *malware* que explorou uma falha desconhecida para o anti-vírus, e o mecanismo anti-spam que permitiu a entrega do e-mail. As normativas e treinamentos de segurança insuficientes dos funcionários, e os avisos de spam pouco significativos para os alvos levaram a abertura do e-mail na caixa de spam, não sendo suficientes estes "dispositivo de alerta" para significar uma ameaça ou risco. A desinformação foi mitigada tardiamente após o sucesso do ataque, impactando no relacionamento da empresa com seus clientes e na sua imagem pública. Dessa forma, a análise com a Escada Semiótica permitiu identificar aspectos sociotécnicos relevantes e a entender fragilidades na Plataforma Tecnologia e Sistema de Informação Humano da organização.

O caso revelou novos aspectos sociotécnicos a serem considerados em análises futuras. Por exemplo, a resposta do sistema para o usuário ao abrir o arquivo desinformava o resultado da ação (o arquivo fechava ao abrir), o que implica na necessidade do usuário de inferir o resultado da interação com o sistema sem nenhum indício de erro. Assim, a experiência do usuário com Tecnologia da Informação é um fator relevante para se considerar no contexto do caso analisado. Assim, selecionar um usuário com um perfil vulnerável, com pouca familiaridade com a tecnologia, por exemplo, indica um ato deliberado para aumentar as chances de sucesso do ataque. A manipulação não se limita ao conteúdo de uma mensagem textual, mas também ao *feedback* de procedimentos comuns que são esperados por usuários, como uma mensagem de erro sobre um arquivo corrompido ou um alerta de suspeita do antivírus/sistema anti-spam. Isso expande a análise e a compreensão do uso da desinformação deliberada para a enganação. Nesse sentido, a análise revelou fragilidades relacionadas à interação entre o usuário e sistema e sua experiência e treinamento.

### 5.2.2 Análise da mensagem - Farsa no WhatsApp

Na análise da Farsa no WhatsApp, o autor avaliou o uso dos artefatos Escada Semiótica, o Diagrama de Partes Interessadas, o Modelo Conceitual do Ciclo de Vida e diferentes técnicas para abordar o contexto: a Análise de Enquadramento e os Princípios de Influência. A análise em profundidade considerou o objeto desinformativo original (a mensagem no WhatsApp), e os artefatos Escada Semiótica e DPI. Ao final dessa análise, foram identificados transformações no objeto desinformativo que levaram a mudança no escopo da análise para abranger outros canais de comunicação. Essa mudança de escopo se refere à análise em largura. A análise sociotécnica foi feita colocando em evidência a perspectiva de cada parte interessada em cada estágio do ciclo de vida. Os resultados da aplicação dos artefatos permitiram compreender aspectos sociotécnicos diretamente relacionados às partes interessadas, revelando indícios de deliberação e intencionalidade.

### 5.2.2.1 Farsa no WhatsApp

Em março de 2020 circulou pelo WhatsApp no Brasil uma mensagem que relacionava o combate do SARS-CoV-2 com o nível de Ph do corpo humano (Uol, 2020b; G1, 2021b). A mensagem apresentava um tratamento onde o indivíduo deveria comer certos alimentos para atingir um nível de Ph que supostamente aniquila a SARS-CoV-2. A mensagem trazia em sua estrutura referência a autoridades científicas, informações sobre alimentos com pH errado, além da orientação de que é possível alterar o Ph do corpo (Johnson, 2020). A Figura A.23 apresenta o objeto desinformativo original, a primeira aparição dessa desinformação no Brasil em março de 2020. Ainda, essa desinformação foi corrigida no final de março por agências brasileiras de checagem de (Lupa e G1) (Uol, 2020b; G1, 2021b), apesar das correções, o objeto desinformativo continuou circulando.

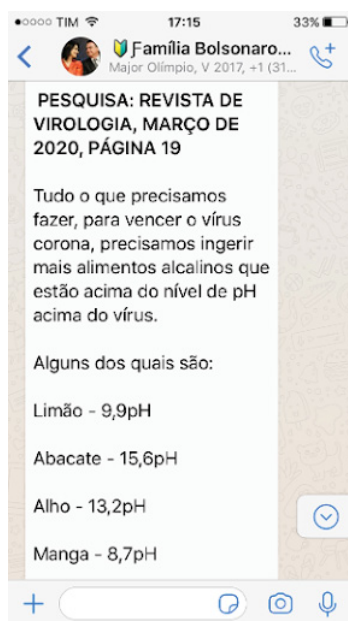


Figura 5.6: Mensagem desinformativa. Fonte: Uol (Uol, 2020b).

A Etapa de Clarificação do Escopo do caso foi realizada com base em reportagens em portais tradicionais de notícia com reputação reconhecida (Uol, 2020a; Alvim, 2020), checadores de fatos (Uol, 2020b; G1, 2021b), relatórios de pesquisas sobre consumo de mídias digitais da população brasileira (Hölig et al., 2020), e portais informativos sobre o tema "saúde" (Johnson, 2020).

A análise revelou 127 aspectos sociotécnicos e suas inter-relações nos estágios do ciclo de vida da desinformação. Alguns aspectos relacionados a Pragmática indicam que a criação da desinformação foi deliberada. Por exemplo, a referência no conteúdo do objeto desinformativo apontando para o artigo científico da Revista da Comunidade de Virologia Americana revela a fonte das informações. Ao coletar informações sobre a referência, descobriu-se que o artigo original é de 1991, não cita o SARS-CoV-2 ou que é possível alterar o pH do corpo humano comendo certos tipos de alimentos, além de ser quase 30 anos antes da pandemia. A introdução de viés e deturpação do conteúdo original representam indícios de deliberação, assim como o uso da referência indicando autoridade certificadora do conteúdo.

Quanto a criação do objeto, a relação do tema com o ambiente da audiência, a popularidade do tema "saúde" em meio a crise sanitária global (Pandemia), e o viés ideológico do grupo onde foi compartilhado - com autoridades popularmente conhecidas por apoiar o tratamento preventivo do SARS-CoV-2 e o anonimato proporcionado pela Internet, são indícios

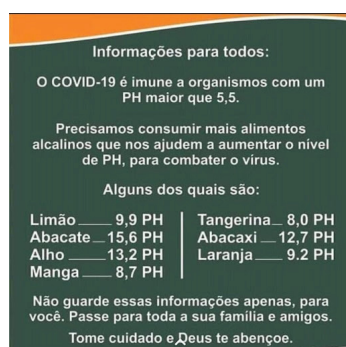
que representam uma propensão para o surgimento da desinformação desta temática neste canal de comunicação. Além disso, estes aspectos auxiliam a entender como o objeto desinformativo é apresentado. Estes aspectos podem representar indícios para uma escolha deliberada de tema em consonância com aspectos da ideologia de grupo onde foi compartilhado, sendo um aspecto sociotécnico inter-relacionado entre partes interessadas e os estágios do ciclo de vida da desinformação e com aspectos sociotécnicos do nível Pragmático ligados a deliberação.

Na difusão, aspectos como a popularidade do aplicativo WhatsApp, suas características que facilitam o compartilhamento de mensagens para usuários em quantidade, e a falta de mecanismos e estratégias efetivas de mitigação dentro do aplicativo se somam aos aspectos do ambiente da audiência, ampliando a capacidade de difusão do objeto desinformativo. O WhatsApp é um canal de comunicação assíncrono bidirecional, possibilitando interações da audiência com o objeto, com o divulgador e entre si. Isso implica na facilidade de discussão e consumo do conteúdo. Além disso, não é possível determinar a origem do objeto desinformativo. Para rastrear a fonte do objeto é necessário acesso a dados restritos que só a parte interessada proprietária "WhatsApp" pode autorizar/prover (caso possuam tais registros), representando uma "fragilidade sociotécnica" que inter-relaciona aspectos dos níveis "Sintático", "Empírico" e "Social" na ES.

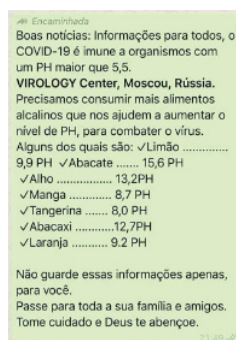
O consumo foi impactado pela incerteza e o medo da audiência relacionados a pandemia (Gedoor, 2021), os hábitos dos brasileiros ao consumir informações sem verificação (Hölig et al., 2020), e aspectos sociotécnicos como os planos de acesso a Internet de dispositivos móveis que limitavam a autonomia na verificação das informações pelo aplicativo (Puska et al., 2020). Outros aspectos que podem impactar o consumo são os vieses cognitivos: viés de confirmação (aceita argumentos alinhados a crenças prévias), efeito manada (favorece opiniões do grupo), e efeito halo (o quanto se gosta de alguém, afeta o julgamento do que ele comunica). Os aspectos sociotécnicos discutidos são exemplos de como o SIH e PT (Escada Semiótica) influenciam os estágios de criação, difusão e consumo da desinformação e como os seus aspectos estão inter-relacionados, representando indícios sobre as motivações e posicionamento das partes interessadas, e como as limitações do canal de comunicação podem impactar no entendimento e expressão dos usuários.

#### 5.2.2.2 Reiteração

Entre abril e maio de 2020 versões diferentes do objeto desinformativo viralizaram via WhatsApp, no Brasil, na Índia, na África do Sul e em Portugal (CE, 2020; Times, 2020; Joel, 2021). O objeto original apareceu pela primeira vez em órgãos de checagem de fatos foi em meados de Março de 2020 na África do Sul (Joel, 2021) e no Brasil (Uol, 2020b). Em Abril de 2020, o objeto desinformativo foi detectado em um novo formato no Facebook, sendo identificada pela agência de checagem de fatos africana Africacheck.org (Joel, 2021). Depois disso, as correções foram difundidas pela aliança internacional de verificadores de fatos do Coronavírus, o CoronavirusFacts (Pointer, 2021; Johnson, 2020). A Figura A.27 mostra as variações da desinformação.



(a) Portugal.



(b) Brasil.

[4/1, 10:04 PM] Karen Allen: This is to inform us all that the pH for corona virus varies from 5.5 to 8.5.

All we need to do, to beat corona virus, we need to take more of an alkaline foods that are above the above pH level of the Virus.

Some of which are:

**Lemon - 9.9pH**  
**Lime - 8.2pH**  
**Avocado - 15.6pH**  
**Garlic - 13.2pH**  
**Mango - 8.7pH**  
**Tangerine - 8.5pH**  
**Pineapple - 12.7pH**  
**Dandelion - 22.7pH**  
**Orange - 9.2pH**

(c) Índia

Figura 5.7: [

Variantes do objeto desinformativo. Fontes: Times of India (Times, 2020), Lupa (Afonso, 2020) e Polígrafo (Sapo, 2021).

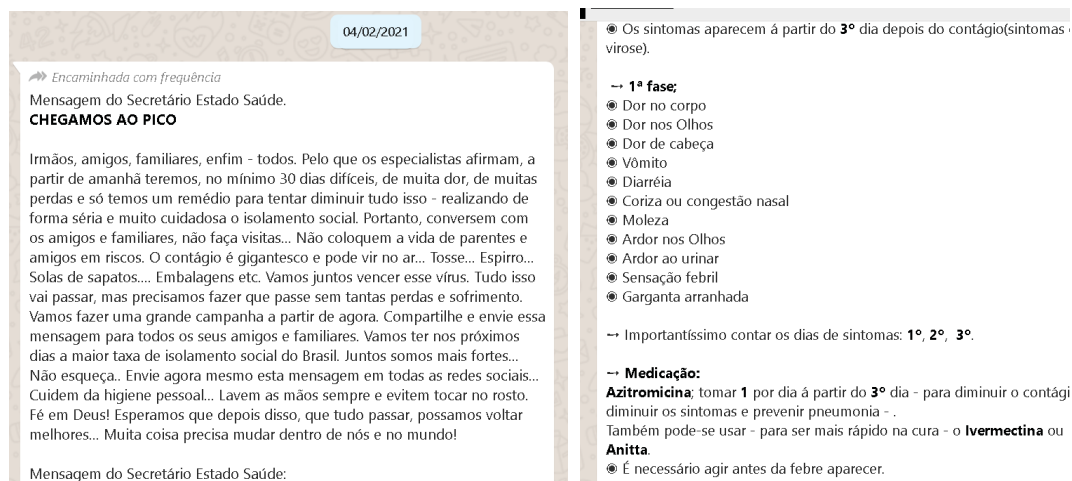
A difusão do objeto desinformativo original para plataformas digitais diferentes pode necessitar de alterações de forma e diagramação, forçando a transformação do conteúdo. A análise do nível "Sintático" revelou aspectos das Plataformas de Comunicação que restringem as propriedades de mensagens, como limites para o número de caracteres e estilo de formatação do texto. Por exemplo, a Figura A.27(a), onde a transformação do objeto desinformativo de texto puro para uma imagem representa a facilidade de difusão entre plataformas digitais sem perder a diagramação do conteúdo. Dessa forma, a estrutura do objeto e os elementos de ênfase que facilitam o consumo se mantêm.

Ainda, em fevereiro de 2021, um ano após a desinformação ter sido identificada e as correções divulgadas por checadores de fatos, foi registrado o reaparecimento no Brasil, em meio a outras informações sobre o tema "tratamento alternativo". A Figura 5.8 apresenta o conteúdo encontrado no novo objeto desinformativo. O objeto continha referências a um documento sobre recomendações para tratamento do SARS-CoV-2 divulgadas pelo Ministério da Saúde (MS) (Almirante e Marés, 2020).

Mesmo após a difusão de correções pelos checadores de fatos, diversos objetos desinformativos relacionados ao tema de tratamento alternativo para SARS-CoV-2 surgiram. Isso representa um indício da insuficiência das intervenções. Os mecanismos propostos pelo WhatsApp para limitar o encaminhamento não apresentam resultados significativos para lidar com os efeitos da desinformação no longo prazo (Melo et al., 2019). Os checadores de fatos frequentemente são descredibilizados por propagarem correções de informações contrárias à ideologia dos grupos, o que prejudica a penetração das correções em câmaras de eco ideológicas (Tang et al., 2017), além das correções serem divulgadas em canais de comunicação que os usuários não acessam regularmente (Fernandez e Alani, 2018). Ainda, a legislação regulatória, sendo um mecanismo de prevenção que busca dissuadir práticas como a criação de desinformação deliberada, sozinha não tem impacto significativo. Os indícios da natureza anônima da Internet, a facilidade de criar uma conta de e-mail, de obter número de celular, e criar uma conta no WhatsApp, dificultam a rastreabilidade da desinformação até sua real origem. A dificuldade de rastrear a desinformação e a dificuldade de fiscalização no WhatsApp permitem que comportamentos anti-sociais, como a criação de desinformação deliberada, ocorra sem punição. Nesse sentido, observa-se a inter-relação de aspectos sociotécnicos que impactam na efetividade das estratégias de mitigação em relação a o estágio de Reiteração.

Percebemos que o caso de desinformação analisado e o objeto desinformativo original fazem parte de um caso maior. Dessa forma, optou-se por ampliar a coleta de informações e objetos desinformativos relacionados ao tema para realizar um estudo amplo do caso "tratamentos





(a) Cabeçalho do objeto desinformativo derivado.

(b) Conteúdo desinformativo no objeto.

(c) Conteúdo do objeto desinformativo original.

Figura 5.8: Transformações do objeto desinformativo. Fonte: *Screenshots*.

alternativos para Covid-19 (SARS-CoV-2)" e verificar a capacidade de transferência<sup>13</sup> da abordagem para analisar diferentes meios de comunicação e tipos de desinformação.

### 5.2.2.3 Implicações: ampliação do escopo de análise do Caso

Com base nos resultados da análise do objeto desinformativo original e seus derivados, ampliou-se a caracterização do caso para englobar o documento oficial (nota informativa Nº 9/2020-SE/GAB/SE/MS) divulgado por autoridade certificadora (Ministério da Saúde) em maio de 2020 e estruturar o estudo a partir dele. Foram coletados objetos desinformativos em diferentes canais de comunicação (Internet, TV, Outdoors, etc.) que tratavam do tema "tratamento alternativo para Covid-19 (SARS-CoV-2)". As análises foram superficiais<sup>14</sup>, buscando investigar o caso em largura e carecem de profundidade para cada objeto desinformativo. No entanto, os resultados preliminares do estudo revelaram aspectos sociotécnicos inter-relacionados entre objetos desinformativos, configurando uma campanha desinformativa - objetos desinformativos com um mesmo tema, objetos com a mesma diagramação e conteúdo idêntico, com ciclos de vida que ocorrem em um mesmo período de tempo, que se difundem por grupos de indivíduos relacionados. As análises em profundidade para cada objeto serão feitas em trabalhos futuros para revelar outros aspectos sociotécnicos inter-relacionados e ampliar o entendimento do Caso.

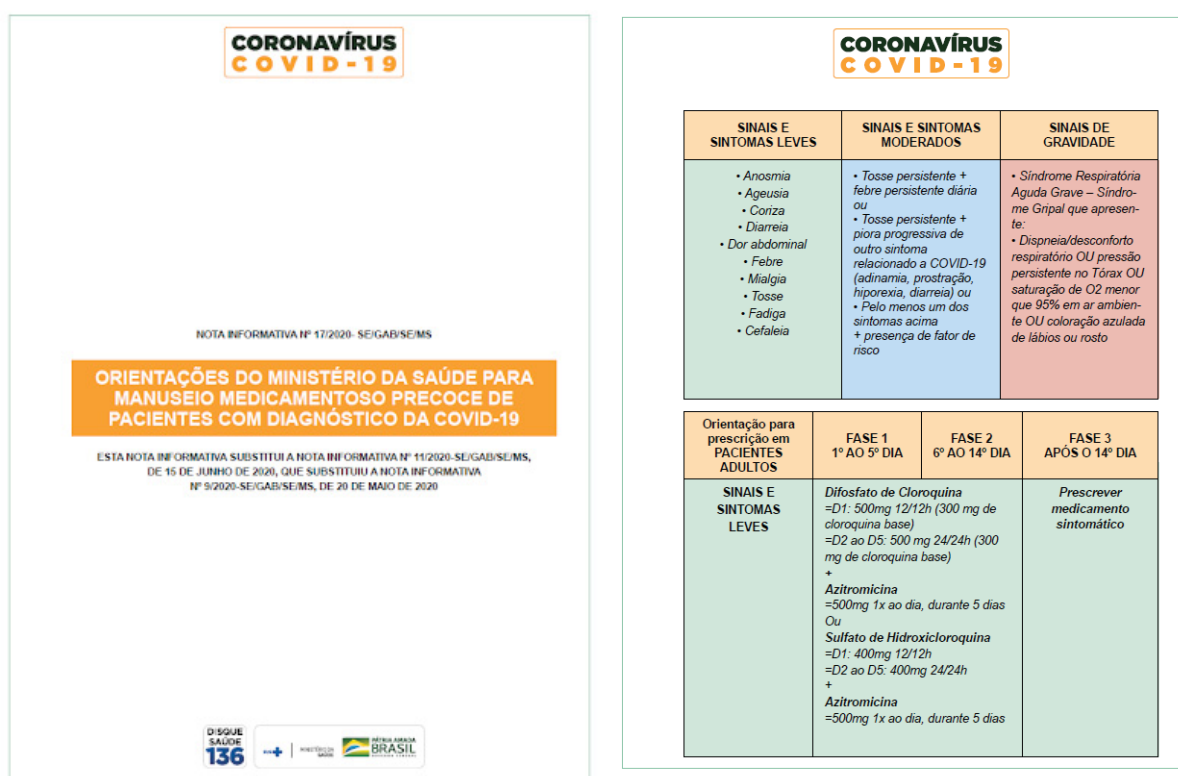
<sup>13</sup>Em relação a capacidade de resolução de problemas(Oulasvirta e Hornbæk, 2016).

<sup>14</sup>O autor não gastou muito tempo na Etapa de Clarificação do Domínio do Caso, coletando poucas informações sobre cada objeto desinformativo descoberto, empregando somente a Escada Semiótica e o DPI para analisar as principais partes interessadas em cada estágio do ciclo, e considerando os Princípios de Influência como material de apoio.



#### 5.2.2.4 Nota informativa Nº 9/2020-SE/GAB/SE/MS

Em 20 de maio de 2020 o ministério da saúde publicou a nota informativa Nº 9/2020-SE/GAB/SE/MS com orientações para manuseio medicamentoso preventivo de pacientes com diagnóstico da COVID-19 (da Saúde, 2021). O material, entre outras indicações, previa a prescrição de medicamentos sem eficácia ou com eficácia não comprovada para tratamento preventivo (Brito, 2021). Essa nota foi substituída em setembro de 2020 pela versão Nº 17/2020-SE/GAB/SE/MS ainda com as indicações de tratamento preventivo. Em janeiro de 2021, o Conselho Nacional de Medicina (CFM) solicitou ao Ministério da Saúde a retirada de todo e qualquer material que mencione o tratamento preventivo da COVID-19 quase um ano depois da divulgação do documento original (dos Anjos, 2021). As Figuras 5.9(a) e 5.9(b) apresentam trechos do objeto desinformativo original.



(a).Capa do documento.

(b) Trecho do documento.

Figura 5.9: Partes da nota informativa Nº 9/2020-SE/GAB/SE/MS.

Alguns aspectos sociotécnicos identificados na análise deste caso revelam a complexidade das atividades desenvolvidas pelas partes interessadas. Por exemplo, a atividade organizada de criar o material, com a parte interessada 'Ministério da Saúde' como a raiz da organização. Nesse contexto, é comum que a equipe delegue funções e determine os procedimentos para a criação do material e determine o conteúdo. Neste caso, existe a configuração de uma atividade organizada relacionada ao estágio de Criação do Ciclo de Vida da Desinformação, que pode incluir diversas outras partes interessadas com funções diferentes. O documento cita quem elabora o texto, quem verifica o conteúdo, os responsáveis pelo objeto digital, e as autoridades que assinam o documento. Nesse sentido, as assinaturas representam autoridades certificadoras para a organização "Sociedade Brasileira", sendo um aspecto sociotécnico ligado ao Princípio de Influência de Autoridade que impacta a forma como a audiência percebe o objeto digital.

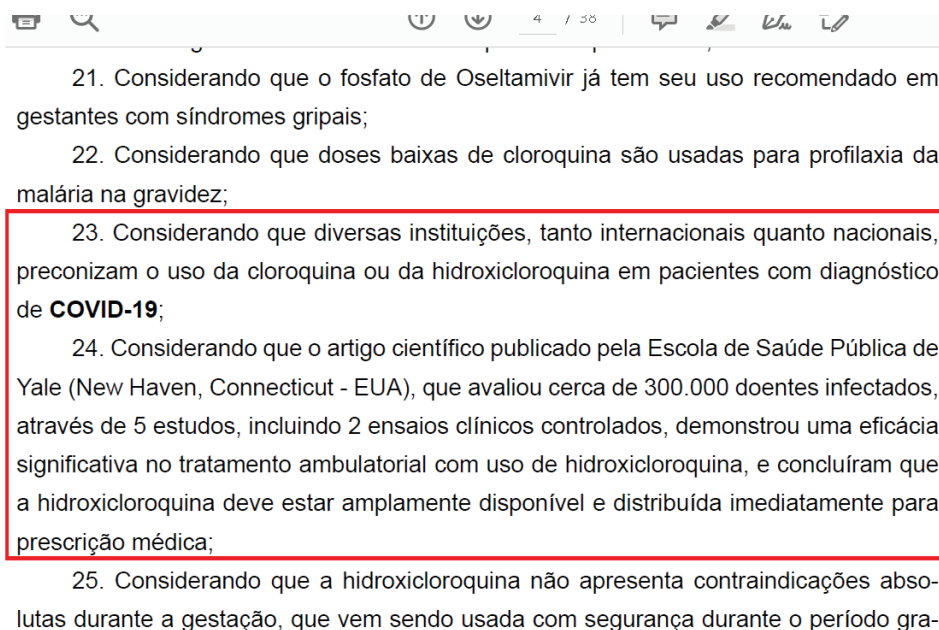
- 
21. Considerando que o fosfato de Oseltamivir já tem seu uso recomendado em gestantes com síndromes gripais;
22. Considerando que doses baixas de cloroquina são usadas para profilaxia da malária na gravidez;
23. Considerando que diversas instituições, tanto internacionais quanto nacionais, preconizam o uso da cloroquina ou da hidroxicloroquina em pacientes com diagnóstico de **COVID-19**;
24. Considerando que o artigo científico publicado pela Escola de Saúde Pública de Yale (New Haven, Connecticut - EUA), que avaliou cerca de 300.000 doentes infectados, através de 5 estudos, incluindo 2 ensaios clínicos controlados, demonstrou uma eficácia significativa no tratamento ambulatorial com uso de hidroxicloroquina, e concluíram que a hidroxicloroquina deve estar amplamente disponível e distribuída imediatamente para prescrição médica;
25. Considerando que a hidroxicloroquina não apresenta contraindicações absolutas durante a gestação, que vem sendo usada com segurança durante o período gra-

Figura 5.10: Exemplos de falácias lógicas: apelo à autoridade e à ignorância. Fonte: Material Oficial (Pág. 4).

O conteúdo deste objeto desinformativo reúne estratégias de manejo e normativas que referenciam **autoridades** para justificar as decisões de escolha e elaboração do conteúdo, como referências à **revistas científicas** e recomendações de **organizações sociais** como o Conselho Federal de Medicina. Neste sentido, alguns aspectos sociotécnicos apontam para um viés ideológico na busca por informações confiáveis que estejam de acordo com a visão ideológica pró-tratamento preventivo. Este tipo de viés é chamado de “viés do gatekeeper/arbitrariedade” (Saez-Trumper et al., 2013), e se caracteriza pela seleção arbitrária de informações para construção da mensagem a ser divulgada, tendendo a um pensamento ideológico específico. A Figura 5.10 exhibe um trecho do material com as falácias lógicas de apelo a ignorância (buscando argumentos sobre fatos, ou possíveis fatos, dos quais a audiência desconhece ou não tem acesso), apelo à Autoridade, Origem, Declive Escorregadio, Meio Termo. Em destaque na Figura 5.10 observa-se a citação de um artigo da Escola de Saúde Pública de Yale como sendo conclusivo e de resultados significativos para amparar a indicação dos passos para manuseio medicamentoso. No entanto, a referência era um artigo de opinião não revisado por pares, embasado por uma literatura sem qualidade (Lupa, 2020). Além disso, os termos “eficácia significativa” são relativos e levam a ambiguidade de interpretação.

As referências e recomendações de autoridades certificadoras são elementos que remetem a credibilidade para o conteúdo, sendo aspectos de influência para a percepção de confiabilidade das informações no objeto. Em destaque na Figura 5.11 estão elementos que configuram o Apelo à Autoridade (falácia lógica) - ao atribuir ao CFM uma proposta (porquê a origem da informação é/parece confiável, então a informação é válida) e Ambiguidade ao afirmar que o CFM “propôs a consideração”(termos que tem várias interpretações) levando a uma conclusão de Falsa Causalidade e Raciocínio Circular (falácias lógicas) para normatizar a implementação e aplicação do tratamento preventivo. Estes dois exemplos constituem aspectos sociotécnicos inter-relacionados dos níveis semióticos “Semântica”, “Pragmática” e “Social” na ES.

A divulgação pelo Ministério da Saúde representa um aspecto do nível semiótico “Mundo Social” que impacta a difusão do objeto desinformativo, a transformações e a reiteração.

30. Considerando que o Conselho Federal de Medicina recentemente propôs a consideração da prescrição de cloroquina e hidroxicloroquina pelos médicos, em condições excepcionais, mediante o livre consentimento esclarecido do paciente, para o tratamento da **COVID-19** (PROCESSO CONSULTA CFM nº 8/2020 – PARECER CFM Nº 4/2020);

31. O Ministério da Saúde, com o objetivo de ampliar o acesso dos pacientes a tratamento medicamentoso no âmbito do SUS, publica as seguintes orientações para tratamento medicamentoso precoce de pacientes com diagnóstico de **COVID-19**.

## 2. ORIENTAÇÕES DO MINISTÉRIO DA SAÚDE PARA MANUSEIO MEDICAMENTOSO PRECOCE DE PACIENTES ADULTOS COM DIAGNÓSTICO DA COVID-19, CONFORME A CLASSIFICAÇÃO DOS SINAIS OU SINTOMAS

Figura 5.11: Recomendações do CFM - Fonte: Material Oficial (Pág. 5)

O surgimento de desinformações sobre o tema "tratamento preventivo" pode ser ligado com objetos desinformativos observados em outros canais de comunicação oficiais. Por exemplo, encontramos aplicativos oficiais do Ministério da Saúde, desenvolvidos em setembro de 2020 embasados no material. Esse período é posterior à análise do objeto desinformativo original (03/2020) e à divulgação do material oficial e os desmascaramentos feitos pela comunidade (05/2020). Os aplicativos TrateCov (G1, 2021a) e Coronavírus - SUS (Olhar Digital, 2021) continham informações falsas sobre tratamento preventivo e o uso de máscaras. As Figuras 5.12(a) e 5.12(b) apresentam as telas com a desinformação nos aplicativos. Assim como as orientações oficiais, os aplicativos tiveram ampla divulgação pelo site do Ministério da Saúde, pela página oficial no twitter (Galo, 2021), no canal de televisão TV Brasil (Abade, 2021) e em lives de Ministros da Saúde (Pancher, 2021), o que amplia o potencial de difusão, e configura uma campanha desinformativa. Isso reflete aspectos do nível semiótico "Mundo Físico", como o custo de se implementar essa campanha.

(a) Tela de diagnóstico (Trate COV).

(b) Tela de dicas (Coronavírus SUS).

Figura 5.12: Desinformação em APPs de Saúde.

A Figura 5.13 mostra discussões sobre tratamento preventivo em sites jornalísticos (Bresilienne, 2021). Podemos notar as consequências de uma desinformação quando propagandeada por autoridade com jurisdição na Sociedade, e o reflexo disso em outras esferas da organização Estado, como o Kit Covid e grupos no WhatsApp de apoio ao tratamento preventivo. A referência ao material em objetos desinformativos encontrados em diversos canais de comunicação, bem como as propagandas e discussões, constituem indícios do impacto da (autoridade da fonte) na Criação, Difusão, Consumo, Reiteração e Mitigação da desinformação. Nesse sentido, a autoridade que endossa o conteúdo do material facilita o surgimento de desinformação nessa temática (Princípio de Afinidade, Autoridade e Unidade) e fomenta discussões em torno de posições argumentativas enviesadas (falácias lógicas), observadas em diversos canais de comunicação.

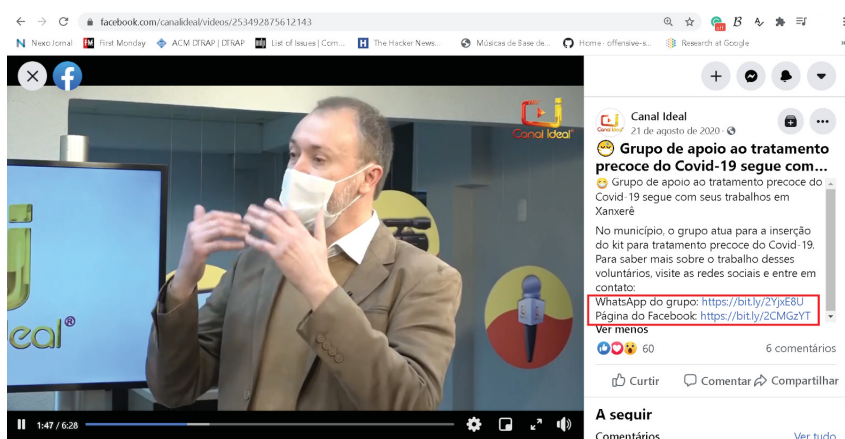


Figura 5.13: Área de discussões sobre notícias. Fonte: Gazeta do Povo (do Povo, 2021).

Redes sociais como o Facebook e Twitter colaboraram para a disseminação de desinformação nesse tema. A divulgação nas redes sociais vai desde anúncios com propagandas de drogas ineficazes (Fávero et al., 2021) à postagem de usuários influentes, como políticos e médicos (Fatos, 2020). A Figura 5.14 apresenta postagens realizadas entre março e agosto de 2020 que indicam a ideologia de autoridades (pessoas públicas e influenciadores), e corroboram as discussões feitas para o objeto original (Figura A.23) em relação a ideologia do grupo que foi registrado. A aceitação de correções que apresentam visões contrárias à ideologia de um grupo tem dificuldade de penetrar e ser consumida (Fernandez e Alani, 2018). Isso fica claro quando a autoridade que desinforma tem alcance e influência para além do grupo com ideologia alinhada a visão da desinformação.

Estes canais são assíncronos e bidirecionais, permitindo interações da audiência com os objetos desinformativos de várias formas diferentes, como reações, discussões e compartilhamentos. Meio bidirecionais permitem que a audiência discuta o conteúdo do objeto entre si e com quem o criou (ou compartilhou), implicando no impacto de aspectos ligados à defesa da ideologia de grupo, a indivíduos indecisos/inseguros/ignorantes podem seguir a ideologia do grupo, a informação corrobora crenças e valores. Além disso, nas redes sociais, cada usuário é um "gatekeeper" e representa um canal de divulgação de informações. Os mecanismos de interação que permitem compartilhar o objeto desinformativo, faz com que um mesmo objeto desinformativo seja observado em diferentes canais de comunicação (diferentes contas de usuários). Isso somado às câmaras de eco ideológicas (Tang et al., 2017) e filtro-bolha (DiFranzo e Gloria-Garcia, 2017), induz a percepção de que o objeto/conteúdo é mais popular do que realmente é, constituindo aspectos relacionados ao Princípio de Afinidade.





Figura 5.14: Postagens em redes sociais (Março/Agosto de 2020). Fonte: AosFatos (Fatos, 2020).

Em outra forma de difusão, entidades privadas (Associação de Médicos pela Vida) investiram recursos em anúncios pagos em jornais (fevereiro de 2021) para incentivar o uso de medicamentos ineficazes e com efeitos colaterais (Uol, 2021b). "Autoridades da saúde" indicam credibilidade, impactando no consumo e difusão da desinformação. A Figura 5.15 apresenta um exemplo de organização promovendo a desinformação.

### Exemplo de distribuição do tratamento em Belém



Figura 5.15: Matéria da BBC sobre comunidade de profissionais da saúde pró-tratamento preventivo (Setembro de 2020). Fonte: BBC (Alvim, 2020).

Em outro meio de divulgação, encontramos uma página web dedicada ao tratamento preventivo da covid. O site <https://covidtemtratamentosim.com.br/><sup>15</sup> estava carregado de desinformação. De narrativas sentimentalistas com conclusões enviesadas usando referências científicas inconclusivas, as desinformações induzem o usuário a consumir a desinformação. O registro do site no Registro.Br foi realizado no dia 12/06/2020 por uma empresa de marketing digital, o que coloca o surgimento em meio a outros objetos desinformativos da temática. A Figura 5.16 apresenta o portal desinformativo na mesma temática.

A desinformação pode se espalhar para além dos limites digitais. A Figura 5.17 apresenta propagandas em outdoors, representando formas de difusão além das mídias digitais, ampliando

<sup>15</sup>Pode ser acessado em <https://web.archive.org/web/20200703002615/https://covidtemtratamentosim.com.br/>



Figura 5.16: Site sobre a temática de tratamento preventivo.

o alcance da desinformação endossada pelo objeto desinformativo digital para o público que não tem acesso aos canais de comunicação.



Figura 5.17: Outdoors em Curitiba. Fonte: Gazeta do Povo (Do Povo, 2021).

#### 5.2.2.5 Considerações sobre o estudo

A análise da farsa no WhatsApp revelou o impacto de aspectos sociotécnicos inter-relacionados na criação, difusão e consumo. Por exemplo, os indícios de deliberação na criação, atribuindo a uma fonte de informações que ela não divulgou (Revista de Virologia Americana). Na difusão, propriedades do canal de comunicação implicam em maior alcance de difusão, como a popularidade do WhatsApp entre os brasileiros. No consumo, falácias lógicas como o apelo à autoridade, e vieses cognitivos como o efeito manada, indicam o potencial de consumo do conteúdo. Além disso, a relação entre o viés ideológico do grupo, restrições que limitam a capacidade de verificação de informações no WhatsApp, e a popularidade do tema da mensagem, apontam para aspectos sociotécnicos que impactam o ciclo de vida, ampliando a capacidade de difusão e consumo.

Foram identificados aspectos no nível Social da Escada Semiótica que influenciaram, tanto a criação do objeto desinformativo quanto suas transformações e outros objetos na temática. O fato da pandemia e da polarização ideológica derivada de conflitos políticos permitiram que a temática fosse tomada como arma para explorar a audiência através de diferentes meios de comunicação (Uol, 2020a). Nesse sentido, a abordagem auxiliou a análise em largura do caso, revelando aspectos sociotécnicos inter-relacionados nos diferentes objetos desinformativos e seus ciclos de vida.



A análise do objeto desinformativo "documento oficial" revelou indícios que apontam para decisões conscientes que implicam em deliberação na escolha do conteúdo para construção do objeto. Por exemplo, as referências a estudos inconclusivos e a canais de comunicação científica sem reputação ou processo de revisão por pares. É importante ressaltar que isso não quer dizer que a intenção da fonte das informações foi de enganar a audiência, mas que a fonte é responsável pelo conteúdo que construiu e divulgou, e pelas possíveis consequências. Entre outros indícios encontrados no decorrer das análises estão o recorrente Apelo à Autoridades duvidosas para o conteúdo a divulgação por autoridades sociais, que se relacionam com a mesma perspectiva ideológica de tratamento preventivo, podendo indicar a intenção de influenciar a audiência a adotar esta ideologia.

Partindo de um escopo de caso de desinformação ampliado, os resultados da análise em largura revelaram diferentes canais de comunicação usados para difundir desinformações na mesma temática. O impacto de aspectos do nível semiótico "Mundo Social", como autoridade da fonte, polarização ideológica, e ignorância sobre o vírus, foram observados em canais de comunicação que variam desde discussões informais em fóruns e redes sociais, até aplicativos e portais dedicados à desinformação. Os aspectos sociotécnicos inter-relacionados indicam o impacto da desinformação oficial divulgada por uma autoridade com jurisdição sobre toda organização social, em fomentar o surgimento de diversos objetos desinformativos e enviesar discussões.

A diversidade de canais de comunicação representa a complexidade de mitigar a desinformação em canais com diferentes propriedades e controlados por diferentes Partes Interessadas. Por exemplo, fóruns utilizam moderadores, WhatsApp utiliza limitação de encaminhamento, caixas de comentários de portais de notícias parecem não ter nenhum mecanismo ativo de mitigação, Twitter utiliza censura, Facebook utiliza rótulos em postagens. Essa diversidade apresenta desafios técnicos e formais para estratégias de mitigação, como a fiscalização e rastreio de desinformação em comunicações criptografadas ou na elaboração de uma legislação apropriada para os requisitos de cada plataforma. Ademais, os impactos de aspectos dos canais de comunicação e a falta de estratégias e mecanismos efetivos de mitigação. Por exemplo, as características, meio assíncrono, agrupamentos, a falta de integração com ferramentas externas e falta de mecanismos de verificação internos do WhatsApp impactam na organização dos usuários como agrupamentos que compartilham vieses ideológicos, autoridades no grupo, além de possibilitar maior controle do fluxo de informações por partes interessadas com influência nas discussões da audiência. Ainda, o uso de canais de comunicação como Outdoors, propagandas pagas em jornais impressos e na televisão, representam formas de superar os limites de difusão das redes sociais digitais, levando a desinformação para além do alcance da população com acesso a Internet e *smartphones*.

A análise em largura possibilitou observar as atitudes e reações de diferentes partes interessadas, tanto as alinhadas ideologicamente com a posição epistêmica de consumir e apoiar o tratamento preventivo quanto partes interessadas que o combatiam. Ao final do estudo, o objeto desinformativo do tipo farsa parece fazer parte de uma teoria da conspiração sobre os tratamentos da Covid-19, que reuniram argumentos em diferentes linhas, desde o uso de medicamentos sem eficácia comprovada, a campanha anti-vacina, o potencial de indução à auto-medicação pelas campanhas desinformativas, gastos privados e públicos com propagandas, e omissão de agências estatais de fiscalização (Federal, 2023).

Como resultado de síntese dos dados e aspectos sociotécnicos identificados, observamos o ciclo de vida da desinformação no caso analisado com três principais fatores relevantes: a) a extensão da autoridade (jurisdição) das partes interessadas da camada Contribuição que criaram a desinformação (com influência na Sociedade Brasileira), b) a difusão por partes interessadas

do grupo Fonte e Mercado em canais de comunicação sem intervenções previstas/apropriadas, c) consumo pelo grupo Fonte e Mercado impulsionou a reiteração de desinformação em grupos de ideologias distintas. Os principais aspectos sociotécnicos que impactaram no ciclo: 1) autoridade da fonte, 2) ignorância sobre ciência, 3) ignorância sobre o vírus, 4) insuficiência/inexistência de intervenções, 5) qualidade das fontes de informação, 6) autoridades influenciadoras e 7) agrupamentos.

### 5.3 OFICINAS DE ANÁLISE SOCIOTÉCNICA DE DESINFORMAÇÃO

Para avaliar a abordagem com o público-alvo, foram realizadas três oficinas de análise sociotécnica de desinformação. O objetivo foi avaliar a utilidade e a facilidade de uso da Dismiss percebidas pelos participantes. As oficinas foram atendidas por um grupo de pesquisadores em computação, representando os projetistas interessados no fenômeno da desinformação. O Quadro 5.4 apresenta um resumo das oficinas.

Tabela 5.4: Resumo das Oficinas

| ID | Data e Duração           | Objetivo                                                                                                                                                                                                                                                                                                                     | Participantes             | Principais apontamentos                                                                                                                                                                                                                                                                                                                                                                   |
|----|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | 07/2022<br>2hs e 30 min. | Avaliar utilidade e facilidade de uso percebidas da abordagem: os artefatos da SO, análise de enquadramento, os materiais de apoio (Princípios de Influência, Quadro de Aspectos da Comunicação, Quadro de Falácias Lógicas e Vieses Cognitivos, Formulários, Questões-guia para uso do DPI e da ES), e o Manual da Dismiss. | 5 Participantes e o autor | <ul style="list-style-type: none"> <li>- Desenvolver materiais de apoio e simplificar os passos para aplicação da Dismiss.</li> <li>- Reduzir o número de formulários e os campos a serem preenchidos.</li> <li>- O Manual precisa ser simplificado e objetivo.</li> <li>- Reestruturar as etapas da Dismiss para dar mais foco e objetividade na análise do contexto do caso.</li> </ul> |
| 2  | 12/2022<br>2hs           | Avaliar utilidade e facilidade de uso percebidas da abordagem: os artefatos da SO, análise de enquadramento, os materiais de apoio (Princípios de Influência, Quadro de Aspectos da Comunicação, Quadro de Falácias Lógicas e Vieses Cognitivos, Formulários, Questões-guia para uso do DPI e da ES), e o Manual da Dismiss. | 5 Participantes e o autor | <ul style="list-style-type: none"> <li>- Reduzir o número de formulários e os campos a serem preenchidos.</li> <li>- Reduzir os passos da exploração do contexto.</li> <li>- Simplificar as explicações.</li> <li>- Relacionar o Ciclo de Vida com a Estrutura do Caso.</li> <li>- Esclarecer as estratégias de aplicação da Dismiss.</li> </ul>                                          |
| 3  | 03/2023<br>2hs           | Avaliar utilidade e facilidade de uso percebidas da abordagem: os artefatos da SO, análise de enquadramento, os materiais de apoio (Princípios de Influência, Aspectos da Comunicação, Questões para uso do DPI e da ES), e o Manual.                                                                                        | 3 Participantes e o autor | <ul style="list-style-type: none"> <li>- Mais tutoriais em vídeo e exemplos de casos analisados.</li> <li>- Mais exemplo de aspectos sociotécnicos relevantes.</li> <li>- Simplificar o manual de uso da Dismiss.</li> </ul>                                                                                                                                                              |

As três oficinas foram elaboradas para avaliar as duas estratégias de aplicação da Dismiss, empregando a análise sociotécnica de objetos desinformativos e do ciclo de vida da desinformação digital no mesmo encontro. As oficinas foram divididas em etapas, como na estrutura de aplicação da Dismiss. Os protocolos para realização dos dois primeiros estudos podem ser encontrados no Apêndice A.8 e do terceiro no Apêndice A.9. Os dados dos estudos foram coletados por formulários usados pelos participantes nas análises (representando os artefatos da SO), questionários respondidos ao final do estudo e por entrevistas feitas no modelo de grupo focal com os participantes após o preenchimento do questionário.

Os resultados das oficinas apoiaram as escolhas do autor na elaboração da Dismiss, em especial no ajuste da quantidade de componentes, passos e estratégias que compõem a abordagem. Na primeira oficina o autor apresentou a primeira versão da Dismiss, que consistia de uma gama de conceitos, artefatos, materiais de apoio à análise, tutoriais e manuais de aplicação da Dismiss que infundiram complexidade e esforço para o entendimento e uso da abordagem. Embora os participantes tenham indicado a utilidade percebida da Dismiss com um potencial

de apoiar o entendimento da natureza sociotécnica do fenômeno da desinformação, a facilidade de uso da abordagem foi o aspecto mais desafiador na sua elaboração. Com os resultados de cada oficina, o autor implementou melhorias para simplificar o entendimento e uso da Dismiss, como a re-elaboração dos materiais de apoio à análise, a re-estruturação dos passos da Dismiss, e a simplificação dos artefatos e manuais. A última oficina foi realizada com a versão mais atual da abordagem, e os resultados mostram a evolução da facilidade de uso percebida da abordagem pelos participantes. Nas próximas seções são apresentadas as oficinas em detalhes, onde são explicados componentes, artefatos e estratégias que fizeram parte de versões anteriores da Dismiss, mas que foram retirados ou re-elaborados para aprimorar a sua facilidade de uso<sup>16</sup>.

### 5.3.1 Oficina I

A primeira oficina ocorreu em 07/2022 e foi atendida por 5 participantes, todos pesquisadores em IHC na UFPR. A oficina teve duração de 02 horas e 30 minutos, sendo realizada em ambiente virtual de modo síncrono. Nessa oficina, os participantes deveriam, individualmente, usar o manual, um passo-a-passo de aplicação, e os formulários de registro de dados da Dismiss para analisar o caso de desinformação apresentado na Figura 5.18. O autor serviu como intermediador, podendo ser consultado pelos participantes para sanar qualquer dúvida que tivessem sobre todos os aspectos da aplicação da abordagem, do preenchimento dos artefatos a explicações sobre os conceitos do Modelo do Ciclo de Vida da Desinformação.

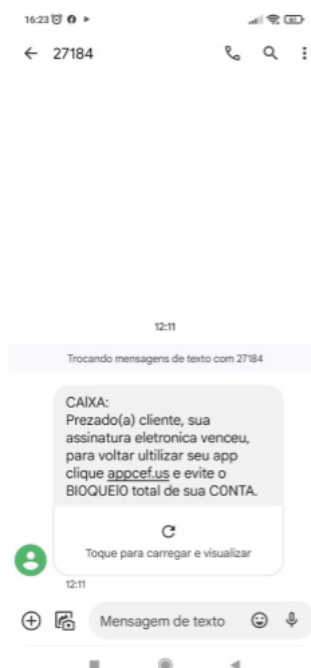


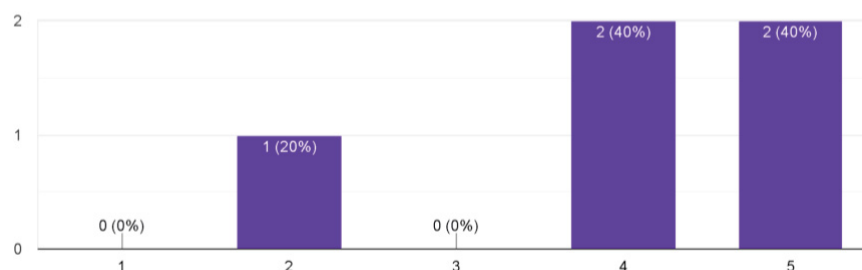
Figura 5.18: Desinformação para análise - Oficina I.

Essa oficina foi planejada para avaliar utilidade e facilidade de uso percebidas da abordagem: os artefatos da SO, análise de enquadramento, os materiais de apoio para identificação de aspectos sociotécnicos (Princípios de Influência, Quadro de Aspectos da Comunicação, Quadro de Falácias Lógicas e Vieses Cognitivos, Questões-guia para uso do DPI e da ES), e o

<sup>16</sup> Alguns componentes e conceitos foram retirados para serem refinados em trabalhos futuros, com o intuito de complementar a Dismiss (e.g Cebola Semiótica e o conceito de campos de informação, discutidos no Apêndice 6.1.

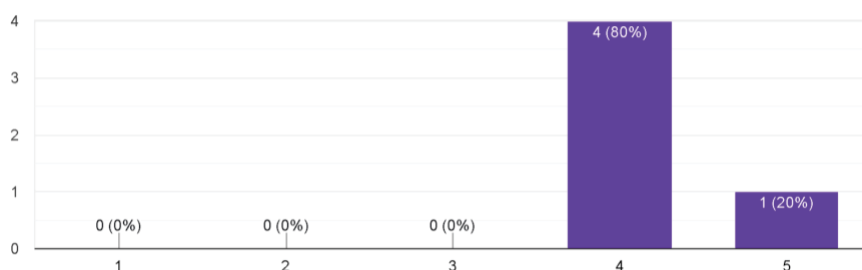
Manual da Dismiss. Durante a aplicação da Dismiss, os participantes preencheram formulários que foram usados como método de coleta de dados, e um questionário no encerramento do estudo. O questionário consistia de questões com a escala Likert (onde "1" significava "discordo completamente" e 5 "Concordo completamente") e perguntas abertas. Os resultados, apresentados na Figura 5.19, mostram a utilidade percebida da Dismiss para apoiar o entendimento da natureza sociotécnica complexa do fenômeno da desinformação, indicada pelas respostas no questionário <sup>17</sup>.

A abordagem me apoiou a identificar aspectos sociotécnicos do evento desinformativo analisado.  
5 respostas



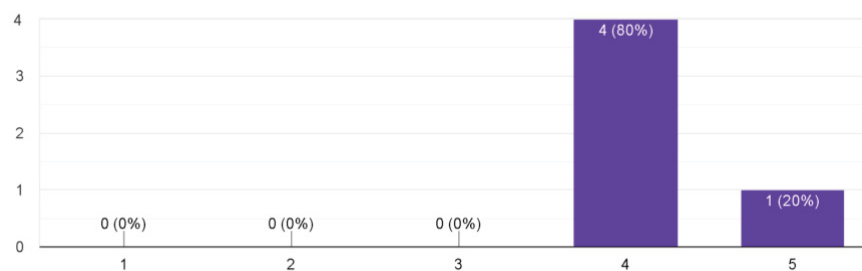
(a) Capacidade de auxiliar a identificar aspectos sociotécnicos relevantes.

A abordagem me fez pensar em problemas que eu dificilmente teria identificado sem usá-la.  
5 respostas



(b) Capacidade de auxiliar a identificar problemas.

A DISMISS me ajudou a compreender a natureza sociotécnica da desinformação no caso analisado.  
5 respostas



(c) Capacidade de apoiar o entendimento da natureza sociotécnica do fenômeno.

Figura 5.19: Sobre a utilidade percebida da Dismiss.

Isso pode ser notado nas respostas dadas pelos participantes no grupo focal. Por exemplo, P3 comentou sobre sua percepção relacionada a utilidade da abordagem:

*"Reconheço a utilidade e potencial profundidade da ferramenta, mas para analistas não familiares com essa ferramenta, ela talvez possa ser de difícil*

<sup>17</sup>As questões não eram de resposta obrigatória, por isso a quantidade de respostas em cada questão pode variar.

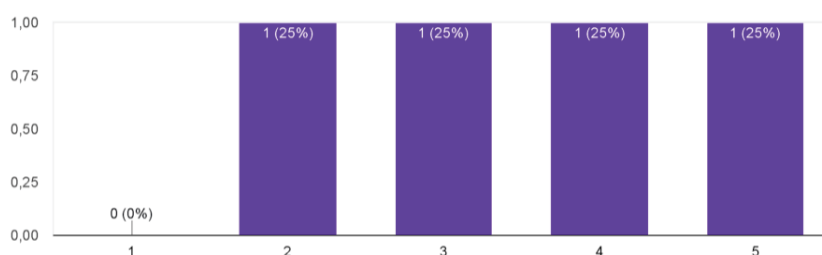
*aplicação. Talvez um manual estendido ou um número maior de exemplos ajudasse, mas ainda seriam necessários exemplos e explicações capazes de dar conta de vários tipos de casos com vários tipos de partes interessadas."*

P5 também se referindo a utilidade percebida da abordagem:

*"Não houve dificuldades, apenas não imaginei que o método fosse tão amplo e abrangente."*

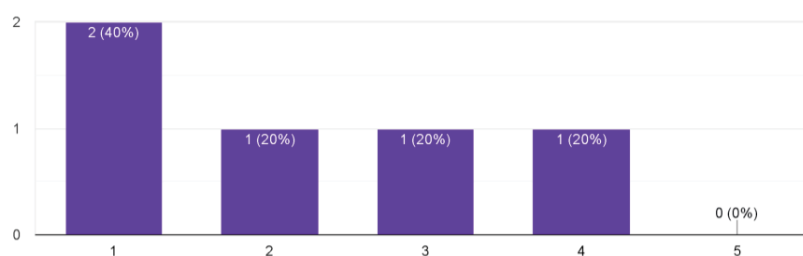
A primeira versão da Dismiss era composta pelo DPI, a ES, uma versão da Escada Semiótica estendida chamada Arcabouço Semiótico Estendido (ASE - usada para caracterização dos estágios do ciclo de vida), e materiais de apoio que eram utilizados na Etapa II de análise sociotécnica do caso de desinformação. Cada artefato da SO foi mapeado para um formulário para preenchimento e registro dos dados. A Figura 5.20 mostra as respostas dos participantes da oficina em relação a utilização destes formulários.

O Formulário de Análise de Enquadramentos auxilia na caracterização dos aspectos sociotécnicos que influenciam o enquadramento de observação usado pelas partes interessadas.  
4 respostas



(a) Problemas com o formulário de caracterização do objeto desinformativo.

Minha interação com o ASE foi clara e compreensível, não exigindo muito esforço mental.  
5 respostas



(b) Facilidade de uso do ASE.

Figura 5.20: Sobre a facilidade de uso dos formulários de registro de dados da Dismiss.

Isso pode ser visto na Figura 5.21 que mostra o preenchimento do artefato ASE pelo participante P4. O formulário cruzava três dimensões de análise sociotecnia (Partes interessadas, Estágio do ciclo de vida e Escada Semiótica) que se mostraram desafiadoras. Além disso, a quantidade de formulários para registro do caso e dos dados de preenchimento dos artefatos tornaram a aplicação da Dismiss confusa. Os participantes tinham que recorrer ao manual e ao tutorial para obterem exemplos de como preencher este formulário, o que causou muito esforço e confusão.

|                     | Criação                                                                                                                     | Difusão                                                                              | Consumo                                                                                                                                                                                          |
|---------------------|-----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                     | golpistas virtuais                                                                                                          | plataforma de SMS                                                                    | usuário de celular                                                                                                                                                                               |
| <b>Mundo Social</b> | - Golpistas tentam se passar por um banco buscando usar o elemento de Autoridade para tirar vantagem de pessoas vulneráveis | - Mensagens são disparadas em massa para contatos que não tem relação com o golpista | - Um usuário que possui conta no Banco em questão acaba sendo mais vulnerável ao golpe<br>- Usuários que não possuem familiaridade com tecnologias também são mais suscetíveis a caírem no golpe |
| <b>Pragmática</b>   | - A mensagem é redigida em forma de ordem, com tom de urgência.                                                             |                                                                                      | - O usuário pode ser enganado devido ao princípio da Escassez, pois a mensagem afirma que o usuário corre o risco de ter sua conta bloqueada se não clicar com urgência no link solicitado       |
| <b>Semântico</b>    | - A linguagem utilizada simula a linguagem padrão formal utilizada em mensagens autênticas de Bancos                        |                                                                                      |                                                                                                                                                                                                  |
| <b>Sintático</b>    | - As palavras 'Bloqueio' e 'Conta' estão em caixa alta com o propósito de chamar a atenção do leitor.                       | - A plataforma não possui mecanismos de mitigação                                    |                                                                                                                                                                                                  |
| <b>Empírico</b>     |                                                                                                                             | - A plataforma é bidirecional e assíncrona, de interação um p/ muitos                | - A mensagem é enviada individualmente para o alvo                                                                                                                                               |
| <b>Mundo Físico</b> | - Pessoas com celulares                                                                                                     | - Celulares que possuem suporte ao envio e recebimento de SMS                        | - Usuários em diferentes localidades do Brasil                                                                                                                                                   |

Figura 5.21: Formulário que mapeia o ASE - Preenchido por P4.

Isso também pode ser notado nas respostas das perguntas abertas; Por exemplo, quando questionado sobre as dificuldades de aplicação da Dismiss, P4 respondeu:

*"Muitos formulários, muitos nomes diferentes. Exemplo: na planilha em que respondemos, a Aba da planilha tem um nome, e a o formulário mesmo tem outro nome. Exemplo: Formulário de caracterização da Mídia (nome da aba), e "Formulário de caracterização do ciclo de vida no topo do formulário. Não consegui entender pq tantos formulários, pq não seria mais fácil fazer vários deles juntos?"*

Na mesma pergunta, P2 respondeu:

*"Acredito que a ferramenta é muito densa. Os materiais de apoio são bons e úteis, mas também são densos e espalhados. Precisei alternar frequentemente entre quatro documentos para me ajudar a preencher os formulários e responder as perguntas... A alternância entre os materiais me deixou perdido..."*

A Dismiss incorpora conceitos da SO, da Análise de Enquadramento, de redes de computadores e da psicologia da persuasão. Essa diversidade de conceitos tornou a Dismiss complexa e difícil de usar. Embora os participantes tenham reconhecido a utilidade dos artefatos e o potencial de aplicação da abordagem para apoiar o estudo sociotécnico do fenômeno da desinformação, a carga conceitual e a densidade teórica da Dismiss foram dois dos pontos levantados pelos participantes de modo recorrente. A Figura 5.22 mostra indicadores nas respostas do questionário sobre a complexidade conceitual e teórica da abordagem.

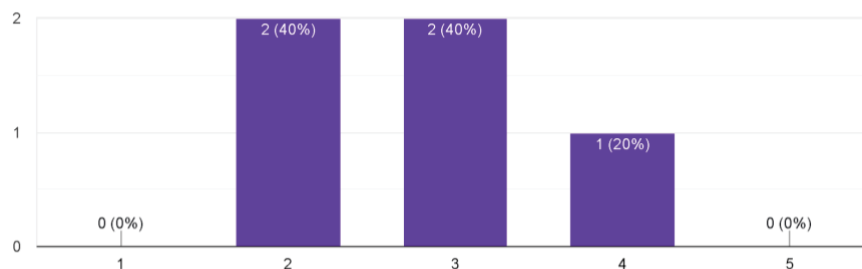
A complexidade conceitual e teórica da primeira versão da Dismiss também pode ser observada nas respostas dos participantes no grupo focal. Quando indagados sobre dificuldades para entender os conceitos da abordagem, P2 respondeu:

*"Não sei se entendi corretamente o conceito de enquadramento e isso junto da minha pouca familiaridade com a escada semiótica acabaram complicando a aplicação do ASE. Conheço os diferentes degraus da escada semiótica, mas por não estar acostumado a aplicá-la na prática, durante a aplicação do ASE tive certa dificuldade em aplicar alguns dos conceitos. Essa dificuldade só aumentou ao tentar cruzar a escada semiótica com o enquadramento de diferentes partes interessadas. Em certos momentos, por exemplo, eu talvez até*



Considero fácil a aprendizagem de uso da DISMISS.

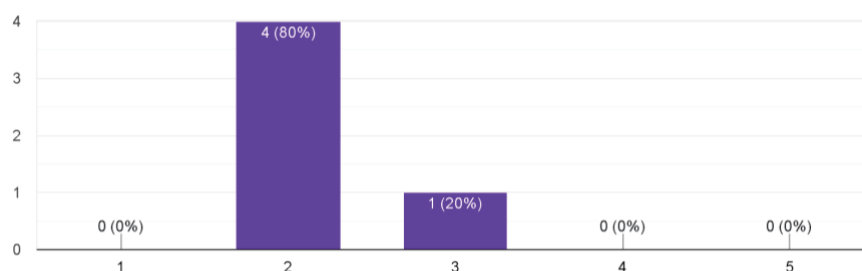
5 respostas



(a) Facilidade de aprendizado da Dismiss.

Considero fácil a aprendizagem de uso do ASE.

5 respostas



(b) Facilidade de aprendizado do ASE.

Figura 5.22: Sobre a complexidade conceitual e teórica da Dismiss.

*conseguisse analisar a mensagem de uma perspectiva pragmática, sintática ou semântica, mas não consegui entender o que eu deveria identificar ao analisar a camada X de um caso desinformativo a partir do enquadramento de parte interessada Y."*

P3 também indica essa complexidade:

"Os conceitos apresentados são bastante complexos, e nem sempre a ligação entre eles está clara (embora esteja descrita, nem sempre é fácil compreendê-la)."

A Figura 5.23 mostra o Formulário de Caracterização do Objeto Desinformativo (que incorpora os conceitos de Enquadramento) preenchido por P2.

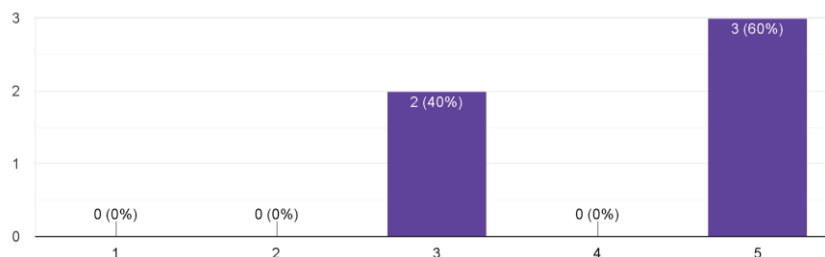
| Formulário de Caracterização do Objeto Desinformativo                    |                                                                                                    |                                                    |
|--------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|----------------------------------------------------|
| Campo                                                                    | Valor                                                                                              | Justificativa                                      |
| Objeto desinformativo:                                                   |                                                                                                    |                                                    |
| Tipo do Objeto:                                                          | Farsa                                                                                              |                                                    |
| Tema do evento:                                                          | Outro                                                                                              |                                                    |
| Tema do objeto:                                                          | bloqueio de conta                                                                                  |                                                    |
| Breve descrição do conteúdo do objeto desinformativo:                    | A mensagem se inicia com a identificação do emissor como a CAIXA                                   |                                                    |
| Breve descrição dos elementos de <b>Ênfase</b> no objeto desinformativo: | frase imperativa "evite o bloqueio total da sua conta"<br>adjetivo "total" intensificando a ameaça |                                                    |
| Breve descrição dos elementos <b>Elementos Visuais</b> :                 | "BIOQUEIO" e "CONTA" em caixa alta                                                                 |                                                    |
| Breve descrição das <b>Ligações com o Ambiente</b> :                     | identificação com o nome da CAIXA                                                                  |                                                    |
| Apresenta alguma falácia lógica?                                         | Apelo à Emoção                                                                                     | apelo ao medo para motivar o alvo a clicar no link |
| Conclusões:                                                              | O alvo deve clicar no link fornecido ou perderá sua conta                                          |                                                    |

Figura 5.23: Formulário que mapeia o ASE - Preenchido por P2.

Em relação ao Modelo Conceitual do Ciclo de Vida da Desinformação Digital, os participantes parecem ter compreendido facilmente seus conceitos e a forma de aplicá-lo no estudo de casos de desinformação. A Figura 5.24 apresenta os indicadores de utilidade e facilidade de uso percebidos do modelo nas respostas dos participantes no questionário.

O Ciclo de Vida da Desinformação auxilia na análise sistêmica da comunicação de desinformação.

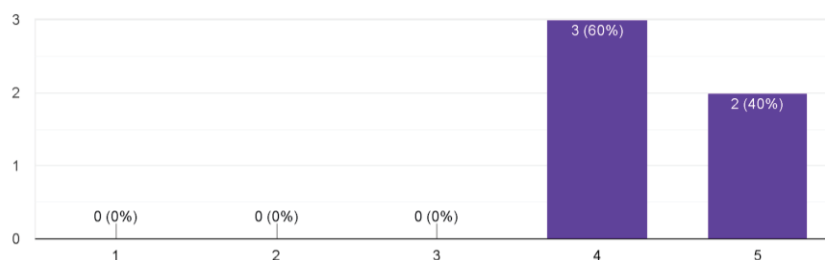
5 respostas



(a) Utilidade percebida do Ciclo.

Minha interação com o Ciclo de Vida foi clara e compreensível, não exigindo muito esforço mental.

5 respostas



(b) Facilidade de uso do Ciclo.

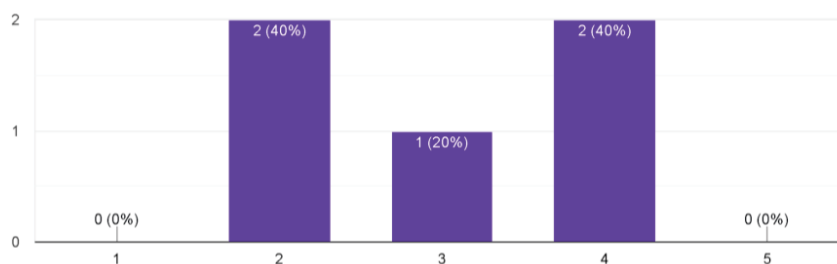
Figura 5.24: Sobre utilidade e facilidade de uso do Modelo Conceitual do Ciclo de Vida da Desinformação.

A Dismiss engloba vários materiais de apoio à análise sociotécnica para serem usados em diferentes etapas da abordagem. São questões-guia e quadros explicativos que funcionam como eixos complementares que dão suporte à análise. Por exemplo, os Princípios de Influência foram proposto para ser utilizado durante a Etapa II da Dismiss, dão um suporte teórico e conceitual para identificação de aspectos sociotécnicos nas dimensões da Escada Semiótica. Eles guiam a identificação de aspectos sociotécnicos que têm potencial de influenciar as percepção de risco e avaliação de veracidade das partes interessadas, e como eles podem aprimorar ou restringir a criação, difusão, consumo e mitigação.

A Figura 5.25 mostra as respostas dos participantes em relação a utilidade percebida destes materiais.

As perguntas de apoio para identificação de aspectos sociotécnicos me ajudaram a utilizar o "Arcabouço Semiótico Estendido".

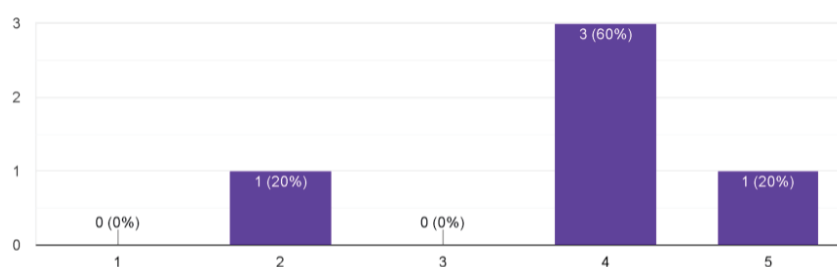
5 respostas



(a) Utilidade das questões-guia para identificação de aspectos sociotécnicos.

As Questões que guiam a Identificação de Partes Interessadas auxiliam a preencher o Diagrama de Identificação de Partes Interessadas.

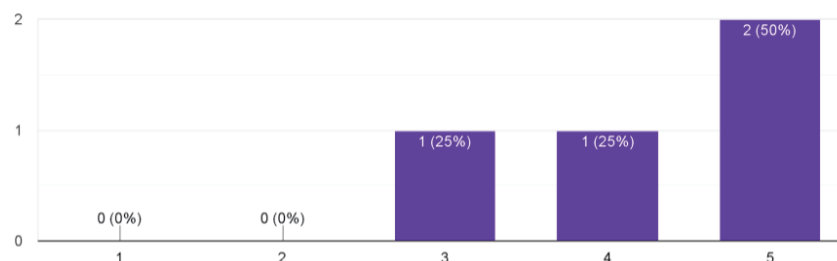
5 respostas



(b) Utilidade percebida das questões-guia para identificação de partes interessadas.

Considero os materiais de apoio (princípios de influência, falácias, lógicas, aspectos da comunicação) úteis para identificação de aspectos ...cos relevantes para o evento analisado (Etapa II).

4 respostas



(c) Utilidade percebida dos materiais de apoio para identificação de aspectos sociotécnicos relevantes.

Figura 5.25: Sobre utilidade dos materiais de apoio.

Embora a percepção de utilidade possa ser considerada positiva, a quantidade de componentes da Dismiss acabou por tornar a utilização destes materiais de apoio confusa e cansativa. Isso pode ser notado nos comentários feitos por P2:

*"Encontrei certa dificuldade em entender a ordem de preenchimento dos formulários. Após a identificação dos stakeholders, por exemplo, não sabia se eu deveria preencher o ASE (que vinha imediatamente depois na planilha), ou preencher o formulário de registro de evento (caso), que requer apenas informações do DPI e me parece mais relacionado à etapa 1 (clarificação)."*

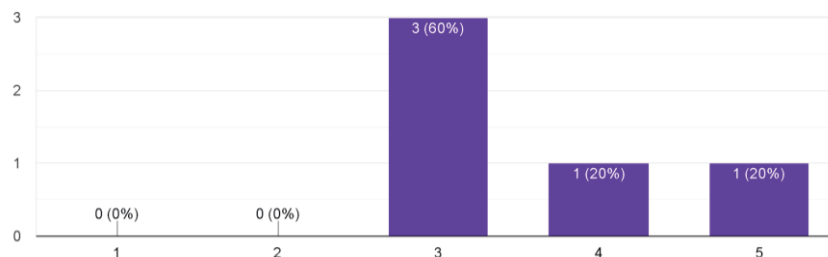
e de P5:

*"O método todo é longo e cansativo, entendo que não tenha como ser de outra forma devido a profundidade com que se busca entender o fenômeno estudado."*

A Dismiss oferece um manual de aplicação que explica os conceitos e artefatos da SO, e da Análise de Enquadramento. A Figura 5.26 apresenta indícios da percepção de utilidade.

O Manual me apoiou a aplicar a DISMISS.

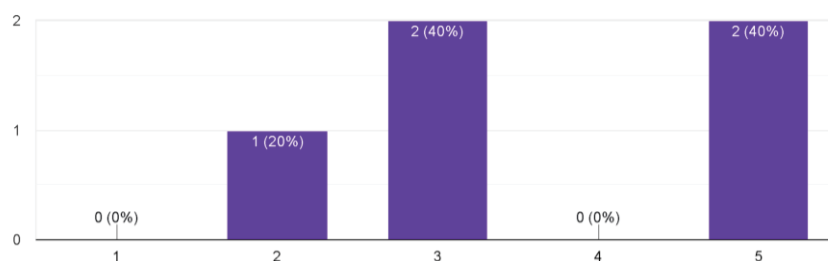
5 respostas



(a) Utilidade do manual em apoiar o uso da Dismiss

As instruções contidas no Manual são suficientes para aplicação da DISMISS.

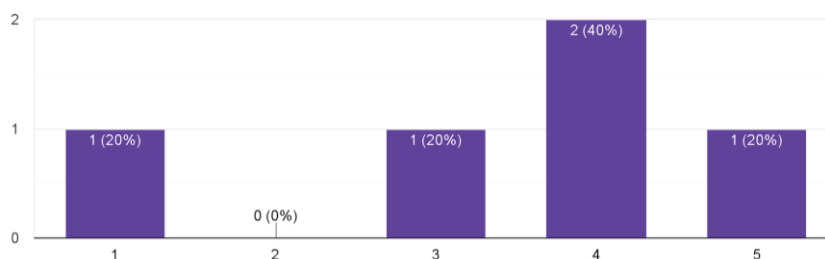
5 respostas



(b) Avaliação da completude do manual.

O exemplo apresentado no manual foi útil na aplicação da DISMISS.

5 respostas



(c) Avaliação da utilidade dos exemplos usados no manual.

Figura 5.26: Sobre utilidade e facilidade de uso do manual de aplicação da Dismiss.

Seguindo na mesma gama de aspectos desafiadores da Dismiss, os participantes comentaram sobre a utilidade do manual.

Por exemplo, P2 indicando a necessidade de mais explicações sobre alguns conceitos e métodos de análise:

*"Não acho que o manual esteja ruim, mas sinto que preciso de mais explicações a cerca da escada semiótica e análise de enquadramento para poder entender como melhor aplicar esses conceitos á análise de um objeto desinformativo. Ainda assim as outras partes do manual e também dos exemplos fornecidos me foram bastante uteis."*

P3, indicando melhorias para o texto do manual:

*"O mesmo já dito. leveza, palavras simples, frases diretas."*

P4, sobre a experiência de usar o manual:

*"É exaustivo utilizar o manual. Talvez mudar a forma de apresentação da informação, como um tutorial por tópicos ou sistema de perguntas-respostas."*

De modo geral, a Dismiss parece útil para apoiar o entendimento da natureza complexa do fenômeno da desinformação. Entretanto, a abordagem é muito "densa", com muitos passos e formulários que confundem a sua aplicação, necessitando de aprimoramentos. Isso pode ser notado nas respostas à pergunta "Na sua opinião, quais os pontos negativos da DISMISS e como podem ser melhorados?"

Por exemplo, P2:

*"Como mencionado, o principal ponto negativo da DISMISS está na dificuldade da aplicação do método. Mais uma vez reconheço a potencial utilidade do método nas mãos de um analista mais familiarizado com seus conceitos. Vejo que parte do apelo do método está justamente em sua capacidade de permitir uma análise em profundidade de um objeto desinformativo. Me pergunto se o método não poderia ser adaptado de forma a se tornar mais modular. Permitindo, por exemplo, que um analista escolha quais artefatos utilizar e o quão profundamente analisar um objeto."*

Com base nas análises dos dados preenchidos nos formulários e no uso artefatos, foram desenvolvidos materiais de apoio e passos simplificados para aplicação da Dismiss. Os dados dessas fontes mostraram dificuldades para o uso do DPI e da ES, refletidos especialmente no preenchimento dos formulários de registro de dados. As respostas dos questionários indicaram o esforço dos participantes para compreender os conceitos e corroboraram as dificuldades no uso e preenchimento dos artefatos. O grupo focal indicou melhorias para o estágio de identificação de aspectos sociotécnicos, bem como revelou a opinião dos projetistas sobre a utilidade da Dismiss.

### 5.3.2 Oficina II

Após a análise dos dados coletados na primeira oficina, o autor implementou melhorias na abordagem considerando as maiores dificuldades relatadas pelos participantes. O número de formulários e os campos a serem preenchidos foram reduzidos, buscando maior clareza para o seu preenchimento. O manual foi reformulado para ser mais simples e objetivo, trazendo os conceitos-chave para a compreensão e uso da abordagem. As etapas da Dismiss foram revistas e reestruturadas para dar mais foco e objetividade na análise do contexto do caso, integrando os formulários no contexto da análise. Após os refinamentos, foi feita a segunda oficina.

A segunda oficina ocorreu em 12/2022 e foi atendida por 5 participantes, todos estudantes de IHC na UFPR. A oficina teve duração de 02 horas, porém foi realizada em ambiente presencial - um laboratório do Departamento de informática da UFPR. Os participantes deveriam utilizar os computadores e, individualmente, usar um passo-a-passo e os materiais de aplicação da Dismiss para analisar um caso real de desinformação. O manual da Dismiss poderia ser consultado em caso de dúvidas. O autor apresentou uma visão geral da aplicação da Dismiss, e permaneceu como intermediador durante todo o estudo, podendo ser consultado pelos participantes para sanar qualquer dúvida relacionada a qualquer aspecto da abordagem e de sua aplicação.

Os participantes poderiam escolher entre dois casos qual analisar (Figuras 5.27 e 5.28).



Figura 5.27: Isca-clique no Youtube - Caso 1.

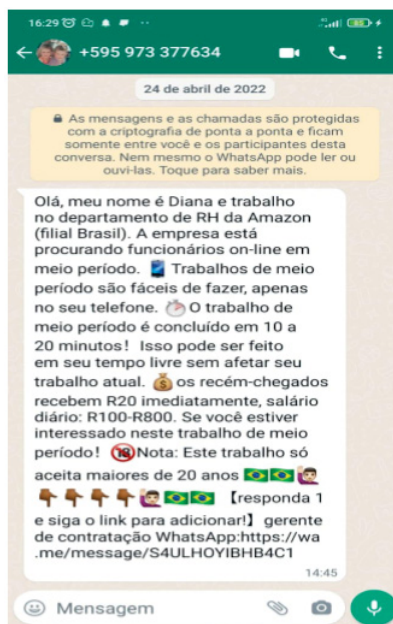


Figura 5.28: Farsa no WhatsApp (Oferta de emprego) - Caso 2.

Essa oficina foi planejada nos mesmos moldes da primeira: para avaliar utilidade e facilidade de uso percebidas da abordagem. O método de coleta de dados foi o mesmo da Oficina I, os participantes preencheram formulários e um questionário no encerramento do estudo. O questionário consistia de questões que usaram a escala Likert (onde "1" significava "discordo completamente" e 5 "Concordo completamente") e perguntas abertas. Ainda, ao final do estudo foi feito um grupo focal para perguntar sobre as dificuldades que os participantes tiveram na oficina e obter sugestões de melhorias.

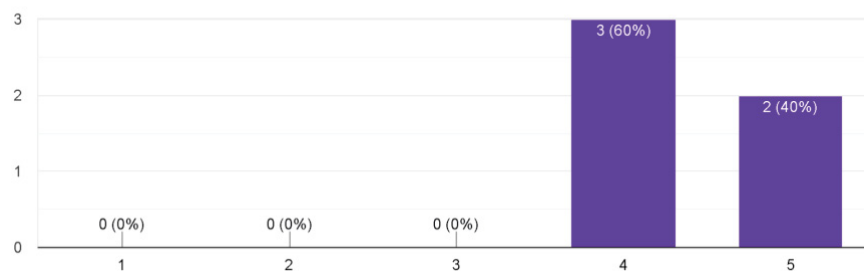
Em relação as respostas no questionário sobre a utilidade percebida da abordagem, a Figura 5.19 mostra as respostas sobre: utilidade percebida da Dismiss para apoiar o entendimento da natureza sociotécnica do fenômeno da desinformação, a capacidade de auxiliar a caracterização sociotécnica do contexto do caso e a identificar problemas, indicadas pelos participantes <sup>18</sup>.

<sup>18</sup>As questões não eram de resposta obrigatória, por isso a quantidade de respostas em cada questão pode variar.



A DISMISS me apoiou a identificar aspectos sociotécnicos relevantes no caso analisado?

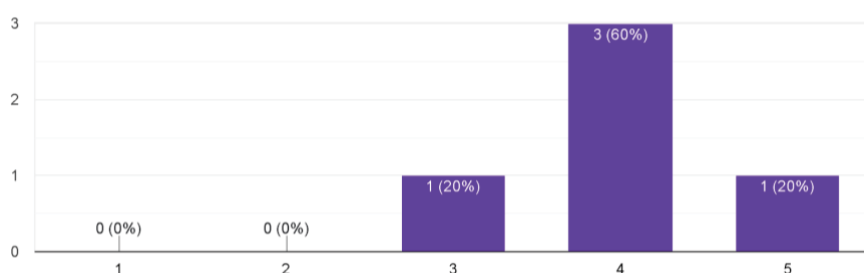
5 respostas



(a) Capacidade de auxiliar a identificar aspectos sociotécnicos relevantes.

A DISMISS me auxiliou a identificar problemas que levaram a ocorrência de desinformação.

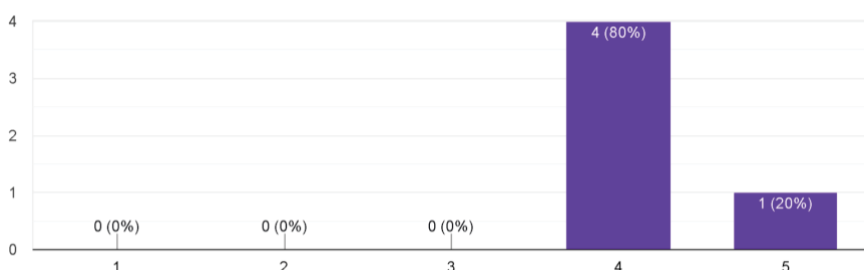
5 respostas



(b) Capacidade de auxiliar a identificar problemas.

A DISMISS me ajudou a compreender a natureza sociotécnica da desinformação no caso analisado.

5 respostas



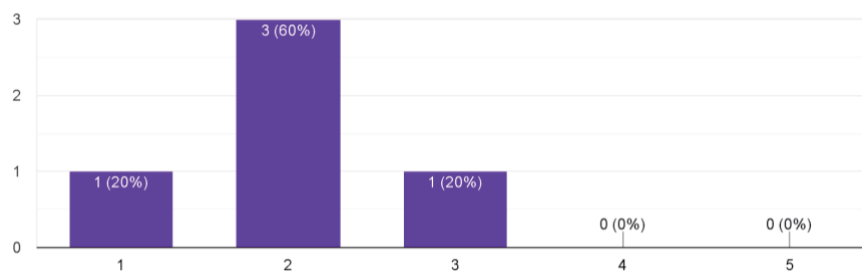
(c) Capacidade de apoiar o entendimento da natureza sociotécnica do fenômeno.

Figura 5.29: Sobre a utilidade percebida da Dismiss.

Embora a utilidade percebida da abordagem seja indicada pelos participantes, e mesmo com as melhorias após a primeira oficina, a Dismiss continuou complexa e cansativa para aplicação. A Figura 5.30 mostra a percepção dos participantes sobre a facilidade de uso geral da abordagem, sobre o aprendizado dos conceitos e artefatos, e sobre o Modelo Conceitual do Ciclo de Vida.

Considero fácil a aprendizagem de uso da DISMISS.

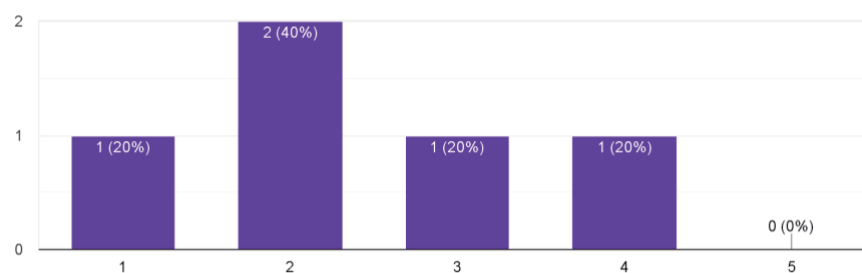
5 respostas



(a) Sobre a facilidade de aprendizado percebida da Dismiss.

Considero que o uso da DISMISS foi claro e compreensível, não exigindo muito esforço mental.

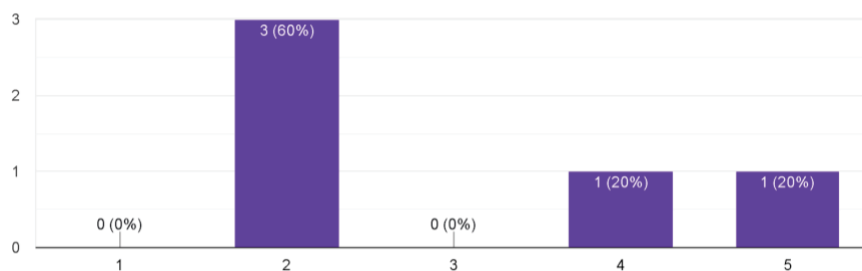
5 respostas



(b) Sobre a facilidade de uso da abordagem.

Minha interação com o Ciclo de Vida foi clara e compreensível, não exigindo muito esforço mental.

5 respostas



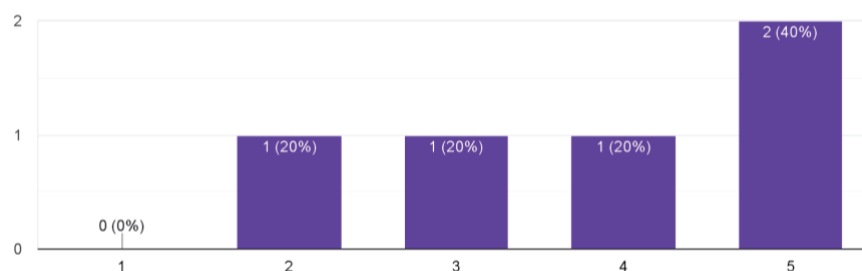
(c) Sobre a facilidade de uso do Modelo Ciclo de Vida da Desinformação.

Figura 5.30: Sobre a facilidade de uso da Dismiss em geral.

Apesar dos indicadores de melhora em relação aos Formulários de Registro do Caso, os demais formulários continuaram complexos e cansativos. A Figura 5.31 apresenta as respostas dos participantes em relação ao preenchimento dos formulários.

Minha interação com o Formulário de Registro de Casos foi clara e compreensível, não exigindo muito esforço mental. .

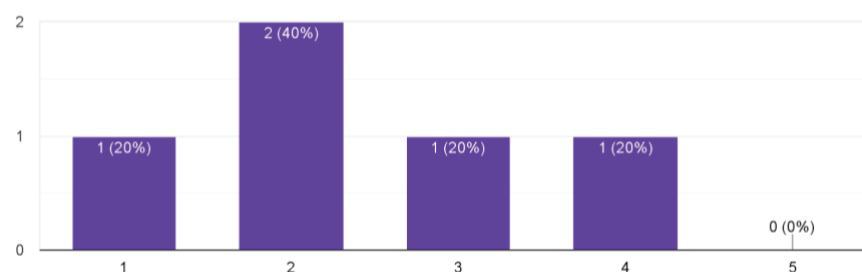
5 respostas



(a) Sobre a facilidade de uso do Formulário de Registro de Casos.

Minha interação com o Formulário de Identificação de Aspectos Sociotécnicos foi clara e compreensível, não exigindo muito esforço mental.

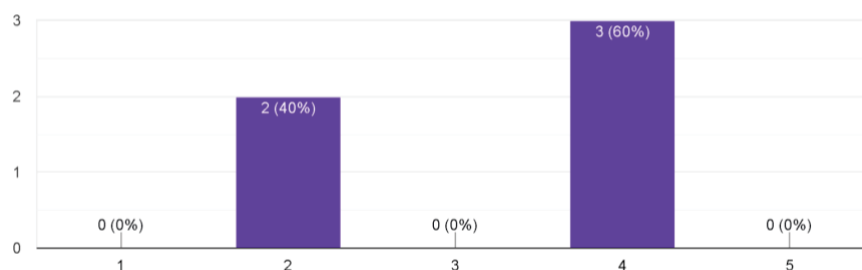
5 respostas



(b) Sobre a facilidade de uso do Formulário de Registro de Aspectos Sociotécnicos.

Minha interação com Formulário de Identificação de Partes Interessadas foi clara e compreensível, não exigindo muito esforço mental.

5 respostas



(c) Sobre a facilidade de uso do Formulário de Registro de Partes Interessadas.

Figura 5.31: Sobre a facilidade de uso dos materiais da Dismiss.

Ao responder as perguntas abertas relacionadas as dificuldades na aplicação, P5 esclareceu as limitações sobre a densidade dos componentes e a estratégia de uso na Dismiss:

*"Acho que ficou muito complexo juntar o ciclo de vida, os seis níveis da escada, e os princípios de influência. Os princípios são muito úteis para identificar problemas. A escada é útil para pensar em diferentes níveis de formalidade do fenômeno. O ciclo de vida é útil para pensar em diferentes problemas que ocorrem em diferentes estágios. Mas para mim, não consegui fazer tudo isso ao mesmo tempo. Tive primeiro que preencher a escada, depois eu pensei nos princípios de influência. Só então eu iria pensar no ciclo de vida, para conseguir pensar no problema em sua criação, depois na difusão, depois no consumo. Não sei se é tão natural (pelo menos pra mim) esse pensamento de*

*conseguir ligar "ciclo de vida com princípios de influência em diferentes níveis de formalidade"...".*

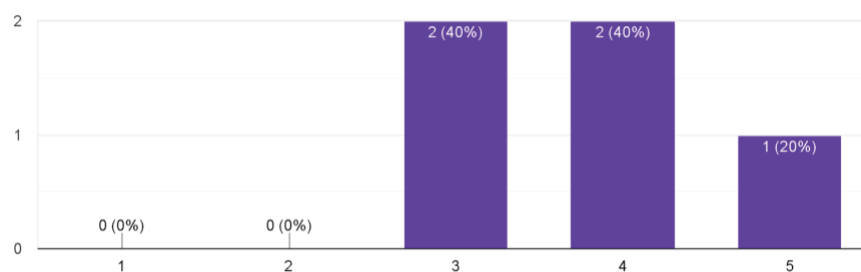
e P1:

*"Está muito densa para ser utilizada. Os formulários estão difíceis, esteticamente, estruturalmente, e em termos de conteúdo. O que eu quero, no final das contas, é um passo-a-passo bem prático para guiar minha análise..."*

Em relação as melhorias no manual, o autor refinou o exemplo de como aplicar a Dismiss, reduziu o tamanho e a complexidade do texto, com menos conceitos e detalhes sobre o fenômeno. A Figura 5.32 mostra os indícios da percepção dos participantes sobre o manual.

O exemplo apresentado no manual foi útil na aplicação da DISMISS.

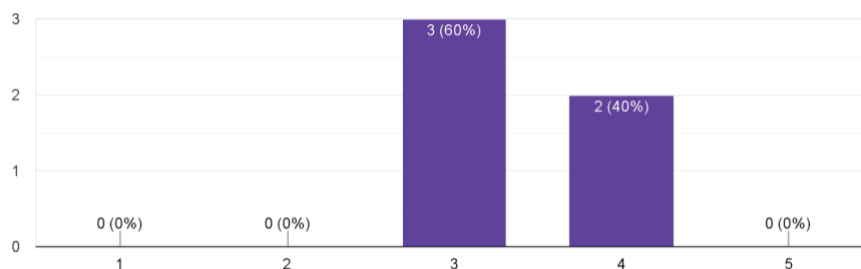
5 respostas



(a) Sobre a utilidade do exemplo de aplicação no manual.

O Manual fornece todas as informações necessárias para Análise Semiótica da Desinformação.

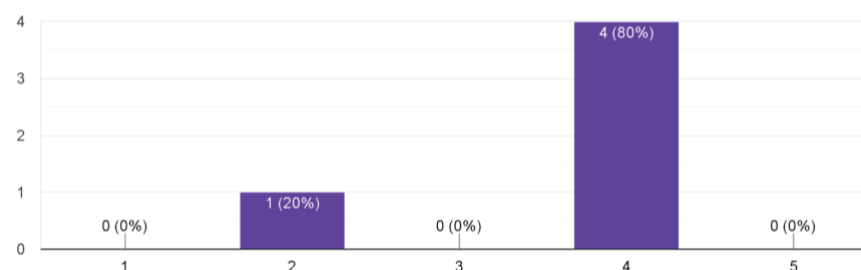
5 respostas



(b) Sobre a completude do manual para realização da análise sociotécnica.

O Manual fornece todas as informações necessárias para entendimento do Ciclo de Vida da Desinformação.

5 respostas



(c) Sobre a completude do manual para entendimento do Ciclo de Vida.

Figura 5.32: Sobre a facilidade de uso dos materiais da Dismiss.

Nas respostas abertas para a pergunta sobre as dificuldades e possíveis melhorias para o manual, P2 respondeu:

*"O manual está bem carregado ainda, longo de acompanhar. Muitos conceitos técnicos."*

P3 respondeu:

*"O problema é que as informações são MAIS DO QUE suficientes, causando uma carga mental alta para o leitor."*

Ao analisar os dados da segunda oficina, os resultados sobre a facilidade de aplicação da Dismiss continuam ligados à densidade de conceitos e artefatos que a Dismiss emprega. Apesar destes conceitos serem parte essencial da abordagem e mapearem a natureza do fenômeno numa tentativa de englobar as suas diversas dimensões de análise, a Dismiss necessita de mais refinamento para simplificar sua aplicação sem perder o seu propósito em relação a abrangência ou limitar a investigação à partes específicas do fenômeno. A complexidade de aplicação parece estar na densidade teórica e conceitual incorporadas em cada etapa de aplicação da abordagem.

### 5.3.3 Oficina III

A análise dos dados coletados na segunda oficina revelou o que pareceu ser um foco dos problemas relacionados ao fluxo de aplicação da abordagem: a estrutura de aplicação continuava densa, com muitos artefatos, formulários e passos inter-relacionados que deixam a aplicação cansativa. Além disso, foi detectado uma fragilidade em relação ao real entendimento da estrutura do ciclo de vida e sua função na Dismiss. Focando em simplificar os passos da Dismiss, o autor implementou melhorias na abordagem, como a redução para quatro formulários de aplicação, sendo dois deles relacionados aos artefatos da SO: DPI e ES. Os passos foram reduzidos e simplificados em relação as explicações e possibilidades de exploração do contexto em cada caso. O ciclo de vida foi apresentado como sendo a estrutura de um caso para ser seguido na exploração do contexto. Além disso, a Dismiss foi dividida em duas estratégias de aplicação: a estratégia de análise de objetos desinformativos e a estratégia de análise do ciclo de vida que complementa a análise considerando a abrangência de um caso de desinformação.

O formato da oficina foi mudado, permitindo aos participantes focarem no preenchimento dos artefatos, ao invés de tentar fazer sentido de todos os conceitos e artefatos da Dismiss apenas com o manual de instrução. Com o autor sendo um mediador, conduzindo os participantes pelos passos da Dismiss, os participantes puderam interagir melhor com os artefatos e dar mais atenção para análise. O questionário também sofreu alterações, sendo reduzido à 10 perguntas, dando mais ênfase para as respostas abertas que permitem maior elaboração dos participantes. Após esses refinamentos (da Dismiss e da oficina), foi feita a terceira oficina.

A terceira oficina ocorreu em 03/2023 e foi atendida por 3 participantes, todos estudantes de IHC na UFPR. A oficina teve duração de 2 horas, sendo realizada de modo síncrono em ambiente virtual. Essa oficina foi proposta com uma estrutura diferente das outras, onde o autor serviu como intermediador da aplicação da abordagem e não só do estudo. O autor apresentou uma visão geral da aplicação da Dismiss<sup>19</sup>, e guiou os participantes como em um passo-a-passo. Nessa oficina, os participantes deveriam analisar uma desinformação em vídeo do tipo Teoria da Conspiração (Figura 5.33)<sup>20</sup>.

<sup>19</sup>O autor gravou um vídeo explicativo (20 minutos) para introduzir a análise sociotécnica de desinformação. Acessível em [https://youtu.be/CqUIK3u\\_BEg](https://youtu.be/CqUIK3u_BEg).

<sup>20</sup>Vídeo acessível em <https://twitter.com/choquei/status/1635793607373254656>.

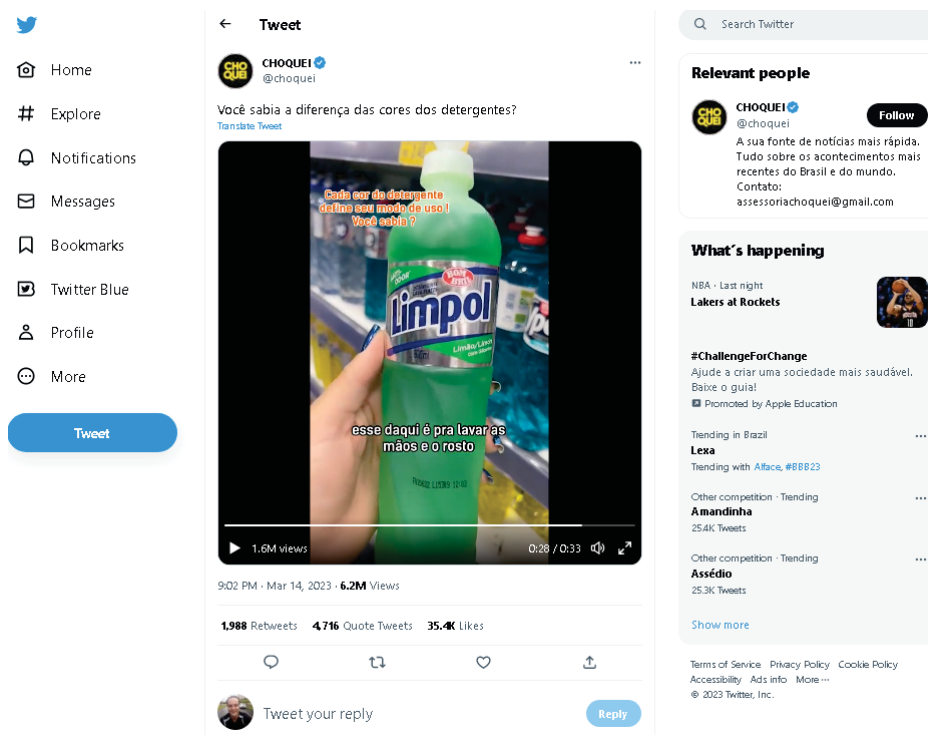


Figura 5.33: Desinformação para análise - Oficina III. Fonte: Twitter.

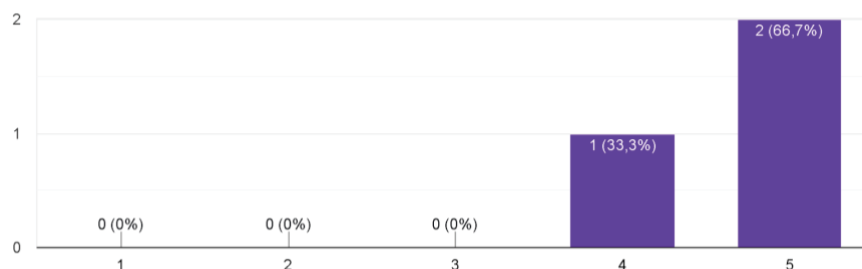
Essa oficina foi planejada para avaliar utilidade e facilidade de uso percebidas da abordagem. O método de coleta de dados foi o mesmo das outras oficinas, os participantes preencheram formulários e um questionário no encerramento do estudo. Esse estudo foi gravado em multimídia, o que permitiu a observação dos participantes aplicando a Dismiss. O questionário consistia de questões que usaram a escala Likert (onde "1" significava "discordo completamente" e 5 "Concordo completamente") e perguntas abertas. Após o preenchimento do questionário os participantes deram suas opiniões em um grupo focal.

Como nas outras duas oficinas, a utilidade percebida da Dismiss foi indicada pelos participantes tanto nas respostas do questionário quanto na entrevista com o grupo focal ao final do estudo. A Figura 5.34 mostra as respostas dos participantes em relação a utilidade percebida da abordagem para apoiar entendimento da natureza sociotécnica da desinformação, a identificação de aspectos sociotécnicos, e a identificação de aspectos sociotécnicos.



A Dismiss me ajudou a entender a natureza complexa do caso de desinformação.

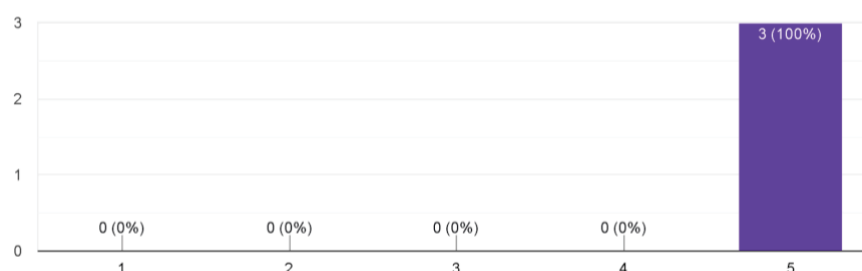
3 respostas



(a) Utilidade em apoiar entendimento da natureza sociotécnica da desinformação.

A Dismiss me ajudou a identificar aspectos sociotécnicos para serem considerados no desenvolvimento de soluções que ficariam ocultos sem o uso da abordagem.

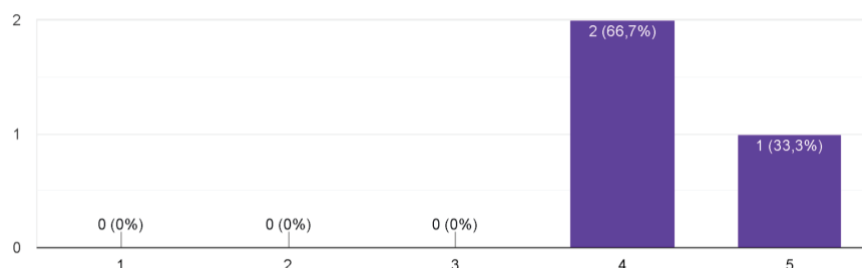
3 respostas



(b) Apoiar a identificação de aspectos sociotécnicos.

A Dismiss foi útil para identificar a problemas no caso de desinformação.

3 respostas



(c) Utilidade para a identificação de aspectos sociotécnicos.

Figura 5.34: Sobre a facilidade de uso dos materiais da Dismiss.

P1 respondeu sobre sua percepção em relação utilidade da abordagem:

*"Considero que a DISMISS tem grante utilidade para diversas partes interessadas, e deveria ser utilizada principalmente por desenvolvedores/funcionários de plataformas que estão sendo empregadas para disseminação da desinformação (Twitter, TikTok, YouTube, etc). Como em muitos casos parece não haver "boa vontade" das plataformas em mitigar o problema, é fundamental o apoio da sociedade civil no combate ao problema, vide exemplo da iniciativa Sleeping Giants. Considero que a DISMISS poderia ser difundida pelos membros da sociedade civil, por meio de divulgação científica, para conscientização do problema da desinformação e para pressionar as plataformas a tomarem medidas efetivas de combate ao problema."*

E P2:

*"Creio que a DISMISS auxilia a identificar aspectos relevantes do contexto do objeto desinformativo. Há vários conceitos relevantes, como os estágios, a influência de vários stakeholders nesses diferentes estágios, as características dos canais e do próprio objeto desinformativo... Me parece que uma abordagem iterativa é mais produtiva para uma análise abrangente e completa do que uma abordagem de preenchimento mais no modelo "cascata" de desenvolvimento. A cada vez que eu olho os artefatos, me dá vontade de voltar e re-preencher artefatos anteriores. Ou seja, no preenchimento de artefatos e na discussão mais posterior, novas descobertas ocorrem sobre o preenchimento de artefatos do início do processo."*

Além disso, a Oficina III foi atendida pelo participantes através da plataforma Miro<sup>21</sup>, onde eles puderam preencher os artefatos de modo direto. A Figura 5.35 mostra o artefato DPI preenchido pelos participantes. Cada participante usou um componente do tipo "Sticky Note" de uma cor diferente na plataforma Miro para preencher os artefatos.

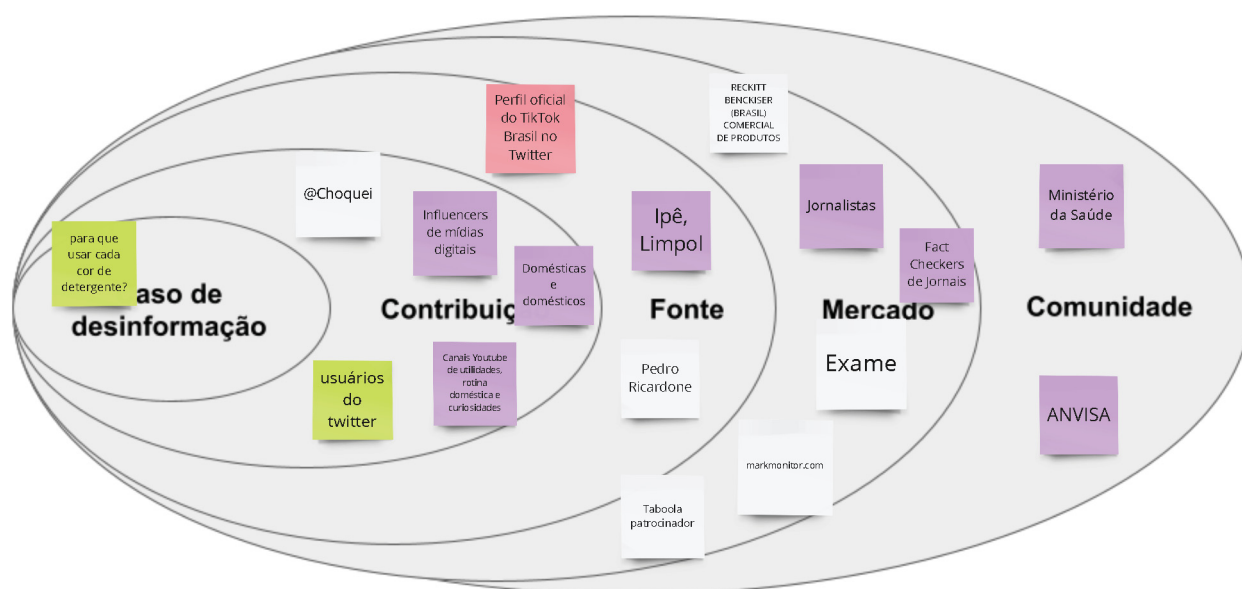


Figura 5.35: Diagrama de Identificação de Partes Interessadas Preenchido.

A Figura 5.36 mostra o artefato Escada Semiótica preenchido pelos participantes.

<sup>21</sup><https://miro.com/pt/>

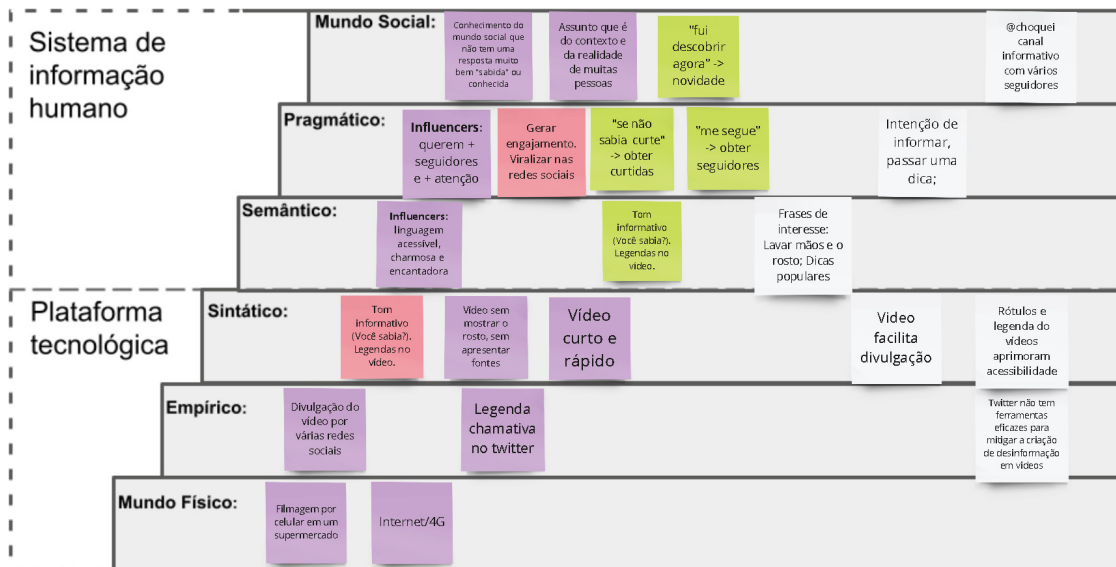


Figura 5.36: Escada Semiótica preenchida para o estágio de Consumo do Ciclo de Vida.

A Figura 5.37 mostra as sugestões de problemas encontrados na análise.

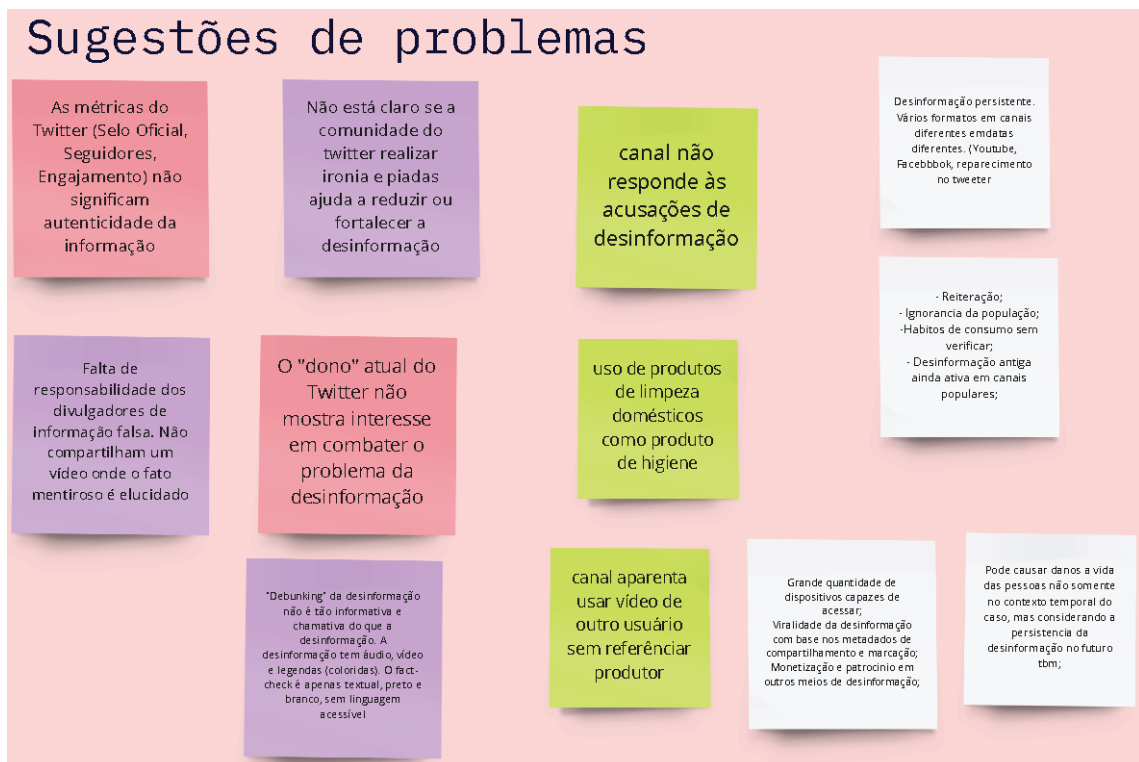
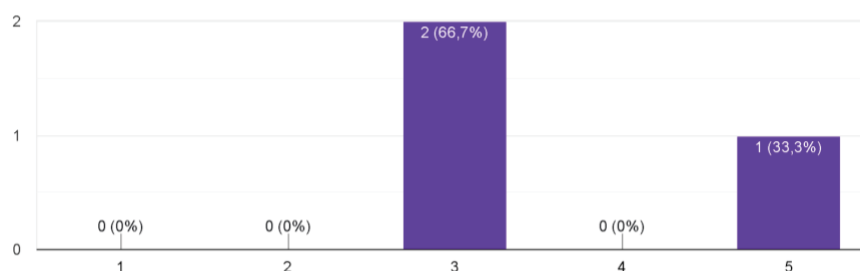


Figura 5.37: Quadro de sugestão de problemas preenchido.

Os materiais de apoio, os conceitos-chave e os artefatos da Dismiss foram simplificados. Os Princípios de Influência e as Propriedades do Canal de comunicação, foram incorporados as questões-guia na forma de exemplos de aspectos sociotécnicos relevantes para caracterização de casos de desinformação. Além disso, os conceitos do Modelo Do Ciclo de Vida foram mantidos e apresentados de modo inter-relacionado progressivamente na aplicação das etapas da Dismiss, com o objetivo de reduzir a carga teórica e conceitual. A Figura 5.38 apresenta os indícios sobre a percepção dos participantes em relação a facilidade de uso e entendimento da Dismiss.

A Dismiss é fácil de usar.

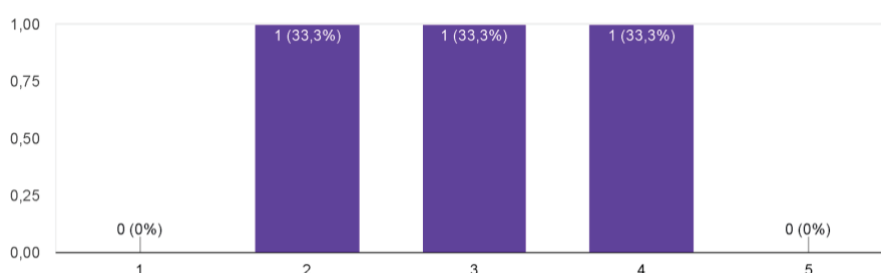
3 respostas



(a) Utilidade em apoiar entendimento da natureza sociotécnica da desinformação.

A Dismiss é fácil de entender.

3 respostas



(b) Apoiar a identificação de aspectos sociotécnicos.

Figura 5.38: Sobre a facilidade de uso dos materiais da Dismiss.

Em resposta ao questionário, P1 comentou sobre sua percepção em relação facilidade de uso da abordagem:

*"Não considero uma ferramenta fácil de usar, porém acredito que a complexidade da ferramenta está diretamente relacionada à complexidade do problema, e justamente por isso considero ser difícil a utilização de ferramentas "fáceis" e "simples" para combater o problema. Porém considero que o workshop foi importante para o entendimento do problema e da solução proposta e somado aos vídeos de Divulgação Científica acredito que auxiliam bastante no entendimento e utilização da DISMISS."*

e P2:

*"A etapa 1 é a mais fácil de lidar. Os estágios são fáceis de entender (mas difíceis de se analisar em um contexto social complexo onde a desinformação surge e habita). Os stakeholders são fáceis de entender, mas demanda um esforço para se identificar. Quando envolve a escada semiótica, o processo começa a ficar difícil. Enfrentar esse preenchimento sozinho, de forma individual, é mais difícil ainda. As seis dimensões semióticas, em diferentes estágios, é complexa. Pensar em seis dimensões semióticas, três vezes, demanda esforço cognitivo alto. Para mim foi uma dificuldade "separar" meu pensamento em três estágios diferentes. Na minha análise, os estágios apareciam todos misturados, e eu não sabia em qual estágio um aspecto realmente teria que ser classificado. Depois, na consolidação, acredito que uma discussão entre o que todos acharam seria benéfico para construirmos juntos uma visão sobre os problemas encontrados. Aí o conhecimento sobre a desinformação é*

*compartilhado entre todos os envolvidos. Entretanto, há uma série de materiais, indicações, recomendações, explicações que facilitam a tarefa. Talvez essa análise possa ser distribuída no tempo, não precisando fazer tudo ao mesmo momento ou no mesmo dia. Acredito que a análise com a DISMISS tem utilidade, e pode ser melhor aproveitada se for realizada de maneira coletiva em grupo, e separada em diferentes momentos (ou dias). Ou então, pode ser preenchida de forma iterativa, preenchendo uma análise rapidamente sem os conceitos de "tipo de canal" e "tipo de objeto", que depois é revisitada posteriormente. Em uma abordagem iterativa, depois o preenchimento dos artefatos da DISMISS pode ser revisitado a partir dos conceitos de "tipo de canal" e "tipo de objeto", deixando a análise ainda mais robusta."*

#### 5.3.4 Discussão sobre as oficinas

A Dismiss foi proposta e elaborada com o propósito de apoiar os projetistas de soluções computacionais no entendimento da desinformação como um fenômeno sociotécnico, auxiliando na caracterização do contexto de um caso de desinformação e na definição de problemas. Os estudos feitos durante as oficinas foram desenhados para avaliar a percepção dos participantes sobre as capacidades da Dismiss em atingir esse propósito. A análise feita nos dados de cada estudo, triangulando três fontes de coleta diferentes (artefatos preenchidos, respostas em questionários, e grupo focal), indicam que a abordagem tem o potencial de apoiar o estudo sociotécnico do fenômeno da desinformação e o entendimento de sua natureza complexa. Apesar disso, os participantes sugerem que a Dismiss necessita de mais refinamento para diminuir sua complexidade de uso e aplicação. Além disso, as oficinas auxiliaram na identificação das dificuldades no uso da abordagem e informaram o desenvolvimento de suas melhorias e refinamentos, continuamente. Nesse sentido, a Dismiss representa um caminho promissor para o estudo sociotécnico do fenômeno, auxiliando no diagnóstico de casos de desinformação e na caracterização do contexto da comunicação.

Sobre a percepção de utilidade da Dismiss, os participantes reportaram a utilidade do Modelo Conceitual do Ciclo de Vida da Desinformação Digital para estruturar o estudo de um caso e compreender melhor o contexto de comunicação desde a primeira oficina. Os artefatos da SO foram úteis na identificação de partes interessadas e aspectos sociotécnicos relevantes em todos os estudos. No entanto, tanto o Ciclo de Vida quanto os artefatos da SO necessitaram de melhorias nas instruções de uso e simplificação de conceitos, buscando reduzir as dificuldades encontradas pelos participantes no uso da Dismiss a cada nova oficina. Nesse sentido, a facilidade de uso da Dismiss ainda é uma de suas fragilidades.

A quantidade e características dos participantes, o formato contínuo do estudo e alterações em variáveis que deixaram a análise dos dados complexa, podem ser elencadas como fragilidades dos estudos que podem impactar os resultados. Os participantes tiveram experiência prévia com a pesquisa do autor em outros momentos de seu desenvolvimento, o que pode impactar na análise dos dados sobre a utilidade da Dismiss. Apesar dessa experiência prévia ser relacionada à avaliação de resultados do mapeamento sobre a desinformação e de aspectos teóricos e conceituais do fenômeno, os participantes não tinham noção de como a Dismiss ligaria estes conhecimentos ou a estrutura e técnicas de análise que a compõem.

Além disso, os participantes tiveram contato com a Dismiss continuamente, tendo a oportunidade de aprender cada vez mais sobre ela em cada nova oficina. Isso pode introduzir viés na análise de dados quanto à utilidade percebida pelos participantes, considerando que a cada nova oficina eles se familiarizavam com os conceitos e artefatos da abordagem. No entanto, as respostas da percepção dos participantes quanto à utilidade e ao apoio que a Dismiss proporciona foram

positivas desde o primeiro contato com a abordagem, indicando a adequação dos componentes da Dismiss para resolver o problema proposto.

Ainda, um dos estudos teve alguns problemas no seu desenvolvimento que não foram detectados antes da sua condução. A possibilidade dos participantes escolherem entre dois casos de desinformação para conduzir a análise sociotécnica dificulta a comparação dos resultados da Oficina II. Entretanto, considerando que o autor buscava por repostas quanto as melhorias na facilidade de uso da Dismiss nas Oficinas II e III, esta fragilidade teve pouco impacto na análise de dados quanto as dificuldades no uso e aplicação da Dismiss, tendo até proporcionado a identificação de dificuldades relacionadas a diferentes tipos de objetos. Além disso, quando triangula-se as respostas das três oficinas, aparece a tendência positiva na percepção de utilidade da Dismiss pelos participantes, indicando sua capacidade de solucionar o problema de pesquisa (apoiar o entendimento sociotécnico do fenômeno).

#### 5.4 CONSIDERAÇÕES SOBRE AS AVALIAÇÕES

As avaliações tiveram o objetivo de apoiar o desenvolvimento constante da DISMISS. A avaliação por grupo focal investigou fragilidades e limitações em diferentes momentos da pesquisa, possibilitando refinar o escopo do problema, os fundamentos teóricos, os métodos de investigação, e a solução proposta. As avaliações em pequena escala investigaram as capacidades de diferentes componentes da Dismiss de resolver o problema, buscando indícios sobre sua eficiência e efetividade. As oficinas avaliaram a capacidade da Dismiss de auxiliar projetistas no entendimento da natureza sociotécnica da desinformação digital.

Os grupos focais que avaliaram a pesquisa em diferentes momentos contribuíram para o desenvolvimento de uma metodologia adequada para a pesquisa, refletidos na revisão dos protocolos do mapeamento e dos estudos, e nas discussões dos resultados de suas aplicações. Além disso, a escolha de componentes teóricos da abordagem priorizou aquelas que foram utilizadas "amplamente" na literatura, como a SO e a AE sendo teorias com aplicações em diversas áreas do conhecimento. A união de diferentes teorias (de uso reconhecido) com o objetivo de mitigar as fragilidades de cada uma, bem como os apontamentos dos grupos focais e as provas de conceito realizadas nos estudos em pequena escala, representam um cuidado com o rigor e abrangência com que a abordagem foi desenvolvida. Esses são indícios da Confiabilidade <sup>22</sup> da solução proposta.

Os resultados dos estudos em pequena escala indicaram a capacidade dos componentes da Dismiss, o Modelo do ciclo do Ciclo de Vida em apoiar o estudo Sistemico do fenômeno e os Artefatos da SO em apoiar o estudo sociotécnico da desinformação. Os aspectos sociotécnicos identificados em cada análise feita nos estudos em pequena escala, auxiliaram na compreensão de fatores contextuais relevantes do fenômeno e em como eles podem ser abordados durante a análise. Por exemplo, aspectos como a intencionalidade, a popularidade, a facilidade de acesso ao meio de comunicação, a estruturação da mensagem, e a sincronicidade no canal de comunicação, revelaram novos papéis que as partes interessadas podem desempenhar, como oportunistas, autoridades influenciadoras e verificadores independentes, além de partes interessadas que podem ser impactadas indiretamente, como no caso dos clientes da empresa RSA que tiveram seus dados roubados. Os estudos em pequena escala e as oficinas contribuíram para que a Dismiss fosse abrangente em relação aos aspectos relevantes do problema que ela busca tratar e indicam a sua Efetividade <sup>23</sup>. Além disso, os estudos também viabilizaram aprimoramentos para o Modelo do

<sup>22</sup>Validade empírica; confiabilidade; replicabilidade; reprodutibilidade; robustez(Oulasvirta e Hornbæk, 2016).

<sup>23</sup>Capturar os aspectos essenciais do problema (Oulasvirta e Hornbæk, 2016).



ciclo do Ciclo de Vida e para a análise sociotécnica, simplificando a DISMISS e auxiliando na elaboração dos materiais de apoio para sua aplicação.

Ainda em relação aos estudos em pequena escala, notou-se o potencial da Dismiss em analisar um caso em profundidade (análise ao nível de postagem). Nesse tipo de análise, a Dismiss pode auxiliar na caracterização detalhada dos elementos que compõem um caso de desinformação. Os artefatos possibilitam desenhar um perfil das partes interessadas, entendendo seu contexto de vida, suas perspectivas quanto o canal de comunicação, e como podem ser impactadas e influenciadas pelos elementos do objeto desinformativo. Isso amplia as possibilidades para estudar fatores humanos e aspectos sociotécnicos relacionados ao fenômeno da desinformação. Além disso, a análise de um ataque cibernético revelou o potencial de se aplicar a Dismiss para outros fins, indicando a sua capacidade Transferência <sup>24</sup>.

As análises em largura dos estudos em pequena escala (análise ao nível de caso) mostraram o potencial da Dismiss em caracterizar de modo abrangente o contexto de um caso de desinformação, desenhando o "ecossistema" situado do fenômeno de modo incremental. Essas análises permitiram construir uma perspectiva da estrutura organizacional de um caso de desinformação e como os elementos de um caso afetam e são afetados uns pelos outros. Isso ampara o entendimento sobre a diversidade de aspectos que podem estar inter-relacionados em diferentes dimensões de abstração, e revela mais indícios sobre a capacidade de Transferência <sup>25</sup> da abordagem.

Sobre a reiteração e a mitigação, os resultados das análises dos estudos em pequena escala apontaram fatores relevantes para acompanhar as transformações de objetos desinformativos e no escopo de casos. Por exemplo, a efetividade de estratégias de mitigação parece depender não só de sua adaptabilidade ao contexto, mas de ações integradas em diferentes dimensões semióticas inter-relacionadas aos estágios do ciclo de vida. Por ser abordado como um fenômeno sociotécnico, o contexto em que a desinformação deliberada ocorre está em constante transformação, sejam significados e consensos, sejam as ferramentas de comunicação, ou mesmo as partes interessadas. Nesse sentido, as análises indicaram a capacidade da Dismiss em revelar diferentes necessidades de medidas de mitigação em diferentes estágios do ciclo, considerando diferentes requisitos relacionados ao contexto e as partes interessadas em cada caso.

Além disso, as análises destes estudos demonstram a complexidade de se lidar com o fenômeno. São diversos aspectos sociotécnicos presentes em diferentes níveis de abstração para cada estágio do ciclo de vida da desinformação ligados a diferentes partes interessadas que podem ser impactadas de modos específicos. Nesse sentido, a Dismiss tem o potencial de auxiliar no diagnóstico de casos de desinformação, estruturando a análise de aspectos sociotécnicos no ciclo de vida da desinformação, apoiando a identificação de partes interessadas e aspectos sociotécnicos relevantes para o entendimento do fenômeno e para o desenvolvimento de intervenções.

As oficinas mostraram a competência da Dismiss como ferramenta para estudar o fenômeno da Desinformação de modo sociotécnico. Nas três oficinas, os participantes foram assertivos em apontar a utilidade percebida da Dismiss e seu potencial para amparar o entendimento sociotécnico de casos de desinformação. No entanto, a granularidade com que se deseja abordar o fenômeno torna a Abordagem complexa e difícil de usar. Nesse sentido, a abordagem foi refinada para buscar simplificar sua aplicação, sem perder a abrangência com que se propõe estudar o fenômeno. As avaliações dos artefatos e o modelo do ciclo de vida serviram para reduzir a complexidade conceitual e a explicação de como usá-los. As avaliações sobre os materiais de

<sup>24</sup>Número de usuários, tarefas e contextos para os quais a solução pode ser aplicado; contextos qualitativamente novos em que a solução pode ser aplicada (Oulasvirta e Hornbæk, 2016).

<sup>25</sup>Número de usuários, tarefas e contextos para os quais a solução pode ser aplicado; contextos qualitativamente novos em que a solução pode ser aplicada (Oulasvirta e Hornbæk, 2016).

apoio e manuais de aplicação indicaram melhorias para resumi-los e torná-los mais objetivos. Como trabalhos futuros, a Dismiss precisa de mais refinamento para simplificar seu uso, como o desenvolvimento de estratégias mais simples para estudar um caso, além de ampliar a abordagem para melhor amparar o entendimento de problemas e a ideação de soluções.

Foram criados materiais de apoio e passos simplificados para a aplicação da Dismiss, com base na análise dos dados preenchidos nos formulários e no uso de artefatos dos participantes das oficinas de análise sociotécnica de desinformação. Os materiais de apoio fornecem um eixo complementar para sustentar na análise sociotécnica, apoiando a identificação de aspectos que de outra forma dependeriam da experiência do analista conduzindo a análise sociotécnica. Entretanto, eles trazem uma carga para aplicação da DISMISS que acarreta em cansaço e facilita a dispersão do analista. Isso representa uma fragilidade da abordagem relacionada a Eficiência <sup>26</sup> da Dismiss, considerando a dificuldade que os participantes das oficinas tiveram para aplicar a abordagem. Por esse motivo optou-se por transformar os materiais de apoio em um componente opcional da aplicação da Dismiss e incorporar parte do conteúdo desses materiais em questões-guia como uma forma de simplificação.

#### 5.4.1 Nota sobre os estudos

O filósofo da ciência, Paul Feyerabend, discute em sua obra (Regner, 1996), a impossibilidade de permanecer “fora” de um tema de estudo como pesquisador. Como um agente que estuda a desinformação, as análises realizadas nos estudos em pequena escala foram moldadas pela subjetividade de minhas experiências na vida, além das lentes práticas e teóricas que adotei. Em vez de tratar esses elementos intersubjetivos como um problema de confiabilidade, busquei prestar atenção analítica a eles para tentar oferecer uma compreensão mais rica da desinformação como um fenômeno sociotécnico. Para fazer isso, busquei seguir algumas tradições de análises qualitativas para promover a reflexão na pesquisa (Petersen et al., 2015; Wieringa et al., 2006; Feyerabend, 1989). Alguns passos específicos tomados para promover a reflexividade incluem: 1) registrar, refletir e dialogar sobre como o trabalho está moldando e sendo moldado por minhas crenças, valores, perspectivas políticas e suposições epistêmicas com outros pesquisadores em IHC (Sessões de grupo focal); e 2) relatar brevemente como esses fatores podem ter entrado em jogo durante o processo de pesquisa.

Um exemplo de discussão registrada: observei que minha própria origem e posição social dificultam a problematização de certos aspectos do fenômeno. Por exemplo, nos estudos em pequena escala, quando as mensagens carregadas de valores éticos e políticos amplificam uma determinada tendência de argumentação das discussões, me vi lutando com dúvidas e ceticismo sobre minhas interpretações dos efeitos de aspectos sociotécnicos em cada estágio do ciclo de vida. Percebi algumas dissonâncias relacionadas a crenças políticas e expectativas pessoais. Por exemplo, mensagens que criticam evidências científicas sobre vacinação na pandemia, com perspectivas negacionistas, geraram tensões consideráveis a respeito de conclusões potencialmente enviesadas nas análises de desinformações com a temática sobre saúde na pandemia. Aflorar essas tensões me ajudou a ser íntimo em minha pesquisa e desenvolver discussões em diversas dimensões do método de pesquisa que adotei, além de me ajudar a identificar pesquisas futuras que pretendo desenvolver.

---

<sup>26</sup>Quanto esforço ou recursos são necessários para criar ou implantar a solução; escalabilidade; tamanho (Oulasvirta e Hornbæk, 2016).

## 6 CONCLUSÃO

Abordar o fenômeno da desinformação digital requer métodos, técnicas e artefatos que apoiem o entendimento de sua natureza sociotécnica complexa. A tecnologia enquadra a informação em formatos que empobrecem a qualidade do que é expressado e limitam a capacidade de interpretação das pessoas que se comunicam. Além disso, ela interconecta partes interessadas de diversas realidades que realizam atividades e possuem entendimentos e objetivos diferentes. A inter-relação entre as partes interessadas e as tecnologias que utilizam incorpora diversos aspectos sociotécnicos em diferentes níveis de abstração que se relacionam aos diferentes estágios do ciclo de vida da desinformação. Nesse sentido, existem diversas dimensões de operação simultâneas no fenômeno, representando um desafio para o entendimento abrangente de problemas da desinformação digital.

O objetivo geral desta pesquisa foi investigar e desenvolver uma abordagem capaz de apoiar o entendimento da desinformação como um fenômeno sociotécnico de forma estruturada. A Dismiss é uma abordagem proposta para auxiliar a análise e o diagnóstico de casos de desinformação, identificando aspectos sociotécnicos relevantes e apoiando a caracterização do contexto da comunicação de desinformação para ampliar o entendimento abrangente do fenômeno. Como parte dos objetivos específicos, mapeamos a literatura em Ciência da Computação para compreender os aspectos desafiadores de abordar a desinformação digital. Com base nos resultados do mapeamento sistemático da literatura, caracterizamos os tipos de desinformação digital pela perspectiva sociotécnica, e a identificação de aspectos relevantes com potencial de influenciar a construção, difusão e consumo (parte do objetivo específico 2). Além disso, o mapeamento informou a construção do Modelo Conceitual do Ciclo de Vida da Desinformação Digital para estruturar o estudo de casos de desinformação e apoiar o diagnóstico de forma sistêmica (objetivo específico 3). O Ciclo foi incorporado a um arcabouço com artefatos da SO (Semiótica Organizacional) e métodos de análise de conteúdo que auxiliam a análise sociotécnica de casos de desinformação digital, identificando e caracterizando os elementos de um caso e o contexto do ciclo de vida da desinformação (objetivos específicos 2 e 3).

Como parte de um método de pesquisa construtivo, a DISMISS foi desenvolvida de forma incremental, sendo avaliada por especialistas em IHC (Interação Humano-Computador) em diferentes fases do seu desenvolvimento (objetivo específico 4). Por meio de análises de diferentes casos de desinformação digital realizadas pelo autor (estudos em pequena escala), e de discussões com um Grupo Focal compostos por pesquisadores em Interação Humano Computador, foram escolhidos os componentes e métodos da DISMISS. Os estudos auxiliaram na caracterização dos casos e na identificação de aspectos relevantes a serem analisados, além de informarem o desenvolvimento dos passos e das técnicas de utilização dos seus componentes. A Dismiss foi avaliada em oficinas de análise sociotécnica de desinformações com representantes do público-alvo. Os resultados dessas oficinas indicam a capacidade da Dismiss em auxiliar na compreensão da natureza sociotécnica complexa do fenômeno, apoiando a caracterização do contexto de casos de desinformação e no entendimento de problemas. Ainda, as oficinas informaram continuamente os refinamentos para facilidade de aplicação e uso da Dismiss.

O desenvolvimento da Dismiss incorporou os elementos básicos de um caso (partes interessadas, canais de comunicação e contextos), os diferentes estágios (criação, difusão, consumo) que possuem aspectos particulares em diferentes níveis de abstração (social, semântica, física, etc.) que podem estar inter-relacionados de diferentes formas. No entanto, a abrangência com que a Dismiss aborda o fenômeno e sua própria natureza complexa levaram

ao desenvolvimento de uma abordagem de difícil aplicação. Nesse sentido, os resultados das avaliações feitas nas oficinas informaram melhorias para a facilidade de uso da abordagem, buscando clarificar o entendimento dos componentes da Dismiss e das diferentes dimensões relevantes ao fenômeno (objetivos específicos 4 e 5). As oficinas se mostraram um caminho promissor para melhorias futuras na abordagem.

Na sua versão atual, a Dismiss possibilita abordar o fenômeno de forma abrangente, identificando partes interessadas em diferentes estágios do ciclo de vida da desinformação digital e aspectos sociotécnicos em diferentes níveis de abstração. O modelo conceitual do Ciclo permite abordar o contexto da comunicação para análise sociotécnica, construindo constantemente o entendimento sobre a estrutura (ecossistema) de um caso em análise. Dessa forma, a abordagem auxilia no entendimento organizacional e na linha de tempo de um caso de desinformação, relacionando objetos desinformativos, partes interessadas e contextos de comunicação diferentes. As duas formas que a Dismiss usa para abordar o fenômeno (por postagem ou por caso) dão flexibilidade à sua aplicação. A análise em profundidade (nível de postagem) auxilia na caracterização detalhada dos elementos diretamente relacionados ao contexto de um caso de desinformação, permitindo análises capazes de delinear o perfil das partes interessadas e investigações sobre como aspectos sociotécnicos podem influenciar a criação, difusão e consumo da desinformação. A análise em largura (nível de caso) permite construir uma visão do ecossistema do caso de desinformação, relacionando aspectos sociotécnicos relevantes nos diferentes estágios do ciclo de vida, identificando padrões em objetos desinformativos e revelando características da difusão. As avaliações da Dismiss nas oficinas indicaram o seu potencial de auxiliar no estudo do fenômeno e no entendimento de sua natureza sociotécnica complexa. Dessa forma, a Dismiss representa uma ferramenta promissora para abordar o fenômeno da desinformação digital de forma sociotécnica e sistêmica, com capacidade para auxiliar no entendimento de problemas de forma abrangente, sendo um caminho propício para compreender detalhes de como ocorrem, seus impactos na sociedade e como podem ser mitigados.

Existem limitações e ameaças à validade da pesquisa em diferentes pontos. Na fundamentação teórica, apesar da Dismiss incorporar artefatos e métodos utilizados amplamente na literatura, os conceitos desenvolvidos pelo autor são genéricos e abrangentes, podendo não serem adequados para casos e situações não consideradas quando de sua criação. Para amenizar os impactos da integração de diferentes teorias, o autor buscou discutir a forma como ele entende a compatibilidade dessas teorias e registrar suas decisões conceituais e teóricas para o desenvolvimento da Dismiss (Apêndice A.1.1). Ainda em relação à fundamentação teórica, a Dismiss é limitada em sua capacidade de lidar com a complexidade de abordar pontos como a subjetividade na interpretação do que é verdade e do que é mentira, as transformações de significado dentro de comunidades, e a adoção de novas crenças - fatores fundamentais para o entendimento abrangente do fenômeno. Nesse sentido, ela representa um esforço inicial para abordar aspectos subjetivos na comunicação através de TIC's, em especial no fenômeno da desinformação, que necessita de adaptações para operacionalizar o estudo aprofundado de aspectos subjetivos, não sendo uma solução especializada, mas sim generalista. Outra limitação da Dismiss reside na abrangência com que ela aborda o fenômeno, trazendo em si uma carga teórica e conceitual densa que sobrecarrega o seu usuário e aumenta a complexidade de seu uso. Em relação ao mapeamento sistemático da literatura, além das ameaças discutidas no Capítulo 3, a validade das conclusões deve ser observada no período em que o estudo foi feito, podendo estar ultrapassadas, carecendo de atualização. Ainda, a avaliação da Dismiss foi realizada considerando uma amostragem pequena de usuários e casos, cujas opiniões representam indícios da utilidade da Dismiss, mas não podem ser considerados evidências absolutas que comprovam sua utilidade para todos os casos do fenômeno da desinformação.

Como contribuições da pesquisa desenvolvida nessa tese, destacamos:

- O mapeamento sistemático oferece uma visão abrangente das abordagens, desafios e oportunidades de pesquisa no estudo da desinformação, considerando a perspectiva sociotécnica. Esse mapeamento possibilitou compreender a importância de investigar os elementos essenciais de um caso de desinformação, a caracterização de diferentes tipos de desinformação digital e os diversos fatores que atuam simultaneamente em diferentes dimensões do fenômeno. Esses resultados são valiosos para diagnosticar casos de desinformação, indo além da aplicação da Dismiss.
- As contribuições conceituais fornecem um entendimento mais sólido do conceito de desinformação, abordando os aspectos subjetivos das partes interessadas, a estruturação da comunicação em casos de desinformação e seus elementos essenciais, além das relações em cada estágio do ciclo de vida da desinformação. Essas contribuições resultam da adaptação de lentes e artefatos da Semiótica Organizacional, do Modelo do Ciclo de Vida da Desinformação Digital, do método de Análise de Enquadramento e dos Princípios de Influência. A tese também discute as limitações de outras teorias e artefatos propostos e utilizados na literatura para lidar com a natureza sociotécnica da desinformação (A.1.1).
- A contribuição de um método de análise de desinformação capaz de abordar os desafios do fenômeno da desinformação digital, relacionados à compreensão das partes interessadas, canais de comunicação e aspectos sociotécnicos em diferentes níveis de abstração, interconectados nos estágios de vida da desinformação. Esse método permite uma compreensão abrangente dos problemas inter-relacionados em diferentes dimensões do fenômeno.
- Os resultados das análises em pequena escala, revelam fatores de influência relevantes em casos reais de desinformação, ampliando o entendimento do fenômeno e fornecendo suporte para a aplicação da abordagem Dismiss, juntamente com o manual e o passo a passo. Essas análises também indicam a necessidade e a importância de soluções de mitigação da desinformação que considerem os aspectos das diferentes dimensões do fenômeno e suas inter-relações entre partes interessadas, tecnologias e os estágios do ciclo de vida de forma integrada.
- As publicações dos resultados de estudos em pequena escala e das discussões conceituais sobre o fenômeno também são contribuições que auxiliam a divulgação científica do conhecimento gerado pela pesquisa.

## 6.1 TRABALHOS FUTUROS

No decorrer do trabalho, surgiram desdobramentos relacionados à fundamentação teórica e à construção de um entendimento do fenômeno da desinformação digital sob a ótica da Semiótica Organizacional, além do potencial de aplicação da abordagem para outros propósitos, como análises forenses relacionadas ao uso de desinformação como arma em segurança digital. A Seção A.1.1 apresenta uma discussão que aborda a clarificação de alguns desafios enfrentados no desenvolvimento da Dismiss. São discutidas adaptações de conceitos como normas, *affordances* tecnológicas e digitais, e os desafios da análise de intenções (Pragmática), além de apresentar uma discussão sobre a integração do método de Análise de Enquadramento (AE) e da Semiótica Organizacional (SO), adaptando-os para análise sociotécnica de desinformações. Esses trabalhos



constituem pontos de interesse do autor e novos caminhos para desenvolver um entendimento mais refinado da relação entre aspectos sociotécnicos e a manipulação da audiência. Em especial, o autor pretende refinar a integração da SO com AE para análise de interação. Outros desdobramentos teóricos são discutidos no Apêndice A.2, onde o autor apresenta um entendimento sobre "vulnerabilidades humanas" passíveis de exploração pela desinformação digital.

Ainda com base nos resultados das avaliações da Dismiss, percebemos que o fenômeno da desinformação exige medidas de mitigação que estejam integradas em diferentes estágios do ciclo de vida da desinformação, como medidas de prevenção, previsão e tolerância, além de estratégias capazes de abordar os aspectos sociotécnicos em diferentes níveis de abstração, para aprimorar a capacidade de lidar com a desinformação na esfera digital. O autor apresenta uma classificação baseada nos resultados da literatura para as Estratégias de Mitigação no fenômeno da desinformação que podem ser úteis no desenvolvimento de intervenções (Apêndice A.2). As inter-relações entre aspectos socioculturais e técnicos/tecnológicos podem catalisar consequências prejudiciais, dependendo das circunstâncias e particularidades de cada uma das partes interessadas (fatores como interesses e as capacidades de ação no meio de comunicação, nível de educação, fatores humanos, tipo do dispositivo, entre outros). Nesse sentido, as estratégias e intervenções precisam ser projetadas considerando as particularidades de cada contexto em relação a cada estágio do ciclo de vida e, principalmente, estar em consonância umas com as outras. Cada estratégia e mecanismo de mitigação encontrado no mapeamento da literatura é capaz de lidar com aspectos específicos do fenômeno e operar em dimensões específicas (Sistema de Informação Humana ou Plataforma Tecnológica) e por isso a integração deles representa maiores chances de gerar soluções efetivas e com danos colaterais reduzidos. Os resultados das avaliações revelaram indícios do potencial da Dismiss para auxiliar no entendimento sistêmico do fenômeno e amparar o desenvolvimento de intervenções integradas nos estágios do ciclo de vida da desinformação. Como trabalho futuro na direção de instrumentalizar e apoiar o desenvolvimento de estratégias de mitigação integradas, o autor pretende incorporar a classificação de Estratégias de Mitigação e uma forma de apoiar o analista a decidir quais mecanismos e estratégias são necessários em cada dimensão do fenômeno, considerando o contexto de aplicação em cada caso de desinformação.

Também haverá trabalhos futuros que investigarão modos de aprimorar a facilidade de uso e a utilidade da Dismiss. São necessários refinamentos para simplificar e direcionar as diferentes formas de uso da abordagem, tornando os materiais de apoio e manuais mais objetivos e contextualizados. Novas oficinas podem auxiliar nessa tarefa. Além disso, existe a possibilidade de avaliar a abordagem com outros públicos, como legisladores e jornalistas, para coletar indícios de sua utilidade para além do público da Computação. Avaliações com especialistas de diferentes áreas podem contribuir para aprimorar a facilidade de uso e utilidade da Dismiss, considerando aspectos que foram ocultados nas considerações do público-alvo. Além disso, é possível avançar no desenvolvimento da abordagem para aprofundar a análise de casos de desinformação, objetivando apoiar a previsão dos impactos de aspectos sociotécnicos, levando em conta os diferentes perfis de partes interessadas e vulnerabilidades humanas.



## REFERÊNCIAS

- Abade, D. (2021). Daniela Abade no Twitter. [Online; accessed 23. Feb. 2021].
- Acemoglu, D. e Ozdaglar, A. (2011). Opinion dynamics and learning in social networks. Dynamic Games and Applications, 1(1):3–49.
- Adams, R. M. (1974). Theories of actuality. Nous, páginas 211–231.
- Addawood, A., Badawy, A., Lerman, K. e Ferrara, E. (2019). Linguistic cues to deception: Identifying political trolls on social media. Em Proceedings of the international AAAI conference on web and social media, volume 13, páginas 15–25.
- Afonso, N. (2020). Agência Lupa. [Online; accessed 8. Jun. 2021].
- Ahmad, A., Webb, J., Desouza, K. C. e Boorman, J. (2019). Strategically-motivated advanced persistent threat: Definition, process, tactics and a disinformation model of counterattack. Computers & Security.
- Ahmed, S., Hinkelmann, K. e Corradini, F. (2019). Combining machine learning with knowledge engineering to detect fake news in social networks-a survey. Em Proceedings of the AAAI 2019 Spring Symposium, volume 12.
- Ahsan, M., Kumari, M. e Sharma, T. (2019). Rumors detection, verification and controlling mechanisms in online social networks: A survey. Online Social Networks and Media, 14:100050.
- Aigner, J., Durchardt, A., Kersting, T., Kattenbeck, M. e Elswiler, D. (2017). Manipulating the perception of credibility in refugee related social media posts. Em Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval, páginas 297–300.
- Aïmeur, E., Díaz Ferreyra, N. e Hage, H. (2019). Manipulation and malicious personalization: exploring the self-disclosure biases exploited by deceptive attackers on social media. Frontiers in Artificial Intelligence, 2:26.
- Akhtar, M. S., Ekbal, A., Narayan, S., Singh, V. e Cambria, E. (2018). No, that never happened!! investigating rumors on twitter. IEEE Intelligent Systems, 33(5):8–15.
- Al-Ash, H. S. e Wibowo, W. C. (2018). Fake news identification characteristics using named entity recognition and phrase detection. Em 2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE), páginas 12–17. IEEE.
- Al-Rawi, A., Groshek, J. e Zhang, L. (2019). What the fake? assessing the extent of networked political spamming and bots in the propagation of# fakenews on twitter. Online Information Review.
- Al-Sarem, M., Boulila, W., Al-Harby, M., Qadir, J. e Alsaeedi, A. (2019). Deep learning-based rumor detection on microblogging platforms: A systematic review. IEEE Access, 7:152788–152812.

- Albahar, M. e Almalki, J. (2019). Deepfakes: Threats and countermeasures systematic review. Journal of Theoretical and Applied Information Technology, 97(22).
- Aldwairi, M. e Alwahedi, A. (2018). Detecting fake news in social media networks. Procedia Computer Science, 141:215–222.
- Alexander, J. e Smith, J. (2010). Disinformation: A taxonomy. IEEE Security & Privacy, 9(1):58–63.
- Alianto, H., Chandra, Y. U. et al. (2017). The smartphone for disseminating of fake news by the university students game player. Em 2017 International Conference on Information Management and Technology (ICIMTech), páginas 14–18. IEEE.
- Alkhair, M., Meftouh, K., Smaïli, K. e Othman, N. (2019). An arabic corpus of fake news: Collection, analysis and classification. Em International Conference on Arabic Language Processing, páginas 292–302. Springer.
- Allport, G. W. e Postman, L. (1947). The psychology of rumor.
- Almaliki, M. (2019a). Misinformation-aware social media: A software engineering perspective. IEEE Access, 7:182451–182458.
- Almaliki, M. (2019b). Online misinformation spread: A systematic literature map. Em Proceedings of the 2019 3rd International Conference on Information System and Data Mining, páginas 171–178.
- Almirante, J. e Marés, C. (2020). Agência Lupa. [Online; accessed 8. Jun. 2021].
- Alvim, M. (2020). O grupo de '10 mil' médicos pró-cloroquina que se aproximou de Bolsonaro com 'evento histórico' - BBC News Brasil. [Online; accessed 26. Feb. 2021].
- Andersen, P. B. (1997). A theory of computer semiotics: Semiotic approaches to construction and assessment of computer systems, volume 3. Cambridge University Press.
- Armano, G., Battiato, S., Bennato, D., Boratto, L., Carta, S. M., Di Noia, T., Di Sciascio, E., Ortis, A. e Recupero, D. R. (2018). Newsvallum: Semantics-aware text and image processing for fake news detection system. Em SEBD.
- Armstrong, C. L. e McAdams, M. J. (2009). Blogs of information: How gender cues and individual motivations influence perceptions of credibility. Journal of Computer-Mediated Communication, 14(3):435–456.
- Arroyo, A. O., Onorati, T. e Diaz, P. (2018). Quality assessment of social media: Lessons learnt from the literature. Em 2018 22nd International Conference Information Visualisation (IV), páginas 278–283. IEEE.
- Assiroj, P., Hidayanto, A. N., Prabowo, H., Warnars, H. L. H. S. et al. (2018). Hoax news detection on social media: A survey. Em 2018 Indonesian Association for Pattern Recognition International Conference (INAPR), páginas 186–191. IEEE.
- Avizienis, A., Laprie, J.-C., Randell, B. e Landwehr, C. (2004). Basic concepts and taxonomy of dependable and secure computing. IEEE transactions on dependable and secure computing, 1(1):11–33.

- Babaei, M., Chakraborty, A., Kulshrestha, J., Redmiles, E. M., Cha, M. e Gummadi, K. P. (2019). Analyzing biases in perception of truth in news stories and their implications for fact checking. Em Proceedings of the Conference on Fairness, Accountability, and Transparency, páginas 139–139.
- Badawy, A., Ferrara, E. e Lerman, K. (2018). Analyzing the digital traces of political manipulation: The 2016 russian interference twitter campaign. Em 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), páginas 258–265. IEEE.
- Baeth, M. J. e Aktas, M. S. (2019). Detecting misinformation in social networks using provenance data. Concurrency and Computation: Practice and Experience, 31(3):e4793.
- Bai, J., Kong, Q., Li, L., Wang, L. e Zeng, D. (2019). Exploring cognitive dissonance on social media. Em 2019 IEEE International Conference on Intelligence and Security Informatics (ISI), páginas 143–145. IEEE.
- Bailey, B. P., Konstan, J. A. e Carlis, J. V. (2001). The effects of interruptions on task performance, annoyance, and anxiety in the user interface. Em Interact, volume 1, páginas 593–601.
- Bailin, S. e Siegel, H. (2002). Critical thinking. The Blackwell guide to the philosophy of education, páginas 181–193.
- Bajaj, P., Kavidayal, M., Srivastava, P., Akhtar, M. N. e Kumaraguru, P. (2016). Disinformation in multimedia annotation: misleading metadata detection on youtube. Em Proceedings of the 2016 ACM workshop on Vision and Language Integration Meets Multimedia Fusion, páginas 53–61.
- Bandeli, K. K. e Agarwal, N. (2018). Analyzing the role of media orchestration in conducting disinformation campaigns on blogs.
- Barfar, A. (2019). Cognitive and affective responses to political disinformation in facebook. Computers in Human Behavior, 101:173–179.
- Bargar, A., Pitts, S., Butkevics, J. e McCulloh, I. (2019). Challenges and opportunities to counter information operations through social network analysis and theory. Em 2019 11th International Conference on Cyber Conflict (CyCon), volume 900, páginas 1–18. IEEE.
- Baroni, L. A., Puska, A. A., de Castro Salgado, L. C. e Pereira, R. (2021). Dark patterns: towards a socio-technical approach. Em Proceedings of the XX Brazilian Symposium on Human Factors in Computing Systems, páginas 1–7.
- Basili, V. R. e Rombach, H. D. (1988). The tame project: Towards improvement-oriented software environments. IEEE Transactions on software engineering, 14(6):758–773.
- Basol, M., Roozenbeek, J., Berriche, M., Uenal, F., McClanahan, W. P. e Linden, S. v. d. (2021). Towards psychological herd immunity: Cross-cultural evidence for two prebunking interventions against covid-19 misinformation. Big Data & Society, 8(1):20539517211013868.
- Bastos, M. T. e Mercea, D. (2019). The brexit botnet and user-generated hyperpartisan news. Social Science Computer Review, 37(1):38–54.
- Basu, K. (2019). Identification of the source (s) of misinformation propagation utilizing identifying codes. Em Companion Proceedings of The 2019 World Wide Web Conference, páginas 7–11.

- Bateson, G. (1955). A theory of play and fantasy. Psychiatric research reports.
- Baxter, G., Marcella, R. e Walicka, A. (2019). Scottish citizens' perceptions of the credibility of online political "facts" in the "fake news" era: an exploratory study. Journal of documentation, 75(5):1100–1123.
- Bay, M. (2018). Social media ethics: a rawlsian approach to hypertargeting and psychometrics in political and commercial campaigns. ACM Transactions on Social Computing, 1(4):1–14.
- Bazan, S. (2017). A new way to win the war. IEEE Internet Computing, 21(4):92–97.
- Bedard, M. e Schoenthaler, C. (2018). Satire or fake news: Social media consumers' socio-demographics decide. Em Companion Proceedings of the The Web Conference 2018, páginas 613–619.
- Berghel, H. (2017a). Lies, damn lies, and fake news. Computer, 50(2):80–85.
- Berghel, H. (2017b). Software sophistry and political sleight of hand. Computer, 50(1):82–87.
- Beskow, D. M. e Carley, K. M. (2019). Agent based simulation of bot disinformation maneuvers in twitter. Em 2019 Winter Simulation Conference (WSC), páginas 750–761. IEEE.
- Bessi, A., Petroni, F., Del Vicario, M., Zollo, F., Anagnostopoulos, A., Scala, A., Caldarelli, G. e Quattrociocchi, W. (2015). Viral misinformation: The role of homophily and polarization. Em Proceedings of the 24th international conference on World Wide Web, páginas 355–356.
- Bevensee, E. e Ross, A. R. (2018). The alt-right and global information warfare. Em 2018 IEEE International Conference on Big Data (Big Data), páginas 4393–4402. IEEE.
- Bhuiyan, M. M., Zhang, K., Vick, K., Horning, M. A. e Mitra, T. (2018). Feedreflect: A tool for nudging users to assess news credibility on twitter. Em Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing, páginas 205–208.
- Boididou, C., Papadopoulos, S., Zampoglou, M., Apostolidis, L., Papadopoulou, O. e Kompatsiaris, Y. (2018). Detection and visualization of misleading content on twitter. International Journal of Multimedia Information Retrieval, 7(1):71–86.
- Bondielli, A. e Marcelloni, F. (2019). A survey on fake news and rumour detection techniques. Information Sciences, 497:38–55.
- Boshmaf, Y., Muslukhov, I., Beznosov, K. e Ripeanu, M. (2013). Design and analysis of a social botnet. Computer Networks, 57(2):556–578.
- Bossetta, M. (2018). A simulated cyberattack on twitter: Assessing partisan vulnerability to spear phishing and disinformation ahead of the 2018 us midterm elections. arXiv preprint arXiv:1811.05900.
- Bresilienne, L. (2021). r/coronabr. [Online; accessed 23. Feb. 2021].
- Brito, P. V. C. e J. (2021). Fabricante diz não haver evidência de que ivermectina funcione contra Covid-19. CNN Brasil.

- Bro, P. e Wallberg, F. (2015). Gatekeeping in a digital era: Principles, practices and technological platforms. Journalism Practice, 9(1):92–105.
- Broniatowski, D. A. e Reyna, V. F. (2019). To illuminate and motivate: a fuzzy-trace model of the spread of information online. Computational and Mathematical Organization Theory, páginas 1–34.
- Buchanan, T. e Benson, V. (2019). Spreading disinformation on facebook: Do trust in message source, risk propensity, or personality affect the organic reach of “fake news”? Social Media+ Society, 5(4):2056305119888654.
- Budak, C. (2019). What happened? the spread of fake news publisher content during the 2016 us presidential election. Em The World Wide Web Conference, páginas 139–150.
- Buffardi, K. e Edwards, S. H. (2014). Responses to adaptive feedback for software testing. Em Proceedings of the 2014 Conference on Innovation Technology in Computer Science Education, ITiCSE '14, página 165–170, New York, NY, USA. Association for Computing Machinery.
- Buntain, C. e Golbeck, J. (2017). Automatically identifying fake news in popular twitter threads. Em 2017 IEEE International Conference on Smart Cloud (SmartCloud), páginas 208–215. IEEE.
- Burbach, L., Halbach, P., Ziefle, M. e Calero Valdez, A. (2019). Who shares fake news in online social networks? Em Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization, página 234–242, New York, NY, USA. Association for Computing Machinery.
- Cafezeiro, I., Viterbo, J., da Costa, L. C., Salgado, L., da Costa Rocha, M. e Monteiro, R. S. (2017). Strengthening of the sociotechnical approach in information systems research. Sociedade Brasileira de Computação.
- Cai, G., Wu, H. e Lv, R. (2014). Rumors detection in chinese via crowd responses. Em 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014), páginas 912–917. IEEE.
- Campan, A., Cuzzocrea, A. e Truta, T. M. (2017). Fighting fake news spread in online social networks: Actual trends and future research directions. Em 2017 IEEE International Conference on Big Data (Big Data), páginas 4453–4457. IEEE.
- Campbell, J. A. (2001). In and out, scream and shout: an internet conversation about stock price manipulation. Em Proceedings of the 34th Annual Hawaii International Conference on System Sciences, páginas 10–pp. IEEE.
- Caputo, D. D., Pfleeger, S. L., Freeman, J. D. e Johnson, M. E. (2013). Going spear phishing: Exploring embedded training and awareness. IEEE Security & Privacy, 12(1):28–38.
- Carlsson, A. e Gustavsson, R. (2017). The art of war in the cyber world. Em 2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), páginas 42–44. IEEE.



- Caulfield, T., Spring, J. M. e Sasse, M. A. (2019). Why jenny can't figure out which of these messages is a covert information operation. Em Proceedings of the New Security Paradigms Workshop, páginas 118–128.
- CE, G. (2020). Fake News: Sistema imunológico em pleno funcionamento é importante, mas não evita Covid-19 - Governo do Estado do Ceará. [Online; accessed 8. Jun. 2021].
- Ceie-Sbc (2021). Design da interação e ambientes de aprendizagem baseados em tecnologia. Informática na Educação.
- Chan, M.-p. S., Jones, C. R., Hall Jamieson, K. e Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. Psychological science, 28(11):1531–1546.
- Chandra, Y. U. et al. (2017). Higher education student behaviors in spreading fake news on social media: A case of line group. Em 2017 International Conference on Information Management and Technology (ICIMTech), páginas 54–59. IEEE.
- Charles, A. C. e de Oliveira Sampaio, J. (2018). Checking fake news on web browsers: an approach using collaborative datasets.
- Che, X., Metaxa-Kakavouli, D. e Hancock, J. T. (2018). Fake news in the news: An analysis of partisan coverage of the fake news phenomenon. Em Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing, páginas 289–292.
- Chen, C., McCloskey, S. e Yu, J. (2018). Focus manipulation detection via photometric histogram analysis. Em Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, páginas 1674–1682.
- Chen, C.-Y., Shae, Z.-Y., Chang, C.-J., Lin, K.-Y., Tan, S.-M. e Chang, S.-L. (2019). A trusting news ecosystem against fake news from humanity and technology perspectives. Em 2019 19th International Conference on Computational Science and Its Applications (ICCSA), páginas 132–137. IEEE.
- Chen, H., Rahwan, I. e Cebrian, M. (2016a). Bandit strategies in social search: the case of the darpa red balloon challenge. EPJ Data Science, 5(1):20.
- Chen, W., Yeo, C. K., Lau, C. T. e Lee, B. S. (2016b). Behavior deviation: An anomaly detection view of rumor preemption. Em 2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), páginas 1–7. IEEE.
- Chen, X., Nong, Q., Feng, Y., Cao, Y., Gong, S., Fang, Q. e Ko, K.-I. (2017). Centralized and decentralized rumor blocking problems. Journal of Combinatorial Optimization, 34(1):314–329.
- Chen, X. e Sin, S.-C. J. (2013). 'misinformation? what of it?' motivations and individual differences in misinformation sharing on social media. Proceedings of the American Society for Information Science and Technology, 50(1):1–4.
- Chen, X., Sin, S.-C. J., Theng, Y.-L. e Lee, C. S. (2015a). Deterring the spread of misinformation on social network sites: A social cognitive theory-guided intervention. Proceedings of the Association for Information Science and Technology, 52(1):1–4.



- Chen, X., Sin, S.-C. J., Theng, Y.-L. e Lee, C. S. (2015b). Why do social media users share misinformation? Em Proceedings of the 15th ACM/IEEE-CS joint conference on digital libraries, páginas 111–114.
- Chen, Y., Conroy, N. J. e Rubin, V. L. (2015c). Misleading online content: recognizing clickbait as "false news". Em Proceedings of the 2015 ACM on workshop on multimodal deception detection, páginas 15–19.
- Cheung, C. M. e Lee, M. K. (2009). Understanding the sustainability of a virtual community: model development and empirical test. Journal of Information Science, 35(3):279–298.
- Chin, Y. S. e Zauddin, H. (2019). New media literacy and media use among university students in malaysia. International Journal of Engineering and Advanced Technology, 8(5C):469–474.
- Cho, J.-H., Alsmadi, I. e Xu, D. (2016). Privacy and social capital in online social networks. Em 2016 IEEE Global Communications Conference (GLOBECOM), páginas 1–7. IEEE.
- Cho, J.-H., Cook, T., Rager, S., O'Donovan, J. e Adali, S. (2017). Modeling and analysis of uncertainty-based false information propagation in social networks. Em GLOBECOM 2017-2017 IEEE Global Communications Conference, páginas 1–7. IEEE.
- Cho, J.-H., Rager, S., O'Donovan, J., Adali, S. e Horne, B. D. (2019). Uncertainty-based false information propagation in social networks. ACM Transactions on Social Computing, 2(2):1–34.
- Cholvy, L. (2014). How strong to believe information? a trust model in the logical belief function theory. Em TRUST@ AAMAS, páginas 44–54.
- Chomphoosang, P., Durresi, A., Durresi, M. e Barolli, L. (2012). Trust management of social networks in health care. Em 2012 15th International Conference on Network-Based Information Systems, páginas 392–396. IEEE.
- Chua, A. Y. e Banerjee, S. (2017). A study of tweet veracity to separate rumours from counter-rumours. Em Proceedings of the 8th International Conference on Social Media & Society, páginas 1–8.
- Chua, A. Y., Cheah, S.-M., Goh, D. H.-L. e Lim, E.-P. (2016). Collective rumor correction on the death hoax of a political figure in social media. Pacific Asia Conference on Information Systems.
- Cialdini, R. B. (1993). Influence: The psychology of persuasion.
- Cobb, M. D., Nyhan, B. e Reifler, J. (2013). Beliefs don't always persevere: How political figures are punished when positive information about them is discredited. Political Psychology, 34(3):307–326.
- CORDEIRO, P. R. D., Pinheiro, V., Moreira, R., Carvalho, C. e Freire, L. (2019). What is real or fake?-machine learning approaches for rumor verification using stance classification. Em IEEE/WIC/ACM International Conference on Web Intelligence, páginas 429–432.
- Cota, W., Ferreira, S. C., Pastor-Satorras, R. e Starnini, M. (2019). Quantifying echo chamber effects in information spreading over political communication networks. EPJ Data Science, 8(1):1–13.

- Coviello, C. A. (2021). Open Letter to RSA Customers. [Online; accessed 18. Nov. 2021].
- Cybenko, A. K. e Cybenko, G. (2018). Ai and fake news. IEEE Intelligent Systems, 33(5):1–5.
- Cybenko, G., Giani, A. e Thompson, P. (2002). Cognitive hacking: A battle for the mind. Computer, 35(8):50–56.
- da Saúde, M. (2021). SEI/MS - 0014934763 - Nota Informativa. [Online; accessed 3. Mar. 2021].
- da Silva Gomes, W. e Dourado, T. (2019). Fake news, um fenômeno de comunicação política entre jornalismo, política e democracia. Estudos em Jornalismo e Mídia, 16(2):33–45.
- Dang, A., Moh'd, A., Milios, E. e Minghim, R. (2016a). What is in a rumour: Combined visual analysis of rumour flow and user activity. Em Proceedings of the 33rd Computer Graphics International, páginas 17–20.
- Dang, A., Smit, M., Moh'd, A., Minghim, R. e Milios, E. (2016b). Toward understanding how users respond to rumours in social media. Em 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), páginas 777–784. IEEE.
- Dawson, A. e Innes, M. (2019). How russia's internet research agency built its disinformation campaign. The Political Quarterly, 90(2):245–256.
- de Freitas Melo, P., Vieira, C. C., Garimella, K., de Melo, P. O. V. e Benevenuto, F. (2019). Can whatsapp counter misinformation by limiting message forwarding? Em International Conference on Complex Networks and Their Applications, páginas 372–384. Springer.
- de Jongab, S. (2009). False information and the emergence of conflict. Em BNAIC 2009 Benelux Conference on Artificial Intelligence, página 107.
- Del Vicario, M., Gaito, S., Quattrocioni, W., Zignani, M. e Zollo, F. (2017). News consumption during the italian referendum: A cross-platform analysis on facebook and twitter. Em 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA), páginas 648–657. IEEE.
- Della Vedova, M. L., Tacchini, E., Moret, S., Ballarin, G., DiPierro, M. e de Alfaro, L. (2018). Automatic online fake news detection combining content and social signals. Em 2018 22nd Conference of Open Innovations Association (FRUCT), páginas 272–279. IEEE.
- Devi, M., Tyagi, H. e Jindal, A. (2019a). Fake news and rumour detection on social media. International Journal of Innovative Technology and Exploring Engineering, 8(11 Special Issue):1172–1175. cited By 0.
- Devi, M., Tyagi, H. e Jindal, A. (2019b). Fake news and rumour detection on social media. [Online; accessed 7. Dec. 2020].
- Devi, P. S. e Karthika, S. (2018). Veracity analysis of rumors in social media. Em 2018 International Conference on Computer, Communication, and Signal Processing (ICCCSP), páginas 1–4. IEEE.
- Dhar, J., Jain, A. e Gupta, V. K. (2016). A mathematical model of news propagation on online social network and a control strategy for rumor spreading. Social Network Analysis and Mining, 6(1):57.

- DiFranzo, D. e Gloria-Garcia, K. (2017). Filter bubbles and fake news. XRDS: Crossroads, The ACM Magazine for Students, 23(3):32–35.
- Dimo, P. (2019). Knowing, believing and thinking we know. Em 2019 22nd International Conference on Control Systems and Computer Science (CSCS), páginas 380–380. IEEE.
- Do Povo, G. (2021). Curitiba manda retirar outdoors que incentivam tratamento precoce contra covid-19. online; acessado 18 de abril. 2021.
- do Povo, G. (2021). Um em cada quatro pacientes intubados por Covid morre por sequelas após alta hospitalar. Gazeta do Povo.
- Dongo, I., Cardinale, Y. e Aguilera, A. (2019). Credibility analysis for available information sources on the web: A review and a contribution. Em 2019 4th International Conference on System Reliability and Safety (ICSRS), páginas 116–125. IEEE.
- dos Anjos, L. (2021). CNS pede que Ministério da Saúde retire publicações sobre tratamento precoce para Covid-19. [Online; accessed 18. Feb. 2021].
- Doty, C. (2015). Social epistemology and cognitive authority in online comments about vaccine safety. iConference 2015 Proceedings.
- D'angelo, P. (2018). Doing news framing analysis ii. Empirical and Theoretical Perspectives.
- Ecker, U. K., Swire, B. e Lewandowsky, S. (2014). Correcting misinformation—a challenge for education and cognitive science.
- Eco, U. (1979). A theory of semiotics, volume 217. Indiana University Press.
- El-Hani, C. N., Queiroz, J. e Stjernfelt, F. (2010). Firefly femmes fatales: A case study in the semiotics of deception. Biosemitotics, 3(1):33–55.
- Elkasrawi, S., Dengel, A., Abdelsamad, A. e Bukhari, S. S. (2016). What you see is what you get? automatic image verification for online news content. Em 2016 12th IAPR Workshop on Document Analysis Systems (DAS), páginas 114–119. IEEE.
- Erickson, T. e Kellogg, W. A. (2000). Social translucence: an approach to designing systems that support social processes. ACM transactions on computer-human interaction (TOCHI), 7(1):59–83.
- Erlandsson, F., Nia, R., Boldt, M., Johnson, H. e Wu, S. F. (2015). Crawling online social networks. Em 2015 Second European Network Intelligence Conference, páginas 9–16.
- Fávero, B., Barbosa, B. e Ely, D. (2021). Facebook exibiu anúncios de drogas sem comprovação contra Covid-19 ao menos 3,9 mi de vezes em 2020. Aos Fatos.
- Fagundes, E. (2023). É falso que governo Lula aumentou auxílio-reclusão para R\$ 1.754. [Online; accessed 3. Apr. 2023].
- Fan, D., Song, B., Zhang, Z.-H., Song, Y.-R. e Jiang, G.-P. (2019). Modeling and analysis of cascading failures considering risk awareness. Em 2019 Chinese Control Conference (CCC), páginas 8895–8898. IEEE.

- Farajtabar, M., Yang, J., Ye, X., Xu, H., Trivedi, R., Khalil, E., Li, S., Song, L. e Zha, H. (2017). Fake news mitigation via point process based intervention. arXiv preprint arXiv:1703.07823.
- Fard, A. E. e Cunningham, S. (2019). Assessing the readiness of academia in the topic of false and unverified information. Journal of Data and Information Quality (JDIQ), 11(4):1–27.
- Fathema, N., Shannon, D. e Ross, M. (2015). Expanding the technology acceptance model (tam) to examine faculty use of learning management systems (lmss) in higher education institutions. Journal of Online Learning & Teaching, 11(2).
- Fatos, A. (2020). Em seis meses de pandemia, Bolsonaro foi quem mais impulsionou apoio à cloroquina no Twitter brasileiro. [Online; accessed 23. Feb. 2021].
- Federal, M. P. (2023). Responsáveis por divulgação de publicidade em defesa do “tratamento precoce” contra covid são condenados a pagar R\$ 55 milhões por danos morais coletivos e à saúde. [Online; accessed 29. May 2023].
- Feng, V. W. e Hirst, G. (2013). Detecting deceptive opinions with profile compatibility. Em Proceedings of the sixth international joint conference on natural language processing, páginas 338–346.
- Fernandes, S., Raj, S., Ortiz, E., Vintila, I., Salter, M., Urosevic, G. e Jha, S. (2019). Predicting heart rate variations of deepfake videos using neural ode. Em Proceedings of the IEEE International Conference on Computer Vision Workshops, páginas 0–0.
- Fernandez, M. e Alani, H. (2018). Online misinformation: Challenges and future directions. Em Companion Proceedings of the The Web Conference 2018, páginas 595–602.
- Fernandez-Luque, L., Karlsen, R. e Melton, G. B. (2011). Healthtrust: trust-based retrieval of you tube’s diabetes channels. Em Proceedings of the 20th ACM international conference on Information and knowledge management, páginas 1917–1920.
- Ferrara, E. (2015). "manipulation and abuse on social media". ACM SIGWEB Newsletter.
- Ferreira, V. A. R. (2015). Uso de tecnologias da informação e comunicação para promover socialização: revisando o conceito de terceiros espaços.
- Feyerabend, P. K. (1989). Contra o método, volume 3. Francisco Alves Rio de Janeiro.
- Figueira, Á., Guimaraes, N. e Torgo, L. (2019). A brief overview on the strategies to fight back the spread of false information. Journal of Web Engineering, 18(4):319–352.
- Figueira, Á. e Oliveira, L. (2017). The current state of fake news: challenges and opportunities. Procedia Computer Science, 121:817–825.
- Finn, S., Metaxas, P. T. e Mustafaraj, E. (2015). Spread and skepticism: Metrics of propagation on twitter. Em Proceedings of the ACM Web Science Conference, páginas 1–2.
- Fiske, J. (2010). Introduction to communication studies. Routledge.
- Flanagin, A. J. e Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. New media & society, 9(2):319–342.

- Flintham, M., Karner, C., Bachour, K., Creswick, H., Gupta, N. e Moran, S. (2018). Falling for fake news: investigating the consumption of news via social media. Em Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, página 376. ACM.
- Fraga-Lamas, P. e Fernández-Caramés, T. M. (2020). Fake news, disinformation, and deepfakes: Leveraging distributed ledger technologies and blockchain to combat digital deception and counterfeit reality. IT Professional, 22(2):53–59.
- Friggeri, A., Adamic, L., Eckles, D. e Cheng, J. (2014). Rumor cascades. Em Eighth International AAAI Conference on Weblogs and Social Media.
- G1 (2021a). Aplicativo do Ministério da Saúde que recomenda 'tratamento precoce' para Covid-19 sai do ar. [Online; accessed 23. Feb. 2021].
- G1 (2021b). É #FAKE que a ingestão de alimentos alcalinos combate o novo coronavírus. [Online; accessed 8. Jun. 2021].
- G1 (2023a). 'Do sonho a pesadelo': casal leva golpe e descobre ingressos falsos na entrada para show do Coldplay. [Online; accessed 30. Mar. 2023].
- G1 (2023b). É #FAKE que governo Lula aumentou auxílio-reclusão para R\$ 1.754,18, valor maior que o salário mínimo. [Online; accessed 3. Apr. 2023].
- Galeotti, A. E. (2019). Believing fake news. Em Post-truth, philosophy and law, páginas 58–76. Routledge.
- Galitsky, B. (2015). Detecting rumor and disinformation by web mining. Em 2015 AAAI Spring Symposium Series.
- Galo, R. (2021). RegisGalo13 no Twitter. [Online; accessed 23. Feb. 2021].
- Gao, M., Xiao, Z., Karahalios, K. e Fu, W.-T. (2018). To label or not to label: The effect of stance and credibility labels on readers' selection and perception of news articles. Proceedings of the ACM on Human-Computer Interaction, 2(CSCW):1–16.
- Gaonkar, S., Itagi, S., Chalippatt, R., Gaonkar, A., Aswale, S. e Shetgaonkar, P. (2019). Detection of online fake news: A survey. Em 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), páginas 1–6. IEEE.
- Garcia-Recuero, A., Morawin, A. e Tyson, G. (2018). Trollslayer: Crowdsourcing and characterization of abusive birds in twitter. Em 2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), páginas 64–71. IEEE.
- Gazendam, H. e Liu, K. (2006). The evolution of organisational semiotics. Em INTERNATIONAL WORKSHOP ON ORGANISATIONAL SEMIOTICS, 6th Reading. Citeseer.
- Gedoor (2021). Conselho Federal de Farmácia - Brasil - Notícia: 30/04/2020 - Levantamento mostra como o medo da Covid-19 impactou venda de medicamentos. [Online; accessed 21. Mar. 2021].
- Ghaisani, A. P., Munajat, Q. e Handayani, P. W. (2017). Information credibility factors on information sharing activities in social media. Em 2017 Second International Conference on Informatics and Computing (ICIC), páginas 1–5. IEEE.



- Gibson, J. J. (1977). The theory of affordances. Hilldale, USA, 1(2):67–82.
- Giddens, A. (1986). The constitution of society: Outline of the theory of structuration, volume 349. Univ of California Press.
- Gilda, S. (2017). Evaluating machine learning algorithms for fake news detection. Em 2017 IEEE 15th Student Conference on Research and Development (SCOREd), páginas 110–115. IEEE.
- Glenski, M., Weninger, T. e Volkova, S. (2018). How humans versus bots react to deceptive and trusted news sources: A case study of active users. Em 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), páginas 654–661. IEEE.
- Goffman, E. (1974). Frame analysis: An essay on the organization of experience. Harvard University Press.
- Goindani, M. e Neville, J. (2020). Social reinforcement learning to combat fake news spread. Em Uncertainty in Artificial Intelligence, páginas 1006–1016. PMLR.
- Goolsby, R., Shanley, L. e Lovell, A. (2013). On cybersecurity, crowdsourcing, and social cyber-attack. Relatório técnico, OFFICE OF NAVAL RESEARCH ARLINGTON VA.
- Gragani, J. (2020). A psicologia por trás da corrida por papel higiênico em meio a “medo contagioso” do coronavírus - BBC News Brasil. [Online; accessed 15. Feb. 2022].
- Granik, M. e Mesyura, V. (2017). Fake news detection using naive bayes classifier. Em 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), páginas 900–903. IEEE.
- Gray, J. e Terp, S.-J. (2020). Misinformation: We’re Four Steps Behind Its Creators. [Online; accessed 30. Nov. 2020].
- Guacho, G. B., Abdali, S., Shah, N. e Papalexakis, E. E. (2018). Semi-supervised content-based detection of misinformation via tensor embeddings. Em 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), páginas 322–325. IEEE.
- Guimarães, M. S., Baranauskas, M. C. C. e Martins, E. (2008). Communication-based modelling and inspection in critical systems. Em ICEIS (5), páginas 215–220.
- Gunawan, F. E. e Suwandi, V. (2020). Identifying the most influencing characteristics of fake news. volume 11, páginas 93–101. ICIC International.
- Gupta, A., Lamba, H. e Kumaraguru, P. (2013). \$1.00 per rt# bostonmarathon# prayforboston: Analyzing fake content on twitter. Em 2013 APWG eCrime researchers summit, páginas 1–12. IEEE.
- Gupta, M. e Han, J. (2011). Heterogeneous network-based trust analysis: a survey. ACM SIGKDD Explorations Newsletter, 13(1):54–71.
- Habib, A., Asghar, M. Z., Khan, A., Habib, A. e Khan, A. (2019). False information detection in online content and its role in decision making: a systematic literature review. Social Network Analysis and Mining, 9(1):50.



- Hailun, D., Shuyi, W. e Guorui, Z. (2014). Multi-agent based research of centralized false information discerning model of micro-blog network. Em 2014 10th International Conference on Natural Computation (ICNC), páginas 1026–1032. IEEE.
- Halbach, P., Burbach, L., Nakayama, J., Plettenberg, N., Ziefle, M. e Calero Valdez, A. (2019). Would i lie to you? how users evaluate faked online content depending on its publication type. Em 2019 IEEE International Professional Communication Conference (ProComm), páginas 56–64.
- Halliday, M. A. K. (2014). Language as social semiotic. The Discourse Studies Reader. Amsterdam: John Benjamins, páginas 263–272.
- Hamborg, F., Meuschke, N. e Gipp, B. (2017). Matrix-based news aggregation: Exploring different news perspectives. Em 2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL), páginas 1–10.
- Hamborg, F., Zhukova, A. e Gipp, B. (2019). Automated identification of media bias by word choice and labeling in news articles. Em 2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL), páginas 196–205. IEEE.
- Hannak, A., Margolin, D., Keegan, B. e Weber, I. (2014). Get back! you don't know me like that: The social mediation of fact checking interventions in twitter conversations. Em ICWSM.
- Hao, K. (2021). How Facebook got addicted to spreading misinformation. MIT Technology Review.
- Harrison, S., Sengers, P. e Tatar, D. (2011). Making epistemological trouble: Third-paradigm hci as successor science. Interacting with computers, 23(5):385–392.
- Hassan, D. (2018). A text mining approach for evaluating event credibility on twitter. Em 2018 IEEE 27th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), páginas 171–174. IEEE.
- Hassan, N., Zhang, G., Arslan, F., Caraballo, J., Jimenez, D., Gawsane, S., Hasan, S., Joseph, M., Kulkarni, A., Nayak, A. K. et al. (2017). Claimbuster: the first-ever end-to-end fact-checking system. Proceedings of the VLDB Endowment, 10(12):1945–1948.
- Hernandez, J. C., Hernandez, C. J., Sierra, J. M. e Ribagorda, A. (2002). A first step towards automatic hoax detection. Em Proceedings. 36th Annual 2002 International Carnahan Conference on Security Technology, páginas 102–114. IEEE.
- Hinz, O. (2007). False information in internet auction communities. Em 2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07), páginas 178b–178b. IEEE.
- Hogan, K. (2010). The psychology of persuasion: how to persuade others to your way of thinking. Pelican Publishing.
- Hölig, S., Hasebrink, U. e Behre, J. (2020). Reuters Institute Digital News Report 2020: Ergebnisse für Deutschland, volume 50. DEU.
- Huang, D., Zhu, Y. e Mustafaraj, E. (2019). How dependable are "first impressions" to distinguish between real and fake newswebsites? Em Proceedings of the 30th ACM Conference on Hypertext and Social Media, páginas 201–210.

- Huh, M., Liu, A., Owens, A. e Efros, A. A. (2018). Fighting fake news: Image splice detection via learned self-consistency. Em Proceedings of the European Conference on Computer Vision (ECCV), páginas 101–117.
- Hui, P.-M., Shao, C., Flammini, A., Menczer, F. e Ciampaglia, G. L. (2018). The hoaxy misinformation and fact-checking diffusion network. Em Proceedings of the International AAAI Conference on Web and Social Media, volume 12.
- Hussain, M. N., Tokdemir, S., Agarwal, N. e Al-Khateeb, S. (2018). Analyzing disinformation and crowd manipulation tactics on youtube. Em 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), páginas 1092–1095. IEEE.
- Hutchins, E. (1995). Cognition in the Wild. MIT press.
- Hwang, T., Pearce, I. e Nanis, M. (2012). Socialbots: Voices from the fronts. interactions, 19(2):38–45.
- Indu, V. e Thampi, S. M. (2019). A nature-inspired approach based on forest fire model for modeling rumor propagation in social networks. Journal of Network and Computer Applications, 125:28–41.
- Introne, J., Iandoli, L., DeCook, J., Yildirim, I. G. e Elzeini, S. (2017). The collaborative construction and evolution of pseudo-knowledge in online conversations. Em Proceedings of the 8th International Conference on Social Media & Society, páginas 1–10.
- Ireland, S. (2018). Fake news alerts: Teaching news literacy skills in a meme world. The Reference Librarian, 59(3):122–128.
- Ishida, Y. e Kuraya, S. (2018). Fake news and its credibility evaluation by dynamic relational networks: A bottom up approach. Procedia computer science, 126:2228–2237.
- Jameel, T., Ali, R. e Malik, K. A. (2019). Social media as an opinion formulator: A study on implications and recent developments. Em 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), páginas 1–6. IEEE.
- Jang, S. M., Geng, T., Li, J.-Y. Q., Xia, R., Huang, C.-T., Kim, H. e Tang, J. (2018). A computational approach for examining the roots and spreading patterns of fake news: Evolution tree analysis. Computers in Human Behavior, 84:103–113.
- Janicka, M., Pszona, M. e Wawer, A. (2019). Cross-domain failures of fake news detection. Computación y Sistemas, 23(3).
- Jansen van Vuuren, A., Jansen van Vuuren, J. e Venter, S. (2012). The susceptibility of the south african media to be used as a tool for information warfare. Em Proceedings of the 11th European Conference on Information Warfare and Security, páginas 127–134. Academic Conferences Limited.
- Java, S., Basheer, F. L., Riaz, S., Kaur, M. J. e Mushtaq, A. (2019). Detection of online manipulation to prevent users victimization. Em 2019 Amity International Conference on Artificial Intelligence (AICAI), páginas 593–599. IEEE.
- Jeong, Y.-S. e Park, J. H. (2019). Learning algorithms in ai system and services. Journal of Information Processing Systems, 15(5):1029–1035.

- Jiang, S., Robertson, R. E. e Wilson, C. (2019). Bias misperceived: The role of partisanship and misinformation in youtube comment moderation. Em Proceedings of the International AAAI Conference on Web and Social Media, volume 13, páginas 278–289.
- Jiang, S. e Wilson, C. (2018). Linguistic signals under misinformation and fact-checking: Evidence from user comments on social media. Proceedings of the ACM on Human-Computer Interaction, 2(CSCW):1–23.
- Jin, Z., Cao, J., Guo, H., Zhang, Y. e Luo, J. (2017). Multimodal fusion with recurrent neural networks for rumor detection on microblogs. Em Proceedings of the 25th ACM international conference on Multimedia, páginas 795–816.
- Jindal, S., Sood, R., Singh, R., Vatsa, M. e Chakraborty, T. (2020). Newsbag: A multimodal benchmark dataset for fake news detection. Em CEUR Workshop Proc., volume 2560, páginas 138–145.
- Joel, M. (2021). Coronavirus doesn't have own pH level, alkaline food won't 'beat' it. [Online; accessed 8. Jun. 2021].
- Johnson, P. J. (2020). Eating alkaline food cannot prevent or cure COVID-19. [Online; accessed 8. Jun. 2021].
- Kaliyar, R. K. e Singh, N. (2019). Misinformation detection on online social media-a survey. Em 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), páginas 1–6. IEEE.
- Kapusta, J. e Obonya, J. (2020). Improvement of misleading and fake news classification for fleective languages by morphological group analysis. Em Informatics, volume 7, página 4. Multidisciplinary Digital Publishing Institute.
- Karadzhov, G., Gencheva, P., Nakov, P. e Koychev, I. (2018). We built a fake news & click-bait filter: what happened next will blow your mind! arXiv preprint arXiv:1803.03786.
- Karduni, A., Wesslen, R., Santhanam, S., Cho, I., Volkova, S., Arendt, D., Shaikh, S. e Dou, W. (2018). Can you verifi this? studying uncertainty and decision-making about misinformation using visual analytics. Em Twelfth international AAAI conference on web and social media.
- Kareem, I. e Awan, S. M. (2019). Pakistani media fake news classification using machine learning classifiers. Em 2019 International Conference on Innovative Computing (ICIC), páginas 1–6. IEEE.
- Karimi, H. F., Masruroh, S. U., Mintarsih, F. et al. (2018). The influence of iteration calculation manipulation on social network analysis toward twitter's users against hoax in indonesia with single cluster multi-node method using apache hadoop hortonworkstm distribution. Em 2018 6th International Conference on Cyber and IT Service Management (CITSM), páginas 1–6. IEEE.
- Karlova, N. A. e Fisher, K. E. (2013). A social diffusion model of misinformation and disinformation for understanding human information behaviour. Information Research, 18(1).
- Kasra, M., Shen, C. e O'Brien, J. F. (2018). Seeing is believing: how people fail to identify fake images on the web. Em Extended abstracts of the 2018 CHI conference on human factors in computing systems, páginas 1–6.

- Katsaros, D., Stavropoulos, G. e Papakostas, D. (2019). Which machine learning paradigm for fake news detection? Em 2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI), páginas 383–387. IEEE.
- Kesumawati, A. e Thalib, A. K. (2018). Hoax classification with term frequency-inverse document frequency using non-linear svm and naïve bayes. International Journal of Advances in Soft Computing & Its Applications, 10(3).
- Khaled, S., El-Tazi, N. e Mokhtar, H. M. (2018). Detecting fake accounts on social media. Em 2018 IEEE International Conference on Big Data (Big Data), páginas 3672–3681. IEEE.
- Khan, S. A., Alkawaz, M. H. e Zangana, H. M. (2019). The use and abuse of social media for spreading fake news. Em 2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS), páginas 145–148. IEEE.
- Khandelwal, S. e Kumar, D. (2020). Computational fact validation from knowledge graph using structured and unstructured information. Em Proceedings of the 7th ACM IKDD CoDS and 25th COMAD, páginas 204–208.
- Khodabakhsh, A., Busch, C. e Ramachandra, R. (2018). A taxonomy of audiovisual fake multimedia content creation technology. Em 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), páginas 372–377. IEEE.
- Kim, J., Tabibian, B., Oh, A., Schölkopf, B. e Gomez-Rodriguez, M. (2018). Leveraging the crowd to detect and reduce the spread of fake news and misinformation. Em Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, páginas 324–332.
- Kokar, M. M. e Endsley, M. R. (2012). Situation awareness and cognitive modeling. IEEE Intelligent Systems, 27(3):91–96.
- Kolkman, M. (1993). Problem articulation methodology.
- Konstantinou, L., Caraban, A. e Karapanos, E. (2019). Combating misinformation through nudging. Em IFIP Conference on Human-Computer Interaction, páginas 630–634. Springer.
- Kopev, D., Ali, A., Koychev, I. e Nakov, P. (2019). Detecting deception in political debates using acoustic and textual features. Em 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), páginas 652–659. IEEE.
- Korshunov, P. e Marcel, S. (2018). Speaker inconsistency detection in tampered video. Em 2018 26th European Signal Processing Conference (EUSIPCO), páginas 2375–2379. IEEE.
- Kotteti, C. M. M., Dong, X., Li, N. e Qian, L. (2018). Fake news detection enhancement with data imputation. Em 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), páginas 187–192. IEEE.
- Kovic, M. e Laissue, N. (2016). Consuming rationally: How marketing is exploiting our cognitive biases, and what we can do about it.
- Krishnamurthy, V. e Hamdi, M. (2013). Mis-information removal in social networks: Constrained estimation on dynamic directed acyclic graphs. IEEE Journal of Selected Topics in Signal Processing, 7(2):333–346.

- Kshetri, N. e Voas, J. (2017). The economics of “fake news”. IT Professional, 19(6):8–12.
- Kugler, L. (2018). Getting hooked on tech. Commun. ACM, 61(6):18–19.
- Kumar, K. K. e Geethakumari, G. (2013). Information diffusion model for spread of misinformation in online social networks. Em 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), páginas 1172–1177. IEEE.
- Kumar, K. K. e Geethakumari, G. (2014). Detecting misinformation in online social networks using cognitive psychology. Human-centric Computing and Information Sciences, 4(1):1–22.
- Kumar, K. K., Srivastava, A. e Geethakumari, G. (2016a). A psychometric analysis of information propagation in online social networks using latent trait theory. Computing, 98(6):583–607.
- Kumar, S., Huang, B., Cox, R. A. V. e Carley, K. M. (2020). An anatomical comparison of fake-news and trusted-news sharing pattern on twitter. Computational and Mathematical Organization Theory, páginas 1–25.
- Kumar, S., Spezzano, F. e Subrahmanian, V. (2014). Accurately detecting trolls in slashdot zoo via decluttering. Em 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014), páginas 188–195. IEEE.
- Kumar, S., West, R. e Leskovec, J. (2016b). Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. Em Proceedings of the 25th international conference on World Wide Web, páginas 591–602.
- Lal, Y. K., Khattar, D., Kumar, V., Mishra, A. e Varma, V. (2018). Check it out: Politics and neural networks. Em CLEF (Working Notes).
- Larson, K. G., Kaleda, K. e Fenstermacher, L. (2019). Applying cognitive psychology principles to the (dis) information environment: an examination of discourse comprehension, memory, and fusion of news articles. Em Signal Processing, Sensor/Information Fusion, and Target Recognition XXVIII, volume 11018, página 110180S. International Society for Optics and Photonics.
- Lazar, J., Feng, J. H. e Hochheiser, H. (2017). Research methods in human-computer interaction. Morgan Kaufmann.
- Leblay, J., Chen, W. e Lynden, S. (2017). Exploring the veracity of online claims with backdrop. Em Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM '17, página 2491–2494, New York, NY, USA. Association for Computing Machinery.
- Lee, J. e Choi, Y. (2018). Informed public against false rumor in the social media era: Focusing on social media dependency. Telematics and Informatics, 35(5):1071–1081.
- Lee, S. (2019). Detection of political manipulation through unsupervised learning. TIIS, 13(4):1825–1844.
- Leite, C., Gondim, J. J., Barreto, P. S. e Alchieri, E. A. (2019). Waste flooding: A phishing retaliation tool. Em 2019 IEEE 18th International Symposium on Network Computing and Applications (NCA), páginas 1–8. IEEE.



- Lemieux, V. e Smith, T. D. (2018). Leveraging archival theory to develop a taxonomy of online disinformation. Em 2018 IEEE International Conference on Big Data (Big Data), páginas 4420–4426. IEEE.
- Li, B. e Chong, A. Y.-L. (2019). What influences the dissemination of online rumor messages: Message features and topic-congruence. Em ICIS.
- Li, D. e Santos Jr, E. (2012). Argument formation in the reasoning process: toward a generic model of deception detection. Em Proceedings of the Workshop on Computational Approaches to Deception Detection, páginas 63–71. Association for Computational Linguistics.
- Liao, J. e Yang, J. (2012). The influence of electronic word-of-mouth authenticity on customers' behavior. Em 2012 Second International Conference on Business Computing and Global Informatization, páginas 427–430. IEEE.
- Lin, T., Etzioni, O. e Fogarty, J. (2009). Identifying interesting assertions from the web. Em Proceedings of the 18th ACM conference on Information and knowledge management, páginas 1787–1790.
- Linstrom, M. e Marais, W. (2012). Qualitative news frame analysis: A methodology. Communitas, 17:21–38.
- Liu, K. (2000). Semiotics in information systems engineering. Cambridge University Press.
- Liu, X., Nourbakhsh, A., Li, Q., Fang, R. e Shah, S. (2015). Real-time rumor debunking on twitter. Em Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, páginas 1867–1870.
- Loui, R. e Hope, W. (2017). Information warfare amplified by cyberwarfare and hacking the national knowledge infrastructure. Em 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), páginas 280–283. IEEE.
- Lozano, M. G., Brynielsson, J., Franke, U., Rosell, M., Tjörnhallmar, E., Varga, S. e Vlassov, V. (2020). Veracity assessment of online data. Decision Support Systems, 129:113132.
- Luciano, F. (1996). Brave.Net.World: the Internet as a disinformation superhighway? The Electronic Library, 14(6):509–514.
- Lupa, A. (2020). Agência Lupa. [Online; accessed 1. Jun. 2022].
- Lycan, W. G. (1988). Judgement and justification. CUP Archive.
- Mack, G. A., Eick, S. G. e Clark, M. A. (2007). Models of trust and disinformation in the open press from model-driven linguistic pattern analysis. Em 2007 IEEE Aerospace Conference, páginas 1–12. IEEE.
- Madhusudan, T. (2003). On a text-processing approach to facilitating autonomous deception detection. Em 36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the, páginas 10–pp. IEEE.



- Maigrot, C., Claveau, V. e Kijak, E. (2018). Fusion-based multimodal detection of hoaxes in social networks. Em 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI), páginas 222–229. IEEE.
- Maigrot, C., Claveau, V., Kijak, E. e Sicre, R. (2016). Mediaeval 2016: A multimodal system for the verifying multimedia use task. Em MediaEval 2016: "Verifying Multimedia Use" task.
- Manzoor, Syed Ishfaq; Singla, J. e Nikita (2019). Fake news detection using machine learning approaches: A systematic review. Em 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), páginas 230–234. IEEE.
- Matthes, J. e Kohring, M. (2008). The content analysis of media frames: Toward improving reliability and validity. Journal of communication, 58(2):258–279.
- McCabe, V. (1982). Invariants and affordances: An analysis of species-typical information. Ethology and Sociobiology, 3(2):79–92.
- McCarthy, L. (2018). 'let's fake news'. Em Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction, páginas 622–625.
- McCombs, M. E. e Shaw, D. L. (1972). The agenda-setting function of mass media. Public opinion quarterly, 36(2):176–187.
- McLuhan, M. (1974). Os meios de comunicação: como extensões do homem. Editora Cultrix.
- Meel, P. e Vishwakarma, D. K. (2019). Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. Expert Systems with Applications, página 112986.
- Melo, P., Messias, J., Resende, G., Garimella, K., Almeida, J. e Benevenuto, F. (2019). Whatsapp monitor: A fact-checking system for whatsapp. Em Proceedings of the International AAAI Conference on Web and Social Media, volume 13, páginas 676–677.
- Mendonça, R. F. e Simões, P. G. (2012). Enquadramento: diferentes operacionalizações analíticas de um conceito. Revista Brasileira de Ciências Sociais, 27:187–201.
- Mendoza, M., Poblete, B. e Castillo, C. (2010). Twitter under crisis: Can we trust what we rt? Em Proceedings of the first workshop on social media analytics, páginas 71–79.
- Metaxas, P. T., Finn, S. e Mustafaraj, E. (2015). Using twittertrails. com to investigate rumor propagation. Em Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing, páginas 69–72.
- Mikko (2021). News from the Lab Archive : January 2004 to September 2015. [Online; accessed 30. Aug. 2021].
- Miletskiy, V. P., Cherezov, D. N. e Stroetskaya, E. V. (2019). Transformations of professional political communications in the digital society (by the example of the fake news communication strategy). Em 2019 Communication Strategies in Digital Society Workshop (ComSDS), páginas 121–124. IEEE.
- Moreno, J. e Bressan, G. (2019). Factck.br: a new dataset to study fake news. Em Proceedings of the 25th Brazillian Symposium on Multimedia and the Web, páginas 525–527.

- Murimi, R. (2018). Online social networks for meaningful social reform. Em 2018 World Engineering Education Forum-Global Engineering Deans Council (WEEF-GEDC), páginas 1–6. IEEE.
- Murungi, D., Purao, S. e Yates, D. (2018). Beyond facts: A new spin on fake news in the age of social media. Em Twenty-fourth Americas Conference on Information Systems.
- Narahari, Y. et al. (2012). Influence limitation in multi-campaign social networks: A shapley value based approach. Em 2012 IEEE International Conference on Automation Science and Engineering (CASE), páginas 735–740. IEEE.
- Nekmat, E. (2020). Nudge effect of fact-check alerts: Source influence and media skepticism on sharing of news misinformation in social media. Social Media+ Society, 6(1):2056305119897322.
- Nekovee, M., Moreno, Y., Bianconi, G. e Marsili, M. (2007). Theory of rumour spreading in complex social networks. Physica A: Statistical Mechanics and its Applications, 374(1):457–470.
- Nestoras, A. (2018). Political warfare: Competition in the cyber era. Em 2018 IEEE International Conference on Big Data (Big Data), páginas 4427–4436. IEEE.
- Neumann, P. G. (1996). Disinformation theory. Communications of the ACM, 39(10).
- Newman, N., Fletcher, R., Kalogeropoulos, A. e Nielsen, R. (2019). Reuters institute digital news report 2019. Relatório técnico.
- Nicol, D. M. (2005). Modeling and simulation in security evaluation. IEEE security & privacy, 3(5):71–74.
- Nikoukhah, T., von Gioi, R. G., Colom, M. e Morel, J.-M. (2018). Automatic jpeg grid detection with controlled false alarms, and its image forensic applications. Em 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), páginas 378–383. IEEE.
- Nilforoshan, H. e Shah, N. (2019). Slicendice: Mining suspicious multi-attribute entity groups with multi-view graphs. Em 2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA), páginas 351–363.
- Noth, W. (1990). Handbook of semiotics. Indiana University Press.
- Nuruddin Hidayat, D., Nurhalimah, Defianty, M., Kultsum, U., Zulkifli e Sufyan, A. (2020). Logical fallacies in social media: A discourse analysis in political debate. Em 2020 8th International Conference on Cyber and IT Service Management (CITSM), páginas 1–5.
- Nussbaumer, A., Verbert, K., Hillemann, E.-C., Bedek, M. A. e Albert, D. (2016). A framework for cognitive bias detection and feedback in a visual analytics environment. Em 2016 European Intelligence and Security Informatics Conference (EISIC), páginas 148–151. IEEE.
- Oehmichen, A., Hua, K., López, J. A. D., Molina-Solana, M., Gómez-Romero, J. e Guo, Y.-k. (2019). Not all lies are equal. a study into the engineering of political misinformation in the 2016 us presidential election. IEEE Access, 7:126305–126314.

- Oh, O., Kwon, K. H. e Rao, H. R. (2010). An exploration of social media in extreme events: Rumor theory and twitter during the haiti earthquake 2010. Em Icis, volume 231, páginas 7332–7336.
- Olhar Digital (2021). App da Saúde se contradiz sobre o uso de máscaras - Olhar Digital. Olhar Digital.
- Ookita, Y. e Fujita, S. (2017). Effective suppression of false rumors in social network service. Em 2017 Fifth International Symposium on Computing and Networking (CANDAR), páginas 243–248. IEEE.
- Orlikowski, W. J. (2007). Sociomaterial practices: Exploring technology at work. Organization studies, 28(9):1435–1448.
- Oulasvirta, A. e Hornbæk, K. (2016). Hci research as problem-solving. Em Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, páginas 4956–4967.
- Özgöbek, Ö. e Gulla, J. A. (2017). Towards an understanding of fake news. Em CEUR Workshop Proceedings - Norwegian Big Data Symposium 2017.
- Pacheco, A. L. P., Brena, D. V. G., Monterrosa, N. T. e Sandoval, C. A. M. (2019). Yakuin: personal content verifier in digital messaging. Em Proceedings of the IX Latin American Conference on Human Computer Interaction, páginas 1–6.
- Pal, A. e Chua, A. Y. (2016). Reviewing the landscape of research on the threats to the quality of user-generated content. Proceedings of the Association for Information Science and Technology, 53(1):1–9.
- Pal, A., Chua, A. Y. e Goh, D. H.-L. (2018). Salient beliefs about sharing rumor denials on the internet. Em Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication, páginas 1–7.
- Pal, A., Chua, A. Y. e Goh, D. H.-L. (2019). Debunking rumors on social media: The use of denials. Computers in Human Behavior, 96:110–122.
- Pal, A. e Loke, C. (2019). Communicating fact to combat fake: Analysis of fact-checking websites. Em Proceedings of the 2019 International Conference on Information Technology and Computer Communications, páginas 66–73.
- Pancher, S. (2021). Samuel Pancher no Twitter. [Online; accessed 23. Feb. 2021].
- Pang, N. e Ng, J. (2017). Misinformation in a riot: a two-step flow view. Online Information Review.
- Parikh, S. B. e Atrey, P. K. (2018). Media-rich fake news detection: A survey. Em 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), páginas 436–441. IEEE.
- Patankar, A., Bose, J. e Khanna, H. (2019). A bias aware news recommendation system. Em 2019 IEEE 13th International Conference on Semantic Computing (ICSC), páginas 232–238. IEEE.

- Patel, I., Nguyen, H., Belyi, E., Getahun, Y., Abdulkareem, S., Giabbanelli, P. J. e Mago, V. (2017). Modeling information spread in polarized communities: Transitioning from legacy media to a facebook world. Em SoutheastCon 2017, páginas 1–8. IEEE.
- Pearce, K. E. e Malhotra, P. (2022). Inaccuracies and izzat: Channel affordances for the consideration of face in misinformation correction. Journal of Computer-Mediated Communication, 27(2):zmac004.
- Peirce, C. S. e Buchler, J. (1902). Logic as semiotic: The theory of signs.
- Pereira, R. e Baranauskas, M. C. C. (2015). A value-oriented and culturally informed approach to the design of interactive systems. International Journal of Human-Computer Studies, 80:66–82.
- Petersen, K., Vakkalanka, S. e Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. Information and Software Technology, 64:1–18.
- Piccolo, L. S., Puska, A., Pereira, R. e Farrell, T. (2020). Pathway to a human-values based approach to tackle misinformation online. Em International Conference on Human-Computer Interaction, páginas 510–522. Springer.
- Pierri, F. e Ceri, S. (2019). False news on social media: a data-driven survey. ACM Sigmod Record, 48(2):18–27.
- Pinheiro, A., Cappelli, C., Aguiar, A. e Maciel, C. (2018). Measuring the accuracy and learnability of tools in the struggle against misinformation in social media applications. Em 2018 13th Iberian Conference on Information Systems and Technologies (CISTI), páginas 1–4. IEEE.
- Pinto, M. R., de Lima, Y. O., Barbosa, C. E. e de Souza, J. M. (2019). Towards fact-checking through crowdsourcing. Em 2019 IEEE 23rd International Conference on Computer Supported Cooperative Work in Design (CSCWD), páginas 494–499. IEEE.
- Poddar, K., Umadevi, K. et al. (2019). Comparison of various machine learning models for accurate detection of fake news. Em 2019 Innovations in Power and Advanced Computing Technologies (i-PACT), volume 1, páginas 1–5. IEEE.
- Pohl, R. e Pohl, R. F. (2004). Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory. Psychology Press.
- Pointer (2021). CoronaVirusFacts Alliance. [Online; accessed 8. Jun. 2021].
- Pollalis, C., Grevet, C., Westendorf, L., Finn, S., Shaer, O. e Metaxas, P. (2018). Classroom activity for critical analysis of news propagation online. Em Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, páginas 1–10.
- Popat, K., Mukherjee, S., Strötgen, J. e Weikum, G. (2017). Where the truth lies: Explaining the credibility of emerging claims on the web and social media. Em Proceedings of the 26th International Conference on World Wide Web Companion, páginas 1003–1012.
- Porto, M. (2004). Enquadramentos da mídia e política.
- Pourghomi, P., Safieddine, F., Masri, W. e Dordevic, M. (2017). How to stop spread of misinformation on social media: Facebook plans vs. right-click authenticate approach. Em 2017 International Conference on Engineering MIS (ICEMIS), páginas 1–8.

- Povo, O. (2020). É falsa informação de que o coronavírus é imune a organismos com PH maior que 5,5. Coronavirus.
- Prasetijo, A. B., Isnanto, R. R., Eridani, D., Soetrisno, Y. A. A., Arfan, M. e Sofwan, A. (2017). Hoax detection system on indonesian news sites based on text classification using svm and sgd. Em 2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), páginas 45–49. IEEE.
- Pratiwi, I. Y. R., Asmara, R. A. e Rahutomo, F. (2017). Study of hoax news detection using naïve bayes classifier in indonesian language. Em 2017 11th International Conference on Information & Communication Technology and System (ICTS), páginas 73–78. IEEE.
- Přibáň, P., Hercig, T. e Steinberger, J. (2019). Machine learning approach to fact-checking in west slavic languages. Em Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019), páginas 973–979.
- Purnomo, M. H., Sumpeno, S., Setiawan, E. I. e Purwitasari, D. (2017). Keynote speaker ii: biomedical engineering research in the social network analysis era: stance classification for analysis of hoax medical news in social media.
- Puska, A., Baroni, L. A., Canal, M. C., Piccolo, L. S. e Pereira, R. (2020). Whatsapp and false information: a value-oriented evaluation. Em Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems, páginas 1–10. ACM.
- Puska, A. e Pereira, R. (2023). Exploring digital misinformation as a sociotechnical phenomenon: Insights from a small-scale study. Em IHC '23: 22nd Brazilian Symposium on Human Factors in Computing Systems, página 12. ACM.
- Puska, A., Piccolo, L. e Pereira, R. (2019). Dismiss false information: A value matter. Human Computer Interaction and Emerging Technologies: Adjunct Proceedings from, página 75.
- Qazvinian, V., Rosengren, E., Radev, D. R. e Mei, Q. (2011). Rumor has it: Identifying misinformation in microblogs. Em Proceedings of the conference on empirical methods in natural language processing, páginas 1589–1599. Association for Computational Linguistics.
- Qiming, S., Wei, Z., Rongli, H. e Yixuan, L. (2021). Strategy on the service mode of university library to help users cope with the false information under the we media environment. Em 2021 9th International Conference on Information and Education Technology (ICIET), páginas 369–373. IEEE.
- Qin, Y., Dominik, W. e Tang, C. (2018). Predicting future rumours. Chinese Journal of Electronics, 27(3):514–520.
- R. Walker, C., Terp, S.-J., C. Breuer, P. e L. Crooks, PhD, C. (2019). Misinfosec, página 1026–1032. Association for Computing Machinery, New York, NY, USA.
- Rahman, M. M., Tajrin, J., Hasnat, A., Uzzaman, N. e Rahaman, G. A. (2019). Smifd: Novel social media image forgery detection database. Em 2019 22nd International Conference on Computer and Information Technology (ICCIT), páginas 1–6. IEEE.
- Rajdev, M. e Lee, K. (2015). Fake and spam messages: Detecting misinformation during natural disasters on social media. Em 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), volume 1, páginas 17–20. IEEE.



- Raman, G., Peng, J. C.-H. e Rahwan, T. (2019). Manipulating residents' behavior to attack the urban power distribution system. IEEE Transactions on Industrial Informatics, 15(10):5575–5587.
- Rana, T., Meel, P. e Roohani (2019). Rumor propagation: A state-of-the-art survey of current challenges and opportunities. Em 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), páginas 64–69. IEEE.
- Raskin, A. (2011). Your life experiences, brought to you by budweiser. IEEE Security & Privacy, 9(2):86–88.
- Rath, B., Gao, W., Ma, J. e Srivastava, J. (2017). From retweet to believability: Utilizing trust to identify rumor spreaders on twitter. Em Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017, páginas 179–186.
- Rath, B., Gao, W. e Srivastava, J. (2019). Evaluating vulnerability to fake news in social networks: A community health assessment model. Em 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), páginas 432–435. IEEE.
- Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Patil, S., Flammini, A. e Menczer, F. (2011). Truthy: mapping the spread of astroturf in microblog streams. Em Proceedings of the 20th international conference companion on World wide web, páginas 249–252.
- Raynal, F. e Gaspard, F. (2010). Small treatise about e-manipulation for honest people. Journal in computer virology, 6(2):143–159.
- Recuero, R., Soares, F. B. e Gruzd, A. (2020). Hyperpartisanship, disinformation and political conversations on twitter: The brazilian presidential election of 2018. Em Proceedings of the international AAAI conference on Web and social media, volume 14, páginas 569–578.
- Regner, A. C. K. P. (1996). Feyerabend e o pluralismo metodológico. Episteme (Porto Alegre): filosofia e história das ciências em revista. Porto Alegre. Vol. 1, n. 2 (1996), p. 61-78.
- Rehm, G. (2017). An infrastructure for empowering internet users to handle fake news and other online media phenomena. Em International Conference of the German Society for Computational Linguistics and Language Technology, páginas 216–231. Springer.
- Resende, G., Messias, J., Silva, M., Almeida, J., Vasconcelos, M. e Benevenuto, F. (2018). A system for monitoring public political groups in whatsapp. Em Proceedings of the 24th Brazilian Symposium on Multimedia and the Web, páginas 387–390.
- Resnik, D. B. (1998). Medical misinformation on the web: mitigation or control? ACM SIGCAS Computers and Society, 28(1):35–37.
- Reuter, C., Kaufhold, M.-A. e Steinfort, R. (2017). Rumors, fake news and social bots in conflicts and emergencies: Towards a model for believability in social media. Em ISCRAM.
- Rieh, S. Y. (2002). Judgment of information quality and cognitive authority in the web. Journal of the American society for information science and technology, 53(2):145–161.
- Rivner, U. (2011). Anatomy of an Attack « Speaking of Security – The RSA Blog and Podcast. [Online; accessed 18. Nov. 2021].



- Robson, C. (2017). Small-scale evaluation: Principles and practice. Sage.
- Rosnow, R. L. (1991). Inside rumor: A personal journey. American psychologist, 46(5):484.
- Ross, B., Pilz, L., Cabrera, B., Brachten, F., Neubaum, G. e Stieglitz, S. (2019). Are social bots a real threat? an agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks. European Journal of Information Systems, 28(4):394–412.
- Rowe, N. C. (2006). A taxonomy of deception in cyberspace.
- Rowe, N. C. e Rothstein, H. S. (2004). Two taxonomies of deception for attacks on information systems. Journal of Information Warfare, 3(2):27–39.
- RSA (2021). About RSA | Cybersecurity and Digital Risk Management. [Online; accessed 18. Nov. 2021].
- Ruan, J., Deng, J., Amariuca, G. A. e Wei, S. (2015). Efficient link cuts in online social networks. Em 2015 IEEE Global Communications Conference (GLOBECOM), páginas 1–6. IEEE.
- Rubin, V. L., Chen, Y. e Conroy, N. J. (2015). Deception detection for news: three types of fakes. Proceedings of the Association for Information Science and Technology, 52(1):1–4.
- Rubin, V. L. e Conroy, N. J. (2012). The art of creating an informative data collection for automated deception detection: A corpus of truths and lies. Proceedings of the American Society for Information Science and Technology, 49(1):1–11.
- Saez-Trumper, D., Castillo, C. e Lalmas, M. (2013). Social media news communities: gatekeeping, coverage, and statement bias. Em Proceedings of the 22nd ACM international conference on Information & Knowledge Management, páginas 1679–1684.
- Safieddine, F., Dordevic, M. e Pourghomi, P. (2017). Spread of misinformation online: Simulation impact of social media newsgroups. Em 2017 Computing Conference, páginas 899–906. IEEE.
- Saha, T., Ramesh Jayashree, S., Saha, S. e Bhattacharyya, P. (2020). Bert-caps: A transformer-based capsule network for tweet act classification. IEEE Transactions on Computational Social Systems, 7(5):1168–1179.
- Samonte, M. J. C. (2018). Polarity analysis of editorial articles towards fake news detection. Em Proceedings of the 2018 International Conference on Internet and e-Business, páginas 108–112.
- Sandifer, A. V., Wilson, C. e Olmsted, A. (2017). Detection of fake online hotel reviews. Em 2017 12th International Conference for Internet Technology and Secured Transactions (ICITST), páginas 501–502. IEEE.
- Santia, G. C., Mujib, M. I. e Williams, J. R. (2019). Detecting social bots on facebook in an information veracity context. Em Proceedings of the International AAAI Conference on Web and Social Media, volume 13, páginas 463–472.
- Santos-Pinto, C. D. B., Miranda, E. S. e Osorio-de Castro, C. G. S. (2021). O “kit-covid” e o programa farmácia popular do brasil. Cadernos de Saúde Pública, 37:e00348020.

- Santoso, I., Yohansen, I., Warnars, H. L. H. S., Hashimoto, K. et al. (2017). Early investigation of proposed hoax detection for decreasing hoax in social media. Em 2017 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom), páginas 175–179. IEEE.
- Sapo (2021). Polígrafo. [Online; accessed 8. Jun. 2021].
- Saquete, E., Tomas, D., Moreda, P., Martinez-Barco, P. e Palomar, M. (2020). Fighting post-truth using natural language processing: A review and open challenges. Expert Systems with Applications, 141:112943.
- Sbisà, M. e Fabbri, P. (1980). Models (?) for a pragmatic analysis. Journal of Pragmatics, 4(4):301–319.
- Schatz, B. R. e Hardin, J. B. (1994). Ncsa mosaic and the world wide web: global hypermedia protocols for the internet. Science, 265(5174):895–901.
- Seo, H., Xiong, A. e Lee, D. (2019). Trust it or not: Effects of machine-learning warnings in helping individuals mitigate misinformation. Em Proceedings of the 10th ACM Conference on Web Science, páginas 265–274.
- Sethi, R. e Rangaraju, R. (2018). Extinguishing the backfire effect: using emotions in online social collaborative argumentation for fact checking. Em 2018 IEEE International Conference on Web Services (ICWS), páginas 363–366. IEEE.
- Sethi, R. J. (2017). Crowdsourcing the verification of fake news and alternative facts. Em Proceedings of the 28th ACM Conference on Hypertext and Social Media, páginas 315–316.
- Shahid, W., Li, Y., Staples, D., Amin, G., Hakak, S. e Ghorbani, A. (2022). Are you a cyborg, bot or human?—a survey on detecting fake news spreaders. IEEE Access, 10:27069–27083.
- Shao, C., Hui, P.-M., Cui, P., Jiang, X. e Peng, Y. (2018). Tracking and characterizing the competition of fact checking and misinformation: Case studies. IEEE Access, 6:75327–75341.
- Sharma, K., Qian, F., Jiang, H., Ruchansky, N., Zhang, M. e Liu, Y. (2019). Combating fake news: A survey on identification and mitigation techniques. ACM Transactions on Intelligent Systems and Technology (TIST), 10(3):1–42.
- Sharma, S. e Sharma, D. K. (2019). Fake news detection: A long way to go. Em 2019 4th International Conference on Information Systems and Computer Networks (ISCON), páginas 816–821. IEEE.
- Shelke, S. e Attar, V. (2019). Source detection of rumor in social network-a review. Online Social Networks and Media, 9:30–42.
- Shen, Q., Wang, L.-Y., Cockerham, D. e Lin, L. (2019). Media multitasking, attention and news evaluation. Em 2019 Eighth International Conference on Educational Innovation through Technology (EITT), páginas 184–189. IEEE.
- Shin, J., Jian, L., Driscoll, K. e Bar, F. (2018). The diffusion of misinformation on social media: Temporal pattern, message, and source. Computers in Human Behavior, 83:278–287.

- Shrestha, A. e Spezzano, F. (2019). Online misinformation: from the deceiver to the victim. Em Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, páginas 847–850.
- Shu, K., Sliva, A., Wang, S., Tang, J. e Liu, H. (2017). Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter, 19(1):22–36.
- Shu, K., Wang, S. e Liu, H. (2018). Understanding user profiles on social media for fake news detection. Em 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), páginas 430–435. IEEE.
- Shu, K., Wang, S. e Liu, H. (2019). Beyond news contents: The role of social context for fake news detection. Em Proceedings of the twelfth ACM international conference on web search and data mining, páginas 312–320.
- Silva, R. M., Santos, R. L., Almeida, T. A. e Pardo, T. A. (2020). Towards automatically filtering fake news in portuguese. Expert Systems with Applications, 146:113199.
- Sirajudeen, S. M., Azmi, N. F. A. e Abubakar, A. I. (2017). Online fake news detection algorithm. Journal of Theoretical & Applied Information Technology, 95(17).
- Søe, S. O. (2018). Algorithmic detection of misinformation and disinformation: Gricean perspectives. Journal of Documentation.
- Soler, U. e Busilo, M. (2019). Education of society as a tool to counteract disinformation in implementing new technologies. on the example of 5g mobile telecommunications network and warsaw sewage system. Em 2019 Applications of Electromagnetics in Modern Engineering and Medicine (PTZE), páginas 210–214. IEEE.
- Spiegelberg, H. (1994). The Phenomenological Movement. Springer, Dordrecht, The Netherlands.
- Srivastava, A., Kannan, R., Chelmiss, C. e Prasanna, V. K. (2018). Factcheck: Keeping activation of fake news at check. Em Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems, páginas 2079–2081. International Foundation for Autonomous Agents and Multiagent Systems.
- Stamper, R. (1973). Information in business and administrative systems. John Wiley & Sons, Inc.
- Stamper, R. (1991). The semiotic framework for informatio systems research. Information systems research: Contemporary approaches and emergent traditions, páginas 515–528.
- Stamper, R. (1993a). A semiotic theory of information and information systems. Em Invited papers for the ICL/University of Newcastle Seminar on Information.
- Stamper, R. (1994). Social Norms in Requirements Analysis - an outline of MEASUR, páginas 107–139. Computers and people series. Academic Press, United States.
- Stamper, R. e Liu, K. (1994). Organisational dynamics, social norms and information systems. Em HICSS (4), páginas 645–654.

- Stamper, R., Liu, K., Hafkamp, M. e Ades, Y. (2000). Understanding the roles of signs and norms in organizations-a semiotic approach to information systems design. Behaviour & information technology, 19(1):15–27.
- Stamper, R. K. (1993b). Signs, norms and information systems. Em Joint ICL/University of Newcastle Seminar on the Teaching of Computer Science 1993: Part IX: Information, páginas 1–35.
- Starbird, K., Arif, A. e Wilson, T. (2019). Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW):1–26.
- Starbird, K., Dailey, D., Mohamed, O., Lee, G. e Spiro, E. S. (2018). Engage early, correct more: How journalists participate in false rumors online during crisis events. Em Proceedings of the 2018 CHI conference on human factors in computing systems, páginas 1–12.
- Tacchini, E., Ballarin, G., Della Vedova, M. L., Moret, S. e de Alfaro, L. (2017). Some like it hoax: Automated fake news detection in social networks. arXiv preprint arXiv:1704.07506.
- Tambuscio, M., Ruffo, G., Flammini, A. e Menczer, F. (2015). Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. Em Proceedings of the 24th international conference on World Wide Web, páginas 977–982.
- Tan, Z., Wu, D., Gao, T., You, I. e Sharma, V. (2019). Aim: Activation increment minimization strategy for preventing bad information diffusion in osns. Future Generation Computer Systems, 94:293–301.
- Tanaka, Y., Sakamoto, Y. e Matsuka, T. (2013). Toward a social-technological system that inactivates false rumors through the critical thinking of crowds. Em 2013 46th Hawaii International conference on system sciences, páginas 649–658. IEEE.
- Tang, S., Liu, Q., McQueen, M., Counts, S., Jain, A., Zheng, H. e Zhao, B. (2017). Echo chambers in investment discussion boards. Em Proceedings of the 17th International AAAI Conference on Web and Social Media (ICWSM 2017).
- Taylor-Smith, E., Smith, C. F. e Smyth, M. (2018). Democratic participation through crocheted memes. Em Proceedings of the 9th International Conference on Social Media and Society, páginas 178–186.
- Teixeira, A., Santos, R. d. C., Pontifícia Universidade Católica de São Paulo. São Paulo, S. e Pontifícia Universidade Católica de São Paulo. São Paulo, S. (2020). Fake news colocam a vida em risco: a polêmica da campanha de vacinação contra a febre amarela no Brasil. [Online; accessed 9. Feb. 2021].
- Teixeira, A. S. e e Silva, M. R. (2007). Diálogo sobre a lógica do conhecimento. UFRJ.
- Theng, Y.-L., Goh, L. Y. Q., Lwin, M. O. e Shou-Boon, S. F. (2013). Dispelling myths and misinformation using social media: A three-countries comparison using the case of tuberculosis. Em 2013 IEEE International Conference on Healthcare Informatics, páginas 147–152. IEEE.
- Thompson, P. (2003). Semantic hacking and intelligence and security informatics. Em International Conference on Intelligence and Security Informatics, páginas 390–390. Springer.

- Times (2020). Fake alert: Viral message misinforms readers about boosting ph levels to beat coronavirus. [Online; accessed 8. Jun. 2021].
- Tong, G., Wu, W. e Du, D.-Z. (2018). Distributed rumor blocking with multiple positive cascades. IEEE Transactions on Computational Social Systems, 5(2):468–480.
- Tong, G., Wu, W., Guo, L., Li, D., Liu, C., Liu, B. e Du, D.-Z. (2017). An efficient randomized algorithm for rumor blocking in online social networks. IEEE Transactions on Network Science and Engineering.
- Torres, R., Gerhart, N. e Negahban, A. (2018a). Epistemology in the era of fake news: An exploration of information verification behaviors among social networking site users. SIGMIS Database, 49(3):78–97.
- Torres, R., Gerhart, N. e Negahban, A. (2018b). Epistemology in the era of fake news: An exploration of information verification behaviors among social networking site users. ACM SIGMIS Database: the DATABASE for Advances in Information Systems, 49(3):78–97.
- Trattner, C. e Kappe, F. (2013). Social stream marketing on facebook: a case study. International Journal of Social and Humanistic Computing, 2(1-2):86–103.
- Traylor, T., Freese, C. e Wong, W. (2017). Psyop, deception, and cyberspace in the open: Analysing fake news in a cyber new normal communications environment. Em European Conference on Cyber Warfare and Security, páginas 488–496. Academic Conferences International Limited.
- Tschiatschek, S., Singla, A., Gomez Rodriguez, M., Merchant, A. e Krause, A. (2018). Fake news detection in social networks via crowd signals. Em Companion Proceedings of the The Web Conference 2018, páginas 517–524.
- Tudjman, M. e Mikelic, N. (2003). Information science: Science about information, misinformation and disinformation. Proceedings of Informing Science + Information Technology Education, 3:1513–1527.
- Tyagi, N., Miers, I. e Ristenpart, T. (2019). Traceback for end-to-end encrypted messaging. Em Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, páginas 413–430.
- Uol (2020a). Como desinformação e grupos antivacina ameaçam combate à covid-19 no Brasil. UOL.
- Uol (2020b). Coronavírus: Grupos usam notícias falsas para apoiar postura de Bolsonaro. UOL.
- Uol (2021a). Cloroquina: por que o medicamento foi indicado contra chikungunya, mas não serve contra covid-19? UOL.
- Uol (2021b). Com anúncio pró-cloroquina, jornais depõem contra trabalho de jornalistas. UOL.
- Valdez, A. C. e Zieffle, M. (2018). Believability of news. Em Congress of the International Ergonomics Association, páginas 469–477. Springer.
- van Breemen, A. J. N. (2010). The semiotic framework - peirce and stamper. Em International Conference on Informatics and Semiotics in Organisations.



- Van Heusden, B. e Jorna, R. (2001). Toward a semiotic theory of cognitive dynamics in organisations. Em Information, Organisation and Technology, páginas 83–113. Springer.
- Vargo, C. J., Guo, L. e Amazeen, M. A. (2018). The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. New media & society, 20(5):2028–2049.
- Varol, O. e Uluturk, I. (2018). Deception strategies and threats for online discussions. Deception strategies and threats for online discussions by Onur Varol and Ismail Uluturk. First Monday, 23(5-7).
- Verdoliva, L. (2020). Media forensics and deepfakes: an overview. IEEE Journal of Selected Topics in Signal Processing, 14(5):910–932.
- Vereshchaka, A., Cosimini, S. e Dong, W. (2020). Analyzing and distinguishing fake and real news to mitigate the problem of disinformation. Computational and Mathematical Organization Theory, páginas 1–15.
- Vicario, M. D., Quattrocioni, W., Scala, A. e Zollo, F. (2019). Polarization and fake news: Early warning of potential misinformation targets. ACM Transactions on the Web (TWEB), 13(2):1–22.
- Vitkova, L., Kotenko, I. e Chechulin, A. (2021). An approach to ranking the sources of information dissemination in social networks. Information, 12(10):416.
- Viviani, M. e Pasi, G. (2017). Credibility in social media: opinions, news, and health information—a survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 7(5):e1209.
- Volkova, S. e Jang, J. Y. (2018). Misleading or falsification: Inferring deceptive strategies and types in online news and social media. Em Companion Proceedings of the The Web Conference 2018, páginas 575–583.
- Volkova, S., Shaffer, K., Jang, J. Y. e Hodas, N. (2017). Separating facts from fiction: Linguistic models to classify suspicious and trusted news posts on twitter. Em Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), páginas 647–653.
- Vraga, E. K. e Bode, L. (2018). I do not believe you: How providing a source corrects health misperceptions across social media platforms. Information, Communication & Society, 21(10):1337–1353.
- Wai, H.-T., Ozdaglar, A. E. e Scaglione, A. (2018). Identifying susceptible agents in time varying opinion dynamics through compressive measurements. Em 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), páginas 4114–4118. IEEE.
- Wandoko, W., Fernando, E. et al. (2019). Analysis social media application message trust factor a case study university student in indonesia. Em 2019 6th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE), páginas 1–5. IEEE.



- Wang, F., Mo, Y. e Huang, B. (2008). Defending reputation system against false recommendation in mobile ad hoc network. Em 2008 IEEE International Conference on Networking, Sensing and Control, páginas 488–493. IEEE.
- Wang, L. e Fussell, S. R. (2020). More than a click: Exploring college students' decision-making processes in online news sharing. Proceedings of the ACM on Human-Computer Interaction, 4(GROUP):1–20.
- Wang, P., Angarita, R. e Renna, I. (2018). Is this the era of misinformation yet: combining social bots and fake news to deceive the masses. Em Companion Proceedings of the The Web Conference 2018, páginas 1557–1561.
- Wang, Q., Lu, T., Ding, X. e Gu, N. (2014). Think twice before reposting it: Promoting accountable behavior on sina weibo. Em Proceedings of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design (CSCWD), páginas 463–468. IEEE.
- Wang, W. Y. (2017). Liar, liar pants on fire: A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648.
- Wang, X., Wang, X., Min, G., Hao, F. e Chen, C. P. (2020). An efficient feedback control mechanism for positive/negative information spread in online social networks. IEEE Transactions on Cybernetics.
- Wason, M., Gupta, S. S., Venkatraman, S. e Kumaraguru, P. (2019). Building sociality through sharing: Seniors' perspectives on misinformation. Em Proceedings of the 10th ACM Conference on Web Science, páginas 321–322.
- Watney, M. (2018). The legal position of social media intermediaries in addressing fake news. Em European Conference on Cyber Warfare and Security, páginas 496–XVI. Academic Conferences International Limited.
- Webb, H., Burnap, P., Procter, R., Rana, O., Stahl, B. C., Williams, M., Housley, W., Edwards, A. e Jirotko, M. (2016). Digital wildfires: Propagation, verification, regulation, and responsible innovation. ACM Transactions on Information Systems (TOIS), 34(3):1–23.
- Westervelt, R. (2011). RSA SecurID breach began with spear phishing attack. SearchSecurity.
- Wieringa, R., Maiden, N., Mead, N. e Rolland, C. (2006). Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. Requirements engineering, 11(1):102–107.
- Wijaya, A. P. e Santoso, H. A. (2018). Improving the accuracy of naïve bayes algorithm for hoax classification using particle swarm optimization. Em 2018 International Seminar on Application for Technology of Information and Communication, páginas 482–487. IEEE.
- Wilder, B. e Vorobeychik, Y. (2018). Controlling elections through social influence. Em Proceedings of the 17th international conference on autonomous agents and multiagent systems, páginas 265–273. International Foundation for Autonomous Agents and Multiagent Systems.
- Wolf, M. (2003). Teorias da comunicação de massa. Martins Fontes.

- Wood, G., Long, K., Feltwell, T., Rowland, S., Brooker, P., Mahoney, J., Vines, J., Barnett, J. e Lawson, S. (2018). Rethinking engagement with online news through social and visual co-annotation. Em Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, páginas 1–12.
- Wu, L., Li, J., Morstatter, F. e Liu, H. (2018). Toward relational learning with misinformation. Em Proceedings of the 2018 SIAM International Conference on Data Mining, páginas 711–719. SIAM.
- Xiong, G., Feng, K. e Liu, Y. (2018). Multi-level gray evaluation method for emergencies event network behavior risk perception. Em Proceedings of the 2018 2nd International Conference on Management Engineering, Software Engineering and Service Sciences, páginas 102–106.
- Xuan, K. e Xia, R. (2019). Rumor stance classification via machine learning with text, user and propagation features. Em 2019 International Conference on Data Mining Workshops (ICDMW), páginas 560–566.
- Yang, L., Wang, J., Gao, C. e Li, T. (2019). A crisis information propagation model based on a competitive relation. Journal of Ambient Intelligence and Humanized Computing, 10(8):2999–3009.
- Yao, Y., Viswanath, B., Cryan, J., Zheng, H. e Zhao, B. Y. (2017). Automated crowdturfing attacks and defenses in online review systems. Em Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, páginas 1143–1158.
- Yasser, K., Kutlu, M. e Elsayed, T. (2018). bigir at clef 2018: Detection and verification of check-worthy political claims. Em CLEF (Working Notes).
- Yu, S. e Han, K.-H. (2018). Silent chatbot agent amplifies continued-influence effect on misinformation. Em Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, páginas 1–6.
- Yu, S., Vorobeychik, Y. e Alfeld, S. (2018). Adversarial classification on social networks. Em Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems, páginas 211–219. International Foundation for Autonomous Agents and Multiagent Systems.
- Yuliani, S., Sahib, S., Abdollah, M. F., Al-Mhiqani, M. N. e Atmadja, A. R. (2018). Review study of hoax email characteristic. International Journal of Engineering & Technology, 7(3.2):778–782.
- Zannettou, S., Caulfield, T., De Cristofaro, E., Kourtellis, N., Leontiadis, I., Sirivianos, M., Stringhini, G. e Blackburn, J. (2017). The web centipede: understanding how web communities influence each other through the lens of mainstream and alternative news sources. Em Proceedings of the 2017 internet measurement conference, páginas 405–417.
- Zannettou, S., Sirivianos, M., Blackburn, J. e Kourtellis, N. (2019). The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans. Journal of Data and Information Quality (JDIQ), 11(3):1–37.
- Zetter, K. (2011). Hacker Spies Hit Security Firm RSA. Wired.

- Zhang, A. X., Ranganathan, A., Metz, S. E., Appling, S., Sehat, C. M., Gilmore, N., Adams, N. B., Vincent, E., Lee, J., Robbins, M. et al. (2018). A structured response to misinformation: Defining and annotating credibility indicators in news articles. Em Companion Proceedings of the The Web Conference 2018, páginas 603–612.
- Zhang, H., Fang, Q., Qian, S. e Xu, C. (2019a). Multi-modal knowledge-aware event memory network for social media rumor detection. Em Proceedings of the 27th ACM international conference on multimedia, páginas 1942–1951.
- Zhang, L., Wang, S.-F., Lin, Z.-Z. e Wu, Y. (2019b). Online ballot stuffing: Influence of self-boosting manipulation on rating dynamics in online rating systems. Telematics and Informatics, 38:1–12.
- Zhang, X. e Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. Information Processing & Management, 57(2):102025.
- Zhang, X., Hu, Y. e Wang, J. (2015). Finding rumor source in weighted scale-free networks. páginas 1–4.
- Zhang, Y., Zhao, Q., Lv, H. e Yu, Y. (2019c). A study of emergency web rumors in weibo based on lda. Em 2019 International Conference on Computer Network, Electronic and Automation (ICCNEA), páginas 240–244. IEEE.
- Zhou, L. e Zhang, D. (2004). Building a misinformation ontology. Em IEEE/WIC/ACM International Conference on Web Intelligence (WI'04), páginas 445–448. IEEE.
- Zhou, L. e Zhang, D. (2007). An ontology-supported misinformation model: Toward a digital misinformation library. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 37(5):804–813.
- Zhou, X. e Zafarani, R. (2018). Fake news: A survey of research, detection methods, and opportunities. arXiv preprint arXiv:1812.00315, 2.
- Zhou, Z., Guan, H., Bhat, M. M. e Hsu, J. (2019). Fake news detection via nlp is vulnerable to adversarial attacks. arXiv preprint arXiv:1901.09657.
- Zhu, X., Hao, J., Shen, Y., Liu, T. e Liu, M. (2018). Diffusion of false information during public crises: Analysis based on the cellular automaton method. Computing and Informatics, 37(1):23–48.
- Zhu, X. e Liu, M. (2018). Intervention strategies for false information on two-layered networks in public crisis by numerical simulations. Soft Computing, 22(16):5467–5477.

## APÊNDICE A – APÊNDICES

### A.1 APÊNDICE 1 - SOBRE AS DECISÕES TEÓRICAS PARA ELABORAÇÃO DA DISMISS

Essa seção me pareceu necessária para explicar as decisões teóricas que tomei para o desenvolvimento da abordagem Dismiss. Como articulado por Harrison et al. (2011), entendo que para avaliar uma pesquisa precisa-se saber não somente quais métodos foram usados, mas como e porque foram usados - "precisamos saber como o conhecimento foi gerado" (Harrison et al., 2011). É preciso considerar alguns aspectos teórico-filosóficos desafiadores de se trabalhar com o fenômeno da desinformação digital pela perspectiva da Semiótica Organizacional (SO), especialmente no que diz respeito a complexidade de abordar pontos como a subjetividade na interpretação do que é verdade e do que é mentira (a epistemologia do ponto de vista), as transformações de significado dentro de comunidades, e a adoção de novas crenças - pontos que apesar de não serem tratados diretamente pela Dismiss, acredito ser fundamental estarem presentes no desenvolvimento de uma abordagem responsável com a natureza abrangente do fenômeno e consciente dos desafios ao se trabalhar com a desinformação digital. Assim, essa seção discute como alguns conceitos e perspectivas da SO foram adaptados nessa tese para o estudo do fenômeno da desinformação e como podem ser explorados em trabalhos futuros para estudar outros aspectos do fenômeno não abordados pela Dismiss.

#### A.1.1 As bases teóricas da Semiótica Organizacional

A epistemologia do fenômeno da desinformação representa um dos desafios de pesquisa para o avanço do conhecimento da Ciência da Computação (R. Walker et al., 2019). Uma perspectiva teórico-filosófica que explique a natureza da relação do ser-humano com o conhecimento é necessária para conceituar informação, desinformação, verdade e mentira, e apoiar o estudo da desinformação e seus impactos na sociedade. Essa discussão surgiu pela complexidade com que a Semiótica Organizacional (SO) aborda a ideia de veracidade, como esta complexidade interfere no estudo do fenômeno da desinformação e a extensão de facetas do fenômeno que ela consegue e não consegue lidar de modo consonante com o entendimento do autor sobre a abrangência com que a Dismiss deve abordar o fenômeno. Essa discussão apresenta as visões ontológicas (a natureza de existir na realidade) e epistemológicas (a natureza do conhecimento e a relação que temos com ele) da SO, como entendido pelo autor. Vou discutir algumas implicações sobre a maneira subjetivista que a SO aborda a natureza da realidade e da veracidade das coisas que observamos. Ademais, a SO se mostrou uma ferramenta conceitualmente complexa e flexível, capaz de ser adaptada para vários desafios do fenômeno da desinformação.

##### *A.1.1.1 Abordando a Realidade - Sobre a natureza Objetiva e Subjetiva da informação*

A SO considera que a informação é uma entidade **subjetiva** que pode ter diferentes interpretações de acordo com as circunstâncias. Os indivíduos em uma comunidade tem perspectivas únicas carregadas de vieses subjetivos particulares que determinam aspectos da comunicação, do entendimento, e da resolução de problemas pela comunidade (Stamper et al., 2000). Essa perspectiva assume diversas realidades (as perspectivas de cada indivíduo) que são compartilhadas. No fenômeno da desinformação, a perspectiva subjetivista sobre a realidade traz

alguns desafios teóricos e conceituais de como entender a veracidade e a responsabilidade para os participantes de um caso de desinformação<sup>1</sup>.

Na SO, a construção daquilo que um indivíduo entende como real (seu entendimento da realidade que percebe), se dá pela significação daquilo que ele percebe no ambiente. Assim, o significado de algo assume a existência de uma realidade, da qual se mapeia percepções em entendimentos. A SO considera a realidade pelo paradigma do Atualismo de Stamper (Stamper et al., 2000). O Atualismo na SO implica que a construção de significado é um processo que requer um indivíduo interpretante, interagindo com o seu ambiente (Adams, 1974), sem ele a realidade não existe. Assim, a SO delimita as propriedades e a extensão da realidade do indivíduo para aquilo que ele consegue perceber. Em relação ao que não está nessa realidade, Stamper conceitua o passado e o futuro como representações da percepção do indivíduo, sendo partes da realidade de algo que existiu ou da possibilidade de ações que existirão (Gazendam e Liu, 2006). Isso significa que a SO enxerga a realidade com um paradigma subjetivista extremo, ou seja, que a realidade de um indivíduo é definida de forma dinâmica e contínua, com base na perspectiva do indivíduo enquanto ele age no ambiente (experiência), percebendo aquilo que o ambiente 'disponibiliza' para esse indivíduo perceber. Assim, o indivíduo constrói uma base de conhecimentos pela interação com o seu ambiente. Dessa forma, a veracidade de algo para um indivíduo é um resultado da significação de sua percepção e é impactada pela experiência e o conhecimento do indivíduo sobre o passado e suas projeções sobre o futuro.

Algumas implicações importantes sobre a visão subjetivista da realidade (Liu, 2000; Gibson, 1977):

- O subjetivismo enfatiza as habilidades dos indivíduos, sua liberdade e responsabilidade em escolher, tanto quanto suas incertezas e conflitos;
- O mundo social é estudado pela busca de um entendimento da forma como o membro da sociedade cria, modifica e interpreta o mundo ao qual pertence;
- A realidade é experiência e é encontrada dentro de um fluxo sensível a transformação;
- "Saber algo ou ser capaz de fazer algo torna-se um item de conhecimento apenas do ponto de vista de um determinado agente. A noção da verdade universal, portanto, torna-se irrelevante neste caso. Em vez disso, o conceito de responsabilidade torna-se tão vital quanto a verdade em uma lógica clássica.- Stamper em Gazendam e Liu (2006).
- O papel da linguagem é visto como descritivo e construtivo.
- As normas fornecem padrões de conduta, mas na maioria das vezes, o desvio de comportamento ocorre na sociedade devido a diferenças em padrões e crenças. Por causa dessas diferenças, o comportamento considerado 'certo' ou 'errado' com base em algum padrão pode ser julgado neutro com base em outros padrões.

#### A.1.1.2 *Semiose e signos*

Para compreendermos como trabalhar com informações e desinformações, partimos da compreensão de como o indivíduo dá significado ao que percebe na realidade que vive. Na Semiótica, o processo de significação - a construção de significado para um indivíduo sobre o mundo que ele percebe - é mediado pela semiose (Liu, 2000). A semiose é uma atividade universal que explica tanto a criação quanto o uso de signos para expressão. Stamper (Stamper,

<sup>1</sup>O conceito de caso de desinformação é definido no Capítulo 2

1991) explica o signo como "algo que representa alguma coisa para alguém em algum aspecto ou capacidade, em alguma comunidade ou contexto social". Assim, a SO entende o signo como a menor partícula concreta de uma representação da realidade, que possui diferentes propriedades que significam coisas particulares para um observador em uma determinada configuração social.

Um signo é representado por três entidades (Liu, 2000): um objeto, um signo que representa esse objeto, e um interpretante. O **Objeto** é o artefato, a forma do signo material, aquilo que transmite uma mensagem para alguém, como marcas de tinta em um papel. O **Signo** é a representação de algo para alguém, um elemento usado na comunicação que referencia o objeto sendo expressado, como na linguagem uma palavra (conjunto de letras) referencia um conceito (entendimento). O **Interpretante** é o efeito para a pessoa que percebe o signo, a inferência do significado do signo para o observador em relação ao objeto (Liu, 2000). A Figura A.1 representa o relacionamento entre essa tríade por meio de um exemplo de um semáforo usados para controlar o trânsito. O objeto é o evento que se deseja transmitir: parar a movimentação de veículos em uma via, o controle do trânsito. O signo/representante é o semáforo com a luz vermelha acesa. E o interpretante é o efeito que a luz vermelha do semáforo causa no observador, ou seja, parar o veículo. O semáforo é um mediador entre o "desejo/intenção" do comunicador e a interpretação do observador.

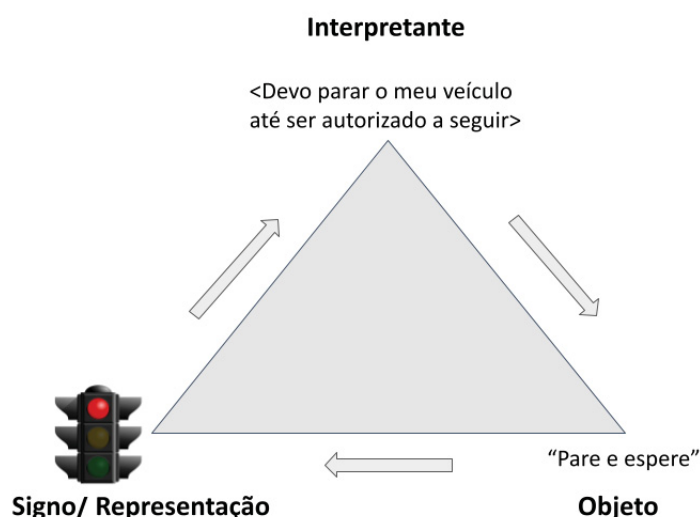


Figura A.1: Relação da tríade semiótica.

A semiose é uma atividade de processamento de signos que constrói o conhecimento da realidade para o indivíduo que a experimenta (Liu, 2000). Essa atividade ocorre o tempo todo, significando percepções, gerando entendimento sobre o mundo que se percebe, e transformando esse entendimento conforme o indivíduo experimenta a realidade (Andersen, 1997). Isso significa que a veracidade de algo percebido por um indivíduo está em constante transformação.

#### A.1.1.3 O papel das *Affordances* e das Normas

A perspectiva de realidade para SO trata o indivíduo como um aprendiz que constrói o conhecimento pela ação. Essa perspectiva é fundamentada pela Teoria das *Affordances* de Gibson, que entende que "nenhuma criatura pode apenas abrir os olhos e ver uma realidade objetiva já existente" (Gibson, 1977). Assim, o indivíduo interage com o ambiente e descobre como ele pode funcionar da melhor maneira possível nesse ambiente. Na Teoria das *Affordances*,



o indivíduo define o mundo real ao fazer sentido dos estímulos externos - como os olfativos e a visuais - e internos - como hormonal e cinestésico - que ele percebe (Gibson, 1977). O termo "*affordances*" é melhor entendido como "aquilo que, na interação com o sujeito, o ambiente potencialmente proporciona, oferece ou convida a agir em razão das propriedades ou características de reconhecimento que possui, ou deveria possuir, para satisfazer certas necessidades de uso" (McCabe, 1982). No entanto para fins de clareza, a definição original, usada na SO é necessária para o entendimento dos desafios de se abordar a subjetividade do fenômeno da desinformação em algumas situações particulares. O conceito original de *affordance* define: 'informações para o organismo', as quais serão percebidas, considerando o que o objeto possibilita e as capacidades do organismo. Cada ambiente disponibiliza um composto de **invariantes** - 'informações sobre o ambiente', padrões de energia percebidos pelos sensores biológicos que carregam informações sobre o ambiente/objeto (Gibson, 1977).

Stamper entende *affordances* como repertórios de comportamento (**invariantes** comportamentais) e distingue as *affordances* físicas e sociais. As *affordances* físicas correspondem aos repertórios de comportamento ligados ao reconhecimento do ambiente físico e a interação do interpretante com os objetos, podem ser vistas como ***affordances* individuais**. Enquanto as *affordances* sociais são repertórios de comportamento associados com o ambiente social e aos seus aspectos particulares como cultura, crenças, expectativas (Stamper et al., 2000). As possibilidades sociais são construídas em conjunto, portanto, são válidas para uma comunidade em particular, têm um tempo inicial e final, e uma autoridade inicial e final (Gazendam e Liu, 2006)<sup>2</sup>.

Em relação ao subjetivismo extremo, Gibson entende que "a criatura precisa aprender o repertório de comportamentos que o ajudam a obter sucesso nas suas necessidades, conforme se desenvolve no ambiente" (Gibson, 1977). Ele apresenta um ótimo exemplo que pode ser usado para ilustrar a ideia de semiose, *affordances* físicas e sociais. Imagine um bebê e uma escova de cabelos sob sua mesa. O bebê não faz ideia do que é o objeto 'escova de cabelo' e vai interagir e criar seu significado (semiose). O bebê vai bater na mesa (*affordance física*), mexer na escova (*affordance física*), colocar ela na boca (*affordance física*), passar na cabeça (*affordance física*), jogar no papai (*affordance física*), etc. Conforme o tempo passa, a escova vai ganhando significados sobre o seu uso como um objeto para o bebê (semiose). Posteriormente, pela observação de outros usando o objeto (*affordance física* e *affordances* sociais) e do aprendizado do bebê (semiose, *affordance física* e *affordances* sociais), o significado do objeto vai ser definido em uma repertório de **invariantes comportamentais**, até o ponto da expressão 'escova de cabelo' denotar o objeto e evocar o repertório de possíveis utilidades (semiose e *affordances*). Essa construção de significado é gradual, ocorrendo desde a infância e os primeiros contatos com esses objetos, através da interação e da descoberta, e se transforma com o passar o tempo (Gibson, 1977). O exemplo explica a ideia de que comportamentos são aprendidos e significados ao ponto de virarem um hábito e serem reconhecidos "automaticamente" no repertório de comportamentos de cada indivíduo<sup>3</sup>.

Então, a SO entende a significação como sendo a relação entre signos e *affordances* (padrões de comportamentos, julgamentos, crenças etc.) que as pessoas podem reconhecer em uma configuração social (organização) (Stamper et al., 2000). Uma organização pode ser vista como um Sistema de Informação sem o computador, onde os indivíduos usam a informação para

<sup>2</sup>*Affordances* físicas e sociais evocam comportamentos nos indivíduos. É nesse sentido que o conceito de *affordances* tecnológicas e digitais são relacionados, como aspectos da tecnologia que informam um tipo de uso, e aspectos de objetos digitais de desinformação que evocam certos comportamentos.

<sup>3</sup>Peirce entende hábitos como tendências adquiridas "para comportar-se de forma similar sob circunstâncias similares no futuro" (Peirce e Buchler, 1902).

realizar ações, como a resolução de problemas, negociações e a execução de tarefas (Andersen, 1997). Dessa forma, Stamper vê uma organização como um sistema social no qual as pessoas se comportam de maneira organizada, seguindo um determinado Sistema de Informações Humano (SIH) (Stamper, 1993a).

As pessoas compartilham certos padrões de comportamento que são governados por um SI social que é construído continuamente pela interação do grupo em uma configuração social (Liu, 2000). Esse SI social é um **sistema de normas** que depende da interação e do consenso da comunidade/organização. Dessa forma, o **consenso** representa um padrão atrelado a um signo, e as *affordances* sociais são um entendimento coletivo (Gazendam e Liu, 2006). De certa forma, como indicado por van Breemen (2010), os Hábitos de Peirce podem ser entendidos como os Consensos em Stamper.

As normas sociais são especificações de padrões normativos de comportamentos vinculados às *affordances* sociais para comunidades (Gazendam e Liu, 2006). Stamper entende as normas como 'conhecimento compartilhado' acumulado por gerações, que condizem com um comportamento desejável, aceitável e exemplar em uma comunidade (Stamper, 1993a). O conhecimento é construído pela interpretação de estímulos da realidade por indivíduos de um grupo através do tempo. É imperativo notar que consensos se transformam com o passar do tempo, assim que um consenso é estabelecido, surgem críticas, dúvidas e mudanças (Liu, 2000). Nesse sentido, surge o conceito de **campo de informação**, que é um conjunto de normas compartilhadas que permite que as pessoas em uma comunidade se comportem de maneira organizada (Gazendam e Liu, 2006). Por exemplo, as crenças de uma comunidade, como os indivíduos que compartilham o consenso de uma terra-plana. Assim, um indivíduo pode simultaneamente fazer parte/possuir de mais de um campo de informação.

Esses conceitos fundamentais são importantes para compreensão de 'significado' na Semiótica Organizacional, no sentido de que cada indivíduo possui um repertório de ações construído continuamente pela sua interação, que representa sua realidade subjetiva, proveniente de *affordances* que reconhece como invariantes comportamentais e está internalizado em um si na forma de um **sistema de normas**. Assim, a interação do observador com o objeto influencia, e é influenciada, pela sua perspectiva particular da realidade, sua cultura, seus valores, crenças, expectativas, experiências, percepções, etc. Assim, a SO trata a realidade subjetiva como uma configuração de invariantes e comportamentos. "Estas invariantes-comportamentais não têm existência independente (não são objetivos), podem ser demonstrados ostensivamente (não são conceitos), incluem coisas que não são signos (não são apenas nomes), e são sempre concretamente instanciados (não estão em uma realidade abstrata)" (Stamper et al., 2000). Isso implica que a semiose é determinante para o impacto da informação na construção da capacidade de interação do indivíduo com seu ambiente, especialmente no seu julgamento de veracidade. Através da semiose, o indivíduo aprende e desenvolve mecanismos de resposta para aquilo que ele consegue perceber e interpretar, construindo sua base de reações significativas de sua interação com o ambiente, sua base de invariantes-comportamentais (*affordances*) ou seja, com base em normas. Dessa forma, a SO representa uma fundamentação adequada para investigar e modelar as ações que um indivíduo executa em relação a sua percepção subjetiva e como sua comunidade impacta suas decisões, julgamentos e ações.

#### A.1.2 Sobre a veracidade e as implicações

O conhecimento compartilhado, seja consensos ou normas, requer algo que certifique sua veracidade (Stamper et al., 2000), uma autoridade certificadora. Finalmente chegamos ao ponto onde podemos discutir a veracidade na SO. Stamper discute a verdade como duas coisas. Na plataforma técnica, a verdade é a matemática e a física, oriundas de uma realidade

objetiva (Stamper et al., 2000). Na plataforma social, a verdade são consensos comunitários, oriundos de uma realidade subjetivista extrema. Para Stamper, o conceito de verdade será algo compartilhado entre os indivíduos da mesma comunidade, um consenso. Ainda, Stamper considera que o conceito de verdade precisa ser substituído pelo de responsabilidade, quando abordando a plataforma social, como uma autoridade certificadora (pode ser uma pessoa, uma lei, um especialista, etc.). Assim, a veracidade é vista como uma propriedade de um signo, o que a coloca sempre em um viés subjetivo de interpretação pelo indivíduo que a interpreta, e que seu valor é determinado como um consenso sobre o que é verdade para a comunidade.

A SO entende que a noção de 'verdade objetiva' - que existe independentemente do que o observador pense - deve ser substituída pela noção de responsabilidade (Stamper, 1993a). A verdade tem a ver com a relação entre as propriedades objetivas da realidade e um modelo desse mundo, enquanto a responsabilidade tem a ver com a relação entre um agente responsável em alguma configuração social e as visões que ele expressa (Stamper, 1993a). Assim, na SO, em vez de um mundo "objetivo", existem várias visões de mundo dos agentes que interagem na organização. Ao explorar esse conceito em conjunto com a ideia de informação, o conceito de verdade objetiva se torna irrelevante, pois o que é importante é aquilo que é significativo para o grupo, aquilo que os agentes acreditam, os seus **consensos**. A verdade para um grupo é um consenso temporário, construído e transformado pela interação em um contexto ao longo tempo.

Essa perspectiva é um tanto desafiadora para investigar a desinformação digital, pois a SO não se aprofunda nos efeitos da mentira na comunicação (não de forma explícita). Nesse sentido, Stamper considera a **precisão uma a propriedade semântica do signo**, que varia de acordo com as *affordances* que as pessoas podem reconhecer (crenças, experiências, etc). Essas *affordances* estão em função de um contexto - um tempo, espaço e um alguém. Isso nos permite instrumentalizar a análise de motivações, causas e efeitos relacionados a desinformação no mundo social - ao menos considerando o campo de informações de uma comunidade. É nesse sentido que a Dismiss incorpora esses fundamentos (os conceitos de semiose, consensos e invariantes-comportamentais, e que percepções evocam comportamentos, sendo possível mapear um campo de informações em um sistema de normas para um indivíduo) para apoiar a investigação de como aspectos físicos e sociais do fenômeno da desinformação podem influenciar os participantes de um caso de desinformação. Porém, é uma perspectiva conceitualmente complexa e que foca nas relações sociais para derivar as ações e atitudes de um indivíduo<sup>4</sup>.

É preciso discutir o que significa uma desinformação, para podermos estudar sua natureza, aspectos e características de uma modo sistêmico. Nesse sentido, essa tese se apoia na Teoria Semiótica de (Eco, 1979). Umberto Eco explica o signo como tudo o que pode ser considerado como substituto significativo (representação) de 'outra coisa'. Para Eco, o signo e a significação implicam que essa 'outra coisa' não precisa existir ou estar em algum lugar no momento em que um signo representa essa 'outra coisa'. Assim, uma ideia representa um signo (Noth, 1990). Para Eco, isso significa que a percepção tem um processo de interpretação de dados desconectados - dados organizados com base em experiências prévias e não somente no estímulo real observado. Como pode não existir uma conexão com algo da realidade, a interpretação invoca outras informações para construir o entendimento do estímulo. "Como quando se observa uma pegada no solo e essa pegada possui indicações de ter sido feita por um gato. Nesse sentido a pegada representa para o interpretante um significado possível que

---

<sup>4</sup>Essa tese entende que esta perspectiva limita a extensão que podemos estudar e modelar o indivíduo e a comunicação, pois não fornece explicitamente uma forma de trabalhar com a factualidade de algo ou o consenso sobre a veracidade de algo. Além disso, é uma perspectiva que torna desafiador modelar estados internos de transformação do conhecimento e do entendimento que ocorrem durante uma interação comunicativa. Aprofundarei minha visão nesse assunto na Seção A.1.3.

foi processado antes do acontecimento (ver o gato)". Essa ideia embasa o que Eco chama de objeto dinâmico e objeto imediato. O objeto dinâmico é a representação construída ligada a um outro signo, o objeto imediato que tem um representante bem definido. O objeto dinâmico está em um processo constante de construção do seu significado. Em algum momento o processo de semiose pode atingir um limite, o que Peirce chama de interpretação lógica final ou **Hábito**, que é uma disposição de agir no mundo, e a possibilidade de ação (Peirce e Buchler, 1902). Isso implica na ideia de Hábito de Pierce ser traduzida na importância do mundo social na construção das normas, ou *affordances* sociais, que são representadas o **consenso** para Stamper (Stamper et al., 2000). Para Stamper, as *affordances* são construídas também com base no conhecimento da comunidade, e assim como quando um indivíduo age sobre o mundo ele assume responsabilidade com base nas normas de sua comunidade, tendo uma predisposição a interpretar ou agir como base nesse Hábito ou norma, à algum estímulo. Dessa forma, essa tese faz uma ligação entre Eco e Stamper de forma minimamente compatível.

Essa tese entende que o significado de algo depende da perspectiva subjetiva de um indivíduo, assim como ocorre para os outros indivíduos que estão experienciando este mesmo algo. O "significado" é algo ao mesmo tempo individual e compartilhado, que se transforma durante a interação, sendo influenciado por diversos fatores, como as experiências prévias, o conhecimento do indivíduo, as suas capacidades de interação com/e do seu ambiente (*affordances*) (Liu, 2000). Assim, aquilo que um indivíduo de uma organização percebe como verdade pode não ser percebido por outro com o mesmo significado, e a sua organização pode influenciar a forma como algo é interpretado (como casos do efeito manada (Nuruddin Hidayat et al., 2020)). Dessa forma, essa tese entende que a veracidade de uma mensagem 'está' para um indivíduo, e é intimamente relacionada às circunstâncias que a mensagem é observada. Nesse contexto, a objetividade factual do conteúdo de uma mensagem pode ter pouco impacto na interpretação das partes interessadas que interagem com ela. Nesse sentido, a desinformação pode ser entendida como *um signo que não corresponde a um estado das coisas na realidade que um indivíduo percebe ou entende* (Eco, 1979).

Considerando a perspectiva de intencionalidade apresentada em (Tudjman e Mikelic, 2003) e as ideias de signo (de Stamper (1991) e de Eco (1979)) podemos definir desinformação acidental como *o uso não intencional de um signo que não corresponde a um estado das coisas na realidade que um indivíduo percebe ou entende em uma determinada circunstância*. Do mesmo modo podemos definir desinformação deliberada como *o uso intencional de signos que não correspondem a um estado das coisas na realidade que um indivíduo percebe ou entende em uma determinada circunstância, com um propósito particular em manipular o entendimento deste indivíduo para obter algum tipo de ganho*.

Essas definições de desinformação intencional e deliberada servem de instrumento na Dismiss para diferenciar aspectos particulares de cada tipo de desinformação, em especial permite identificar padrões nos objetos desinformativos entre aspectos inter-relacionados nos degraus da escada semiótica.

### A.1.3 Modelos de Comunicação, Semiótica Organizacional e os Desafios da Análise Pragmática

A SO dispõe de artefatos e métodos aptos à análise e modelagem das interações e SIH's. Liu (2000) sugere o uso do modelo de comunicação proposto pela Teoria do Ato de Fala (TAF) para análise da comunicação, e o uso da Lógica Deôntica (LD) para o estudo das relações estabelecidas entre os agentes em um determinado meio e para antecipar os potenciais efeitos de signos na comunicação. No entanto, algumas premissas da TAF e da LD trazem complicações para abordar determinados contextos de comunicação e tipos de interação, em



especial os contextos "não confiáveis" e sua relação com o comportamento dos indivíduos (Sbisà e Fabbri, 1980).

Em especial, estas teorias podem abordar causa e efeito como sendo o impacto de um sistemas de normas no comportamento dos indivíduos como sendo garantido pela natureza destas próprias normas. No entanto, elas não garantem sozinhas a coerência entre ato/fala e intenção dos agentes, e dificultam a previsibilidade de como um ato/fala será compreendido, ou das transformações sociais decorrentes (Sbisà e Fabbri, 1980). Essa perspectiva sera explicada e discutida a seguir.

A SO foi proposta para analisar e modelar contextos organizacionais com características específicas, como ambientes de negócios, que assumem premissas como a honestidade e a responsabilidade de agentes sobre os seus "atos e falas", e sobre seguir ou não as normas e consensos (responsabilidade). Estas premissas (por exemplo, seguir as normas) são consideradas nas pressuposições sobre os efeitos dos signos na interpretação e comportamento dos agentes envolvidos, em especial quais são os seus papeis na comunicação, como eles esperam que um agente se comporte e como uma determinada informação poderá ser interpretada, antes mesmo da comunicação ocorrer (Liu, 2000) - basta saber alguns valores de variáveis contextuais (Sbisà e Fabbri, 1980), como as normas formais e informais no campo de informação do individuo, e as mecânicas e procedimentos da organização. Nesses contextos, a TAF e a LD fornecem o aparato necessário para análises e modelagens.

No entanto, na comunicação de desinformação, os agentes atuam/comunicam mensagens que potencialmente podem diferir de seus estados mentais (intenções). Nesse **ambiente de comunicação não confiável**, os agentes entram em uma negociação (de regras, normas, comportamentos, papeis, interpretação, etc.) que ocorre durante o ato/fala (interação social), e que muda conforme a interação se desenvolve (Sbisà e Fabbri, 1980). Nesse ambiente, as normas e regras (culturais, legislativas, etc), sozinhas, não são determinantes para a forma como uma comunicação será entendida, tampouco implicam na honestidade/ sinceridade dos agentes envolvidos na comunicação. Elas funcionam como dispositivos de força ilocucionária <sup>5</sup> que são empregados de formas subjetivas por cada agente ouvinte/observador e que podem afetar, em maior ou menor grau, suas interpretações nas circunstância em que ocorre a interação/comunicação (Sbisà e Fabbri, 1980). Nesse contexto, o agente que escuta executa procedimentos de avaliação e negociação de significado (semiose) enquanto a comunicação/ato está ocorrendo<sup>6</sup>, e falseia a extensão da veracidade do que ele observa e do seu próprio entendimento (Sbisà e Fabbri, 1980). O ouvinte executa esses procedimentos quando bem entende e os estende até que a **extensão da veracidade daquilo que ele avalia seja aceitável**, e ele se dê por satisfeito, aceitando (ou rejeitando) o ato de fala (Sbisà e Fabbri, 1980). <sup>7</sup> como honestos ou "feliz"/ ou rejeitando/avaliando como "infeliz" (Sbisà e Fabbri, 1980).

Nesta circunstancias, as **normas** culturais são melhor vistas como **princípios** - " não como normas que governam o comportamento, mas como dispositivos interpretativos usados para compreensão (Sbisà e Fabbri, 1980)"- que são empregados da maneira que o ouvinte deseja para compreender o ato/comunicação, juntamente com diversos outros dispositivos que ele avalie relevantes (Sbisà e Fabbri, 1980). Por este motivo, as pressuposições de um sistema

<sup>5</sup>Elementos usados para interpretação e avaliação dos efeitos de um ato/fala (por exemplo, a força de uma ordem, uma promessa, um pedido de desculpas, uma declaração...) (Liu, 2000).

<sup>6</sup>Decidindo por exemplo, o papel do falante e o seu próprio, o grau da credibilidade do falante, ao que determinado signo representa quando usado pelo falante naquela situação, as próprias normas que ele quer seguir, o que ele infere/deduz como objetivo da comunicação, etc.

<sup>7</sup>Por ato ilocucionário essa tese entende: "o aspecto particular do ato de fala que pode ser abstraído de seu todo na medida em que o ato de fala conta como tendo uma certa força (por exemplo, a força de uma ordem, uma promessa, um pedido de desculpas, uma declaração...) (Sbisà e Fabbri, 1980)."

de normas e do efeito de variáveis contextuais como determinante para prever o efeito da interação/comunicação na dimensão Social, tornam o modelo de comunicação da TAF, como um modelo de relação de poder unilateral com um sistema normativo predominantemente daquele profere o ato de fala (Sbisà e Fabbri, 1980), um tanto limitado para analisar o fenômeno da desinformação. Ele não permite abordar facilmente o ouvinte como um participante ativo na interação com suas próprias avaliações e decisões interpretativas, com suas regras pessoais e internas que impactam sua ação tanto quanto regras externas. Isso levou a desafios durante a elaboração da Dismiss relacionados a organizações onde não há consensos formados ou os consensos estão sendo revisado (como na polarização política de uma sociedade).

Além disso, modelar as mudanças que o ato/fala ocasionam - transformação social que ocorre por efeito do ato de fala, o efeito perlocutório<sup>8</sup> (Sbisà e Fabbri, 1980) - durante o próprio ato requer uma análise constante das transformações que ocorrem na situação de interação comunicativa - de modo situado, empírico e sincronizado. Assim, os aspectos **circunstanciais** (Os aspectos decididos durante a interação, como as regras, entendimentos, negociações, etc.) da comunicação de desinformação que estão inter-relacionados e a sua relevância para o impacto social são difíceis de considerar. Por isso, comunicações que ocorrem em um ambiente de interação como uma rede social digital, um grupo de comunicação no WhatsApp, ou um micro blog como o Twitter, são melhor entendidos por esta perspectiva dualista que integra fatores internos e externos da comunicação, na causa e efeito de atos ilocucionários.

Por isso, essa tese entende que em qualquer interação social existe um "Sistema de Informação com as normas socioculturais" da organização, integrando campos de informações, consensos e procedimentos, e introduz o entendimento de **Sistema de Informação de normas circunstanciais/situacionais** da interação, onde os participantes se transformam constantemente, por exemplo, reavaliando o valor do sistema de normas naquele contexto de interação (Sbisà e Fabbri, 1980), ou se apresentam com "segundas intenções". Um exemplo desse tipo de fator seria as regras internas de transformação que um indivíduo usa para aprender e justificar uma nova crença. Na descrição do conhecimento no Atualismo de Lycan (Lycan, 1988), "adquirir uma nova crença é adquirir uma nova representação interna real instanciada na arquitetura funcional de alguém. E uma nova crença é justificada se, melhor do que seus concorrentes, aumentar a coerência explicativa de todo o conjunto de crenças da pessoa". Esta é a teoria explicativa da justificação de Lycan, onde "os aspectos do valor epistêmico usados para identificar graus de coerência explicativa são eles próprios justificados pelo apelo à teleologia natural"<sup>9</sup>. Essa tese entende que o estudo desses fatores é importante para o desenvolvimento responsável de intervenções para a desinformação digital e a mitigação de seus efeitos prejudiciais.

Por exemplo, nessa nossa interação temos um sistema de normas que nos permite realizar a comunicação com um certo nível de qualidade de expressão e entendimento. A linguagem utilizada é de comum entendimento, assim como o objetivo dessa comunicação, por exemplo. Nessa interação, temos papéis bem definidos pelo sistema de normas. Eu sou o autor e vocês são a audiência. Além disso, eu sou um especialista em computação autenticado por uma autoridade certificadora (Diploma universitário, especialização, mestrado, etc.), o que me coloca como uma autoridade diante de vocês para assuntos que sou certificado. Além disso, eu sou professor e cientista em formação, que são minhas posições dentro da organização, vocês são advogados, programadores, policiais, professores e qualquer quer seja a profissão de vocês. Entretanto, conforme eu vou me expressando, alguns de vocês podem se interessar na ideia de que eu possa estar mentindo, por exemplo quando algo que eu expresse entra em conflito com algo que vocês tem certeza. Nesse momento, nossos papéis nessa interação comunicativa podem ser reavaliados

<sup>8</sup>efeito do ato produzido no ouvinte (Liu, 2000).

<sup>9</sup>O estudo filosófico do propósito e da finalidade.



e a percepção de vocês sobre mim pode mudar. Outros de vocês podem ter uma personalidade que desconfia de tudo, e que me coloca meu papel e autoridade em questionamento antes mesmo de receberem uma mensagem minha. Essa é a ideia dos níveis de Sistema de Informação Social e do Sistema de Informação Circunstancial.

Neste "ambiente circunstancial"<sup>10</sup>, um agente pode manipular a atenção, percepção, e avaliação da audiência através da comunicação, empregando deliberadamente a desinformação como um dispositivo de influência em diversos níveis de abstração<sup>11</sup>. Nesse sentido, esse conceito possibilita estudar os fatores internos de um agente que o levam a interpretar algo de uma maneira (Como vieses cognitivos e heurísticas) e apoiam a busca pelo entendimento de como um crença pode ser transformada. Por isso, essa tese entende que o modelo de comunicação da TAF e a LD não são facilmente empregadas para análise pragmática de casos de **desinformação deliberada** na comunicação. Entretanto, por motivos de simplicidade, a Dismiss aborda somente o conceito de SIH com as normas socioculturais. O conceito de SIH de normas circunstanciais é uma ideia a ser desenvolvida em trabalhos futuros.

Sbisà e Fabbri (1980) sugerem que é preciso analisar a comunicação enquanto ocorre e coletar empiricamente os aspectos que possam potencialmente estar ligados com a interpretação do ouvinte e as transformações sociais que decorrem do ato/fala, ante a predeterminação de efeitos sociais embasados em teoria e pressuposições. Como isso nem sempre é possível, é preciso ao menos considerar esses aspectos que impactam na interpretação e ação dos ouvintes no ato de comunicação de desinformação deliberada. Os autores apontam as teorias de comunicação oriundas da Sociologia da Interação (micro-sociologia) como alternativas que oferecem modelos de comunicação mais capazes de lidar com os aspectos da pragmática e da articulação de aspectos circunstanciais/situacionais na comunicação. Estes modelos tem a sua disposição dispositivos analíticos de comunicação que permitem ampliar a investigação da interação entre pessoas através de TICs, como a interação estratégica e a retrospectiva (Goffman, 1974) - dispositivos de força ilocucionária que mudam a compreensão dos participantes de um ato/fala enquanto ele ocorre (Sbisà e Fabbri, 1980).

Por esses motivos, a Dismiss investigou e propôs um modelo conceitual de comunicação capaz de estruturar o estudo das diferentes perspectivas que integram a comunicação de desinformação e que permite contemplar qualquer tipo de meio e canal de comunicação, de redes sociais a e-mails. Além disso, a Dismiss incorpora a Análise de Enquadramento (Goffman, 1974) como um método para apoiar a caracterização das perspectivas dos participantes de um caso de desinformação, como uma forma de olhar para os signos que constituem uma mensagem em uma situação de uso. Ainda, a Dismiss se ampara na Semiótica Organizacional para sustentar a análise de aspectos sociais na construção das perspectivas dos participantes - como como uma forma de observar os significados atribuídos coletivamente aos signos que constituem uma mensagem. Como trabalhos futuros o autor pretende continuar com a teorização e desenvolvimento de uma abordagem capaz de caracterizar o SIH de normas socioculturais e o SIH de normas circunstanciais/situacionais dos envolvidos na comunicação de desinformação. No momento a Dismiss lança uma luz sobre os aspectos que impactam a comunicação de desinformação

<sup>10</sup>Entendo o termo "ambiente circunstancial" como um ambiente não confiável, não seguro, incerto, onde o ouvinte avalia constantemente a credibilidade e veracidade das mensagens que recebe

<sup>11</sup>As limitações do TAF (e LD) em abordar a pragmática da comunicação, em especial a análise pragmática em contextos "não confiáveis" e da pressuposição de efeitos perlocutórios com base em premissas de honestidade, tornam a análise pragmática da comunicação de desinformação um desafio complexo e exaustivo. No capítulo de Avaliação 5 o autor discorre sobre suas tentativas de abordar a comunicação de desinformação usando o TAF, a Análise Semântica, a modelagem de normas pela lógica da LD e esquemas ontológicos durante a modelagem do Ciclo de Vida da Desinformação Digital.

de diferentes perspectivas (externas e internas a um sistema de configuração social), mas não aprofunda a caracterização de estados mentais, e causas e efeitos na transformação de crenças.

#### A.1.4 A Análise de Enquadramento

Esta teoria estuda a organização da experiência de interação social (como uma experiência é organizada em uma determinada situação) (Mendonça e Simões, 2012). Ela compreende a forma como uma informação é organizada em um determinada circunstância e quais são os elementos na comunicação que influenciam a interpretação dos participantes do ato de comunicação (Goffman, 1974). A TE auxilia a modelar a estrutura funcional do comportamento social pela ótica da interação (Goffman, 1974).

Na TE a comunicação se divide em três níveis de elementos (Bateson, 1955): denotativo (conteúdo da mensagem), metalinguístico (a capacidade de transformação da linguagem) e metacomunicativo (elementos que definam as relações daqueles que interagem - enquadramentos). Nesta tese, abordamos os níveis denotativo e metacomunicativo.

Na interação social, os participantes de uma comunidade compartilham certos entendimentos do seu ambiente que permitem que eles entendam o que está ocorrendo em um determinada situação (Bateson, 1955). Por exemplo, distinguir uma brincadeira de uma briga, uma sátira de uma ofensa, etc. Bateson desenvolveu a ideia de enquadramento ao observar a interação de macacos em um zoológico, onde ele percebeu que mesmo que os sinais e ações da interação fossem semelhantes, mas não idênticos, aos de um combate, os animais sabiam que se tratava de uma brincadeira (Bateson, 1955). Isso só poderia acontecer se os animais fossem capazes de algum grau de metacomunicação, ou seja, de sinais que indicassem a mensagem "Isto é uma brincadeira".

Formalmente, enquadramentos são os *"elementos de suporte" da comunicação que permitem entender o que está acontecendo em determinada interação* (Bateson, 1955). Enquadramentos não são inventados pelos agentes, mas mobilizados na interação comunicativa, dependendo de sentidos partilhados (Mendonça e Simões, 2012). Eles são marcos interpretativos generalizados em uma organização, construídos socialmente, que permitem as pessoas entender situações circunstanciais (Porto, 2004). Dessa forma, a TE aborda a cultura como sendo constituída de dispositivos de interpretação usados pelos sujeitos para entender as situações de interação - princípios (Goffman, 1974).

Por isso, a TE busca entender o posicionamento dos participantes de uma interação comunicativa (**footing**). Os **Footings** "são introduzidos, negociados, ratificados (ou não), co-sustentados e modificados na interação. Podem sinalizar aspectos pessoais (uma fala afável, sedutora), papéis sociais (um executivo na posição de chefe do setor), bem como intrincados papéis discursivos (o falante enquanto animador de um discurso alheio)." (Mendonça e Simões, 2012). Além disso, a TE traz o conceito de **chaves**, que é um conjunto de regras e convenções a partir das quais uma atividade é transformada em outra, partindo de um quadro inicial primário e transformando-o em outro conforme a comunicação se desenvolve.

Dessa forma, os enquadramentos apoiam a identificação dos princípios de organização que governam uma situação e o engajamento dos atores nela, os *footings* referem-se de modo mais específico ao posicionamento de tais atores em uma interação com um enquadramento definido, mas passível de transformações (Mendonça e Simões, 2012). Nesse sentido, enquadramentos e *footings* devem ser analisados em articulação nas reflexões sobre as interações entre os atores. Ainda, é preciso considerar que os atores não são completamente livres e independentes na situação de interação. Eles são configurados pela situação e atuam sobre ela, transformando o significado da interação que experienciam (Mendonça e Simões, 2012).

Assim, em qualquer interação social existe o Sistema Humano de Informações com as normas culturais da organização, integrando campos de informações e consensos, e o Sistema Humano de Informações de normas circunstanciais da interação, onde os participantes se transformam constantemente, por exemplo, reavaliando o valor do sistema de normas naquele contexto de interação, e a posição deles na situação (Sbisà e Fabbri, 1980). Neste "ambiente circunstancial"<sup>12</sup>, um agente pode manipular a atenção, percepção, e avaliação da audiência através da comunicação, empregando deliberadamente a desinformação como um dispositivo de influência em uma estratégia de interação.

Nesse sentido, uma parte importante do fenômeno da desinformação e do ciclo de vida da desinformação é compreender os impactos de aspectos sociotécnicos na interação entre pessoas e desinformações através de Tecnologias de Informação e Comunicação. Nesse sentido, a teoria do enquadramento dá um suporte conceitual para investigar em profundidade como diferentes dimensões da experiência de comunicação de desinformação das partes interessadas impactam a observação de um mesmo caso de desinformação. Por exemplo, como o objeto desinformativo foi enquadrado (quais aspectos podem representar indícios de deliberação?), como o canal de comunicação influencia no enquadramento do objeto, como a audiência percebe o objeto e o canal de comunicação, quais aspectos podem influenciar interpretação de diferentes partes interessadas, e as transformações que a interação acarreta no caso através da experiência das partes interessadas.

#### A.1.4.1 Enquadramentos para a desinformação

A ocorrência de um signo desinformativo pode ser acidental ou deliberada (Tudjman e Mikelic, 2003). A desinformação acidental ocorre devido a falhas na representação ou interpretação de modo não intencional (Tudjman e Mikelic, 2003). Por exemplo, falhas de memória, um erro de digitação, ou uma tela quebrada podem prejudicar a expressão ou entendimento de algo. Ainda, aspectos circunstanciais como emoções, multitarefas e privacidade (Shen et al., 2019) podem atrapalhar a comunicação. Por exemplo, quando um agente comunica algo com um objetivo, e outro agente interpreta como se tivesse sido expressado com outro. A Figura A.2 apresenta uma representação de desinformação acidental.

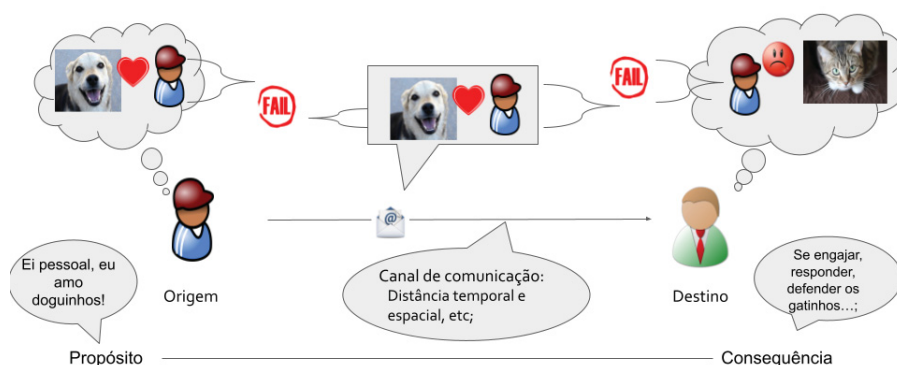


Figura A.2: Diferentes enquadramentos na comunicação que levam ao desentendimento (acidental).

<sup>12</sup>O termo "ambiente circunstancial", não confiável e adversário são usados como sinônimos quando se referindo ao contexto da interação e comunicação de desinformação digital (o meio de comunicação).

Em outra perspectiva existe a desinformação deliberada, que ocorre quando um signo desinformativo é empregado com a intenção de enganar (Tudjman e Mikelic, 2003). Por exemplo, uma sátira, uma falsificação, ou difamação, são desinformações deliberadas construídas com propósitos diferentes. Figura A.3 apresenta uma representação de desinformação deliberada, onde o ouvinte é levado a acreditar em algo.



Figura A.3: Diferentes enquadramentos na comunicação que levam ao desentendimento (deliberado).

Considerando essa perspectiva de intencionalidade podemos definir desinformação acidental como *o uso não intencional de um signo que não corresponde a um estado das coisas na realidade de um indivíduo ou de seu grupo em uma determinada circunstância*.

Do mesmo modo podemos definir desinformação deliberada como *o uso intencional de signos que não correspondem a um estado das coisas na realidade de um indivíduo ou de seu grupo em uma determinada circunstância, com um propósito particular de um agente mal-intencionado em manipular a semiose de um indivíduo ou grupo, para obter algum tipo de ganho*.

Dessa forma, um objeto digital não é só um conjunto de atributos físicos, mas um agregado de atributos relacionados a percepção, julgamentos, emoções, etc. Estes atributos são exemplos de aspectos relevantes para entendimento do fenômeno da desinformação, como motivações, sensações e intencionalidades, que constituem situações na comunicação de desinformação que vão além de um fato como percebido objetivamente. Assim, um objeto digital é um agrupamento de aspectos perceptivos e funcionais que implicam uns aos outros sob a "ideia de um objeto" (Spiegelberg, 1994), ou seja, uma representação ou signo (Peirce e Buchler, 1902). Com base nessa perspectiva podemos instrumentalizar o estudo do fenômeno sem que a noção de verdade seja removida do signo em análise, mas posta entre parênteses como sendo a perspectiva de quem experiencia o signo de desinformação em um determinado contexto de interação.

#### A.1.5 A Cebola Organizacional

A SO se apoia na constatação de que todo comportamento organizado é afetado pela representação e interpretação dos signos pelas pessoas, de modo particular ou em organizações (Liu, 2000). Uma organização é um sistema social no qual as pessoas se comportam de maneira organizada, seguindo um determinado Sistema de Informações Humano (SIH) (Stamper, 1993a). No SIH a informação é criada, processada, distribuída, armazenada e usada por diferentes indivíduos com diferentes atribuições e propósitos.

A SO entende a significação como sendo a relação entre signos e padrões (comportamentos, julgamentos, crenças etc.) que as pessoas podem reconhecer na organização (Stamper et al., 2000). Uma organização pode ser vista como um SI sem o computador, onde os indivíduos usam a informação para realizar ações, como a resolução de problemas, negociações e a execução de tarefas (Andersen, 1997).

Uma organização possui uma arquitetura composta por sub-sistemas de comunicação e operação divididos em três níveis: informal, formal e técnico. Liu (2000) explica que o SI Informal é onde significados e crenças são construídos, e responsabilidades e intenções são compreendidas. No SI Formal, as regras e normas são formuladas, substituindo alguns significados e intenções do SI Informal. No SI Técnico, algumas partes do SI formal são automatizadas por mecanismos e artefatos técnicos. Assim, esta tese aborda o fenômeno da desinformação como sociotécnico, envolvendo aspectos técnicos, formais e informais inter-relacionados (Puska et al., 2019). A Figura A.4 apresenta o artefato da SO Cebola Organizacional (CO), concebido para auxiliar no levantamento de aspectos sociotécnicos do SIH no estudo de uma organização.

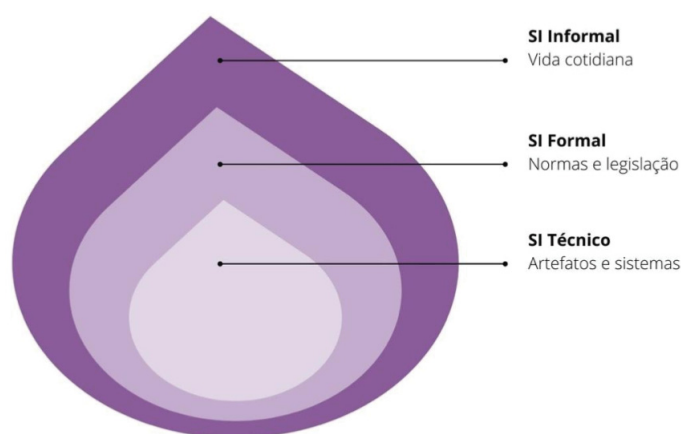


Figura A.4: Representação da Cebola Organizacional. Inspirado em (Ceie-Sbc, 2021)

O seguinte exemplo será usado para discussão de um conceito para desinformação pela ótica da SO. Um agente mal intencionado troca todas as luzes de um semáforo pela cor verde, objetivando o caos no trânsito e a confusão na organização social. Nesse exemplo, o SI Informal constitui a parte mais abrangente da arquitetura de uma organização, representando os consensos e ações aceitáveis para os seus indivíduos. Nesse sentido, os agentes possuem expectativas em relação ao uso correto da sinalização, estão habituados ao funcionamento correto do semáforo, tendem a parar quando o sinal está vermelho, etc. No SI Formal, a burocracia mecaniza as interpretações e ações das pessoas, que seguem as leis que regulamentam o trânsito e indicam o significado das luzes na sinalização, o que fazer quando o semáforo está inoperante, etc. O SI técnico é representado pelo artefato tecnológico que automatiza a sinalização e controla o trânsito de veículos.

Podemos usar a CO para estudar os contextos das partes interessadas do fenômeno da desinformação e identificar aspectos informais, formais e técnicos relacionados a fatores de influência no fenômeno. Uma desinformação poderia ocorrer por diferentes causas em cada sub-sistema do SIH. Por exemplo, No SI técnico, um *bug* no semáforo ou a falta de energia elétrica, podem confundir os motoristas, causar acidentes, etc. No caso da desinformação deliberada, um agente mal-intencionado deturparia a operação do semáforo, trocando todas as luzes pela cor verde ou reprogramando o dispositivo. No SI formal, desinformação acidental pode ser causada pela má-interpretação dos sinais do semáforo, como quando um indivíduo desconhece as leis de trânsito, ou um estrangeiro cujas leis sejam diferentes, etc. De outra forma, um agente



mal-intencionado pode explorar o SI formal para manipular a ação dos motoristas, como explorar o significado da luz verde que induz os motoristas a seguirem. No SI informal, a desinformação acidental pode ser fruto de uma falha de percepção ou maus hábitos, por exemplo, quando um motorista decide avançar no sinal vermelho e outros motoristas seguem essa atitude sem perceber o sinal vermelho. No caso da desinformação deliberada, um agente mal-intencionado pode explorar o hábito ou expectativa dos motoristas de que o semáforo é confiável, que os outros motoristas seguem as regras, etc. Nessa tese, a CO foi testada como um artefato para apoiar a análise sociotécnica da desinformação, como um artefato para identificar aspectos relevantes do contexto das partes interessadas em um evento desinformativo, e apoiou a caracterização dos estágios do ciclo de vida da desinformação com base nos resultados do mapeamento sistemático da literatura.

#### A.1.6 Eixos complementares que sustentam a análise sociotécnica de desinformação

##### A.1.6.1 *Princípios da influência*

Aspectos técnicos, humanos, sociais e organizacionais contribuem para interpretação de um signo de uma determinada forma (Liu, 2000). Sendo assim, é preciso critérios para identificar aspectos sociotécnicos relevantes para entender os enquadramentos usados pelas partes interessadas que interagem com o objeto digital no contexto que ele se propaga. Considerando o entendimento de desinformação para esta tese, a veracidade da mensagem não é um aspecto objetivo da comunicação de desinformação, mas definida por aspectos sociotécnicos do contexto das partes interessadas incorporados na comunicação, impactando a experiência de interação com a mensagem e a interpretação de algo como sendo verdade. Nesse sentido, buscamos amparo na literatura que estuda influência para mapear as características informais e formais da comunicação que influenciam a percepção da audiência e que fazem parte tanto da desinformação acidental, quanto são explorados ativamente na ocorrência de desinformação deliberada. Para isso, adotamos as lentes da Teoria da Persuasão (Cialdini, 1993).

A influência e a persuasão constituem o processo de induzir alguém a fazer, reagir, pensar, ou acreditar em algo que se deseja (Cialdini, 1993). Na comunicação de desinformação, um parte interessada induz uma determinada interpretação (ou é induzida a uma determinada interpretação), seja por acidente ou de modo deliberado. Nesse sentido, aspectos de influência são aqueles tocam no que é importante para as partes interessadas, como emoções, crenças, valores, e no que e como as pessoas estão pensando em um determinado grupo em um determinado momento (Cialdini, 1993). Por isso, é necessário identificar as partes interessadas e mapear o contexto da situação de comunicação onde a desinformação ocorre para identificar aspectos sociotécnicos relevantes e analisar a desinformação digital de modo abrangente.

Os fundamentos da influência e persuasão são baseados em: manipular a atenção, construir confiança, identificar o estado emocional e os interesses da audiência, e explorar estes aspectos ao se comunicar (Hogan, 2010). Segundo (Cialdini, 1993), os princípios da influência - determinantes que influenciam a interpretação das partes interessadas - são:

- **Afinidade:** Pessoas tendem a gostar daquilo que as representa, que se parece com elas. Partes interessadas mal intencionadas exploram elementos como o uso de uma linguagem significativa para uma audiência, uma aparência agradável, usar opiniões positivas, a similaridade na estrutura/diagramação de um conteúdo, o tamanho de uma mensagem, as cores, etc.
- **Aprovação social:** Pessoas querem se sentir validadas pelo que as outras pessoas estão fazendo ao seu redor, em especial quando em dúvida. Eventos ocorrendo no ambiente de um grupo têm impacto no que o grupo percebe como comportamento correto. Elementos



deste princípio são homofilia, confiança, consensos, popularidade de um tema ou pessoa, valores morais, credíes, etc. Este princípio ganha força na presença de incerteza, induzindo a tendência de um indivíduo em agir como o grupo está agindo - o efeito manada.

- **Autoridade:** Pessoas tendem a seguir figuras de autoridade. Partes interessadas mal-intencionadas exploram elementos como a popularidade de indivíduos para o grupo, como celebridades e influenciadores digitais. Outros exemplos são pessoas públicas, especialistas, pessoas que são referência nas atividades que realizam, como médicos, cientistas etc. Este princípio é potencializado por aspectos ligados a incerteza e a ignorância, como a escolaridade, a experiência etc.
- **Comprometimento:** Pessoas querem que seus compromissos sejam consistentes com os seus valores. Comprometimento tem base no que as pessoas estão fazendo, e no que as outras pessoas pensam que ela está fazendo. Assim, partes interessadas mal-intencionadas exploram aspectos como a coerência, a fluência, domínio de um assunto, a expectativa da audiência de ter suas necessidades atendidas, ou defender as ideologias do grupo para gerar afinidade e comprometimento.
- **Escassez:** A tendência das pessoas de temerem perder algo ser pior que a possibilidade de deixar de ganhar. Algo que se torna (ou pode se tornar em breve) escasso ou indisponível, tem mais valor do que quando era abundante - como as corridas por papel higiênico no início da pandemia (Gragnani, 2020), a inexistência de uma vacina/conhecimento sobre o vírus, ou um determinado grupo que luta por direitos. Agentes mal-intencionadas exploram temas como a perda a liberdade (censura, restrições do direito de ir e vir etc.).
- **Reciprocidade:** É a expectativa inerente de quando se recebe algo com valor positivo, um sentimento de dívida surge. Explorado por mensagens ou informações não solicitadas, que geram uma ligação com a necessidade de responder e se engajar, ou atender a pedidos posteriores. É usada para construir um senso de consenso. Alguns elementos explorados são a exposição seletiva, percepção seletiva, e controle do fluxo de informações.
- **Unidade:** A tendência de buscar nos outros, traços de identidade que nos definem, e levam a formação de grupos com indivíduos ideologicamente alinhados. É um princípio que deriva da interação social e da criação de laços de confiança entre indivíduos em grupos. Uma parte interessada mal-intencionada explora aspectos tribais da sociedade, como gírias, sentimentos de familiaridade, proximidade, ou "nós" contra "eles".

Ao analisar a ocorrência de desinformação na comunicação, os Princípios da Influência guiam a identificação de aspectos sociotécnicos que têm potencial de influenciar a semiose das partes interessadas, e como eles podem aprimorar ou restringir a criação, difusão, consumo e mitigação. Estes elementos embasam o estudo do fenômeno no sentido de reduzir o viés de interpretação do analista ao observar a ocorrência de desinformação na comunicação, e mitigar os impactos do "achismo" em relação a um determinado aspecto observado. No entanto, eles não são obrigatórios e podem ser escolhidas outras lentes teóricas para identificar outros tipos de aspecto, ou em uma aplicação mais informal, o conhecimento de especialistas ou do próprio analista. Por isso, a investigação de outros fatores humanos e situacionais, bem como outras teorias da sociologia e psicologia que possam integrar a abordagem, serão investigado em trabalhos futuros.

#### *A.1.6.2 Falácias lógicas e Vieses Cognitivos*

A Tabela A.1 apresenta um catálogo de falácias lógicas usadas em estratégias de manipulação pela desinformação.

Tabela A.1: Catálogo de Falácias Lógicas

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                                                                                                                                                                                                                                                                                                |                                                                                                                                                                                                                                                                                                                                                               |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Falácia do Espantalho:</b> Distorção de argumento particular. Exagerar, distorcer, ou inventar um argumento, tornando muito mais fácil apresentar a sua própria posição como sendo razoável. Este tipo de desonestidade prejudica o debate racional. Exemplo: A teoria de Darwin explica que seres humanos não são muito diferentes de macacos. Entretanto, os seres humanos são mais inteligentes. Logo, eles são diferentes. Assim, a teoria de Darwin é falsa.</p> | <p><b>Raciocínio Circular:</b> Um argumento circular inclui a conclusão na premissa. Esse argumento logicamente equivocado frequentemente nasce em situações em que pessoas possuem uma suposição forte e há longo enraizada, como um axioma, verdades inquestionáveis universalmente válidas. Exemplo: Pensamento circular é ruim porque não é muito bom.</p> | <p><b>O atirador do Texas (Viés):</b> Filtrar propositalmente um conjunto de dados para poder encaixá-lo em seu argumento, ou encontrar um padrão que se encaixa em uma presunção. Exemplo: Esse estudo aqui comprova que o lockdown não funciona. As cidades que fizeram lockdown tiveram uma diminuição insignificante no número de casos, coisa de 5%.</p> |
| <p><b>Presumir Causalidade:</b> Uma correlação, que são coisas que acontecem juntas ou sequencialmente, pode ser uma coincidência. Exemplo: Um novo vírus mortal foi detectado em uma cidade na China. A cidade tem um laboratório de virologia. Logo, a China criou o vírus.</p>                                                                                                                                                                                           | <p><b>Apelo à Emoção:</b> Manipulação de uma resposta emocional. Apelos ao medo, ciúme, ódio, piedade, etc. A falácia ocorre quando a emoção é usada no lugar da argumentação lógica. Exemplo: É fácil falar em lockdown se você não precisa trabalhar. O trabalhador precisa comer, nem todo mundo tem a mesma condição econômica.</p>                        | <p><b>Ad Hominem (Atacar o argumentador):</b> Desmoraliza o oponente, atacando características pessoais e seu caráter. Exemplo: João é virologista e defende a vacinação em massa contra o coronavírus. Não podemos acreditar em João pois ele não toma café e não penteia o cabelo.</p>                                                                      |
| <p><b>Incredulidade Pessoal:</b> A ignorância sobre um fenômeno tende a irracionalidade. Só porque você acha que alguma coisa é difícil de entender ou não sabe como funciona, então você acredita que provavelmente não deve ser verdade. Exemplo: Desconfiança nas vacinas contra o coronavírus porque “fazer vacina demora anos”.</p>                                                                                                                                    | <p><b>Declive Escorregadio:</b> Se A acontecer, então B vai acontecer também. Logo, A não pode acontecer. Falácia lógica que evita a ação com base em presunções e racionalidade não concretas. Exemplo: Com o lockdown, o país vai falir. Não podemos aceitar o lockdown.</p>                                                                                 | <p><b>Apelo à Hipocrisia:</b> Responder a uma crítica com uma outra crítica. Manipula o fluxo do debate e o foco da conversa. Exemplo: Debatedor A diz, - O candidato à presidência B é um mentiroso que não realizou nenhuma ação produtiva em 20 anos como deputado. Debatedor B: - E daí? O candidato B também não é muito produtivo.</p>                  |

|                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                         |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Ônus da Prova:</b> Dizer que quem deve provar a veracidade de uma afirmação não é quem afirma, mas quem contraria. Descartar algo por não estar provado além de toda e qualquer dúvida é um raciocínio falacioso. Exemplo: Claro que o tratamento precoce contra a covid funciona. Então, prove que não!</p> | <p><b>Falácia Anedótica:</b> Utilizar uma experiência pessoal ou exemplo isolado ao invés de um argumento sólido ou evidência convincente. Exemplo: Ivermectina cura a COVID. Todo mundo que eu conheço que tomou, se curou em 7 dias.</p>                                                       | <p><b>Apelo à popularidade:</b> Fazer um apelo à popularidade ou ao fato que muitas pessoas fazem alguma coisa, como uma tentativa de validação desta coisa. Exemplo: Claro que o tratamento precoce contra a COVID funciona. Todo mundo que eu conheço se curou com Ivermectina. Lá em casa não deu uns 10 dias todos ficaram bem.</p> |
| <p><b>Composição/Divisão:</b> Presumir que uma parte de alguma coisa precisa se aplicar ao todo, ou que o todo necessariamente deve aplicar-se às suas partes. Exemplo: Cloroquina é antiviral, usado para tratar vírus. O Coronavírus é um vírus, logo a cloroquina serve pra ele também.</p>                     | <p><b>Falsa Dicotomia:</b> Apresentar dois estados alternativos como as únicas possibilidades para alguma coisa, quando, de fato, existem outras possibilidades. Exemplo: Ou você está alinhado a minha visão de não recomendar lockdown e apoiar o tratamento precoce, ou você é comunista.</p> | <p><b>Apelo à Natureza:</b> Argumentar que já que algo é “natural”, então é válido, justificado, inevitável, bom ou ideal. Exemplo: O PH do corpo mata a COVID. Não precisa de remédios químicos, é só tomar suco de limão e comer abacaxi para aumentar o PH do corpo e ficar imune.</p>                                               |

A Tabela A.2 apresenta um catálogo de vieses cognitivos explorados em estratégias de manipulação pela desinformação.

Tabela A.2: Catálogo de vieses cognitivos

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Ancoragem:</b> A primeira coisa que se julga, influencia o julgamento de tudo o que se segue. As mentes humanas são associativas por natureza, de modo que a ordem em que recebemos as informações ajuda a determinar o curso de nossos julgamentos e percepções. Exemplo: As primeiras informações oficiais sobre o vírus ser uma “gripezinha”, “ancora” a percepção do quão racional ou irracional outras avaliações vão parecer.</p> <p><b>Como evitá-lo:</b> Atenção com aquilo que não se sabe ou não se tem experiência no assunto. “A primeira impressão é a que fica”, então cuidado com as certezas que se tem. Uma análise cuidadosa de situações como essa é essencial para não ser vítima desse viés.</p> | <p><b>A Falácia do Custo Irrecuperável:</b> Você irracionalmente se apegar a coisas que já perdeu, que já lhe custaram algo. Quando investimos nosso tempo, dinheiro ou emoção em algo, é doloroso perder. Exemplo: se você gastou dinheiro em um empreendimento, mas só a pandemia prejudicou a sua rentabilidade por causa do isolamento social. É irracional continuar com a empresa aberta só porque “você já pagou por isso e vai perder o negócio”; especialmente quando isso só lhe faz perder tempo e piorar a difusão do vírus.</p> <p><b>Como evitá-lo:</b> Com uma perspectiva objetiva. O que você aconselharia para uma pessoa querida?</p> | <p><b>A heurística da disponibilidade:</b> Seus julgamentos são influenciados por aquilo que vem à mente com mais facilidade. O quão recentes, emocionalmente poderosas ou incomuns são suas memórias pode torná-las mais relevantes. Exemplo: quando vemos notícias sobre investimentos em Cloroquina e Ivermectina, mudanças de leis e propaganda oficial de ‘Kit Covid’. Podemos ser influenciados a acreditar que estes medicamentos são mais frequentes e importantes no nosso contexto do que realmente são.</p> <p><b>Como evitá-lo:</b> Com uma perspectiva mais ampla, com diferentes fontes de informação e focos. Não confiar apenas nos primeiros julgamentos e nas influências emocionais.</p> |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>A Maldição do Conhecimento:</b> Depois de entender algo, você presume que seja óbvio para todos. Construimos redes complexas de compreensão e esquecemos quão intrincado é realmente o caminho para o nosso conhecimento disponível. Esse viés está intimamente relacionado ao viés retrospectivo, no qual você tenderá a acreditar que um evento era previsível o tempo todo, uma vez que ocorreu. <b>Como evitá-lo:</b> Ao ensinar algo novo a alguém, vá devagar, repita os pontos-chave e facilite a prática ativa para ajudar a incorporar o conhecimento.</p> | <p><b>Viés de Confirmação:</b> Você favorece coisas que confirmam suas crenças existentes. Estamos preparados para ver e concordar com ideias que se encaixem em nossos preconceitos e para ignorar e descartar informações que conflitem com eles. Você poderia dizer que esta é a mãe de todos os preconceitos, pois afeta muito o nosso pensamento por meio do raciocínio motivado. <b>Como evitá-lo:</b> Falseabilismo. Pense nas suas ideias e crenças como se tivessem problemas a serem descobertos por você.</p> | <p><b>O Efeito Dunning-Kruger:</b> Quanto mais você sabe, provavelmente menos confiante você será. Como os especialistas sabem o quanto não sabem, eles tendem a subestimar sua capacidade; mas é fácil ficar confiante demais quando você tem apenas uma ideia simples de como as coisas são. Exemplo: “Eu confio no tratamento precoce- Luciano Hang. A opinião de um leigo não deve ter valor sobre a de especialistas. <b>Como evitá-lo:</b> Cuidado com a abstração como aprendizado. Procure não confundir a cautela dos especialistas com a falta de compreensão, nem dar muito crédito à confiança de leigos.</p> |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Viés de Crença:</b> Se uma conclusão apoiar suas crenças existentes, você racionaliza qualquer coisa que a apoie. É difícil deixarmos de lado nossas crenças existentes para considerar os verdadeiros méritos de um argumento. Exemplo: Acreditamos no tratamento precoce. Ele faz sentido e funcionou na minha comunidade. Qualquer ideia contrária será resistida. <b>Como evitá-lo:</b> Uma pergunta útil é 'quando e como adquiri essa crença?' Temos a tendência de defender automaticamente nossas ideias, sem nunca realmente questioná-las.</p> | <p><b>O efeito de Tiro pela Culatra:</b> Quando algum aspecto de suas crenças centrais é desafiado, isso pode fazer com que você acredite ainda mais fortemente. Podemos experimentar estar errados sobre algumas ideias como um ataque a nós mesmos ou à nossa identidade tribal. Isso pode levar a um raciocínio motivado que causa um reforço de crenças, apesar de evidências não confirmadas. A maioria das pessoas aceitará uma correção relacionada a fatos específicos, no entanto, o efeito de tiro pela culatra pode reforçar uma crença relacionada à medida que as pessoas tentam reconciliar uma nova narrativa em seu entendimento. <b>Como evitá-lo:</b> Correções de amigos são mais aceitas do que de estranhos. Incremento gradual da narrativa com a nova argumentação.</p> | <p><b>O Efeito Halo:</b> O quanto você gosta de alguém, ou o quão atraente ele é, influencia seus outros julgamentos sobre ele. Nossos julgamentos são associativos e automáticos e, portanto, se quisermos ser objetivos, precisamos controlar conscientemente as influências irrelevantes. Se alguém é bem-sucedido ou falha em uma área, isso também pode influenciar injustamente nossas expectativas em relação a essa pessoa em outra área. Exemplo: A autoridade do poder executivo recomendando kit covid. <b>Como evitá-lo:</b> Se você perceber que está dando notas consistentemente altas ou baixas em todas as áreas, vale a pena considerar que seu julgamento pode estar sofrendo com o efeito halo.</p> |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|



|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>O Efeito Barnum:</b> Você vê especificidades pessoais em declarações vagas, preenchendo as lacunas. Como nossa mente é dada a fazer conexões, é fácil pegarmos afirmações nebulosas e encontrar maneiras de interpretá-las de modo que pareçam específicas e pessoais. Videntes, astrólogos e outros usam esse preconceito para fazer parecer que estão dizendo a você algo relevante. Como evitá-lo: Considere como as coisas podem ser interpretadas para se aplicarem a qualquer pessoa, não apenas a você.</p> | <p><b>Pensamento de Grupo:</b> Você permite que a dinâmica social de uma situação de grupo anule os melhores resultados. A discordância pode ser desconfortável e perigosa para a posição social de uma pessoa e, com frequência, a voz mais confiante ou primeira voz determinará as decisões do grupo. Por causa do efeito Dunning-Kruger, as vozes mais confiantes também são frequentemente as mais ignorantes. <b>Como evitá-lo:</b> Em vez de contradizer os outros abertamente, procure facilitar meios objetivos de avaliação e práticas de pensamento crítico como uma atividade de grupo.</p> | <p><b>O Efeito de Enquadramento:</b> Você se permite ser indevidamente influenciado pelo contexto e pela entrega. Todos nós gostamos de pensar que pensamos de forma independente, mas a verdade é que todos nós somos, de fato, influenciados pelo contexto. A formulação de como uma questão é colocada mostrou ter um efeito significativo no resultado. <b>Como evitá-lo:</b> Somente quando temos a humildade intelectual para aceitar o fato de que podemos ser manipulados, podemos esperar limitar o quanto somos. Tente estar ciente de como as coisas estão sendo colocadas para você.</p> |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

## A.2 APÊNDICE 2 - DESDOBRAMENTOS

Esta seção apresenta discussões sobre a implicação de conceitos da Semiótica Organizacional, como *affordances* e consensos, e como podem servir para estudar causas e efeitos da desinformação deliberada. Além disso, o autor delinea um conceito para Vulnerabilidades Humanas no fenômeno de desinformação e como estudá-las.

### A.2.1 Vulnerabilidades Humanas no Fenômeno da Desinformação Digital

A SO considera propriedades do signo que indicam uma perspectiva para conceituar informação. Por exemplo, a precisão representa uma qualidade semântica do signo em detrimento do que objeto que ele representa. A SO entende que a noção de 'verdade objetiva' - que existe independentemente do que o observador pense - deve ser substituída pela noção de responsabilidade (Stamper, 1993a). A verdade tem a ver com a relação entre as propriedades objetivas da realidade e um modelo desse mundo, enquanto a responsabilidade tem a ver com a relação entre um agente responsável em alguma configuração social e as visões que ele expressa (Stamper, 1993a). Em vez de um mundo "objetivo", existem várias visões de mundo dos agentes que interagem na organização. Ao explorar esse conceito em conjunto com a ideia de informação, o conceito de verdade objetiva se torna irrelevante, pois o que é importante é aquilo que é significativo para o grupo, aquilo que os agentes acreditam, os seus **consensos**. Assim, a verdade para um grupo é um consenso temporário, construído e transformado pela interação em um contexto ao longo tempo. Os consensos representam um aspecto que uma parte interessada com intenções de manipulação pode explorar para induzir comportamentos em um grupo.

O semiótico Umberto Eco (Eco, 1979), explora em sua teoria semiótica a relação entre um objeto e sua representação na forma de signos. Eco explica o signo como tudo o que pode ser considerado como substituto significativo (representação) de 'outra coisa' (objeto/ideia/nomes), e que essa 'outra coisa' não precisa existir ou estar em algum lugar no momento em que um signo a representa. Assim, Eco entende que "cada vez que há possibilidade de mentir, há uma função-signo (semiose) que é: significar (e então comunicar) algo ao qual nenhum estado real de coisas corresponde" (Eco, 1979). Estendendo esse entendimento, o conceito proposto por essa tese: *a desinformação deliberada configura o uso intencional de signos que não correspondem a um estado real das coisas no ambiente ou dos consensos do grupo, com um propósito particular de um agente mal intencionado em enganar e manipular o entendimento ou ações dos indivíduos deste grupo.*

Pra Eco, a percepção tem um processo de interpretação de dados desconectados - dados organizados com base em experiências prévias e não somente no estímulo real observado. Como pode não existir uma conexão com algo da realidade, a interpretação invoca outras informações para construir a representação do estímulo. Como quando se observa uma pegada no solo e essa pegada possui indicações de ter sido feita por um gato. Nesse sentido a pegada representa para o interpretante um significado possível que foi processado antes do acontecimento (ver o gato). Essa ideia embasa o que Eco chama de objeto dinâmico e objeto imediato. O objeto dinâmico é a representação construída ligada a um outro signo, o objeto imediato que tem um representante bem definido. O objeto dinâmico está em um processo constante de construção do seu significado. Em algum momento o processo de semiose pode atingir um limite, o que Peirce chama de interpretação lógica final ou **Hábito**, que é uma deposição de agir no mundo, e a possibilidade de ação (Peirce e Buchler, 1902). Isso implica na ideia de Hábito de Pierce ser traduzida na importância do mundo social na construção das normas, ou *affordances* sociais, que são representadas como o **consenso** para Stamper (Stamper et al., 2000). Para Stamper, as *affordances* são construídas também com base no conhecimento da comunidade, e assim como

quando um indivíduo age sobre o mundo ele assume responsabilidade com base nas normas de sua comunidade, tendo uma predisposição a interpretar ou agir como base nesse Hábito ou norma, à algum estímulo. Esta condição pode ser explorada pela desinformação, pois em uma organizações as normas sociais podem ser mapeadas em um modelo da organização. Sabendo as normas sociais e suas operações, essa tese entende ser possível construir a desinformação para explorar a organização de alguma forma, de acordo com alguma propósito, se aproveitando de hábitos e comportamentos formais e informais, em conjunto com sua capacidade operacional técnica.

Nesse contexto, as ideias de Eco sobre a semiose e como o processo de interpretação funcionam, representam um caminho para explorar a desinformação deliberada pela ótica da semiótica. Van Heusden e Jorna (2001) entende o conhecimento como uma construção cumulativa de significados, considerando a memória e expressão dessas representações pelo uso de signos. Van Heusden e Jorna (2001) explica as dinâmicas cognitivas nas organizações como processos de semiose embasados em três perspectivas de objetos semióticos multidimensionais: perceptual (objetos de uma dimensão), codificação (objetos de duas dimensões) e abstrativo (objetos de três dimensões). O perceptual constrói o conhecimento tácito, é ativado por alterações entre as representações armazenadas na memória e os estímulos percebidos ou também quando se encontra um objeto sem reapresentação na memória. Isso ativa uma análise e construção do significado (semiose), que dependendo do estímulo, se pautará em outras informações pra construir sua representação (objeto dinâmico de Eco). Como quando se come um fruta inédita, a primeira coisa que lhe vem na cabeça é ... "tem um gosto parecido com...", é o uso de experiências prévias, normas, hábitos, memórias e ideias, como signos, para representar algo novo. Nessa construção, um conhecimento pode ser construído com base num representação pouco precisa dos estímulos percebidos. O segundo tipo, provem da codificação desse conhecimento tácito usado para comunicação. No exemplo da fruta, códigos e categorias são empregadas pra expressar essa representação de um entendimento abstrato. Como quando alguém nos diz que a fruta é uma manga que se desenvolve em regiões tropicais. Agora, categorias foram atribuídas ao estímulo, enriquecendo a enciclopédia de conhecimento do interpretante. Nesses caso, a desinformação pode atuar na representação de códigos que desviam das normas sociais, como explorar a semelhança entre conhecimentos tácitos e categorias. Por exemplo, no caso de um estímulo falso, como cloroquina ser um remédio de categoria antiviral, e por isso se aplicada a uma doença que não se relaciona com a função original do remédio, ou não se sabe se funciona. Um terceiro processo corresponde ao conhecimento abstrativo, como quando o interpretante usa seu conhecimento tácito em conjunto com os códigos para entender as estruturas entre classes. A desinformação aqui também pode explorar a semelhança entre categorias, como a diagramação de uma fake news com uma notícia real induz a uma percepção de credibilidade ou veracidade.

Podemos incorporar essas ideias com as ideias de hábitos e consensos, e de que eles representam normas que são *affordances* reconhecidas em estímulos percebidos na realidade, e que elas evocam comportamentos previsíveis. Se *affordances* são parte de um sistema de normas, e se o indivíduo se baseia nesse sistema para tomar decisões, avaliar situações, e julgar informações, é possível prever com um certo grau de precisão como um determinado estímulo pode afetar um determinado indivíduo, basta saber quais normas (internar e externar) ele segue e como ele as entende. Em relação as normas sociais, a Semiótica Organizacional tem um ferramental para guiar a investigação, permitindo: caracterizar um campo de informações de um indivíduo/comunidade, modelar normas e procedimentos de uma comunidade/organização, e identificar possíveis causas e efeitos no nível do Sistema de Informação Social.

"As normas fazem os membros de um comunidade tender a se comportar ou pensar de uma certa forma"(Stamper, 1993b). Em uma relação parecida com Van Heusden e Jorna (2001),

Stamper explica as normas em um campo de informação podem ser perceptivas, cognitivas, avaliativas e comportamentais (Gazendam e Liu, 2006). Normas perceptuais são consensos implícitos de como ver o mundo, sobre o que existe no contexto compartilhado pelos observadores, suas capacidades sensoriais. Normas cognitivas são as crenças e conhecimentos padronizados de um grupo, o entendimento de como as coisas acontecem ou devem acontecer. Normas avaliativas conduzem o grupo em direção a consensos, como julgar a credibilidade/precisão de alegações. Normas comportamentais regem o comportamento social adequado dentro de um grupo, são o modo de agir na comunidade. Ainda, as normas podem ser explícitas e formais, como documentos do governo e os livros de leis, ou até um acordo verbal explícito e informal, como quando todos os indivíduos em um grupo de pesquisa concordam em contribuir para o café do laboratório. As normas também podem ser implícitas e informais, como o comportamento de não 'furar' uma fila.

O Sistema de normas internalizado ou individual mapeia estímulos em padrões de comportamento invariante. Isso quer dizer que é possível manipular ou evocar um determinado comportamento se usarmos um estímulo específico. Por exemplo, um alarme de incêndio faz com que todos em uma organização tomem uma ação, mesmo se o alarme for falso. O mesmo pode ocorrer com um aviso de vírus no computador, vão desencadear um processo de semiose em uma cadeia de comportamentos dependendo de quem percebe esses estímulos. Esses comportamentos estão ligados a atribuições sociais formais e informais e a capacidades físicas e técnicas de cada um. O ponto é que sabendo um **hábito ou consenso** é possível interferir na dimensão social. Sabendo um procedimento é possível interferir no certame formal e técnico. Sabendo como estímulos são interpretados e como o mundo técnico opera, podemos nos intrometer e controlar o que alguém vê sem que ele saiba disso. Essa ideia fica clara com a análise do caso de desinformação "Spearphishing da RSA" feita usando a Dismiss e apresentada de modo resumido no Capítulo 5 e em profundidade na Seção A.7.1.1<sup>13</sup>.

Uma parte interessada mal intencionada explora o processo de inferência de significado (semiose), para manipular a perspectiva de um observador/grupo e induzir outras partes interessadas à ações prejudiciais, visando obter algum ganho próprio. A desinformação deliberada interfere na semiose de um observador para tentar direcionar a significação dos estímulos percebidos por ele. Na semiose a criação de um signo e a inferência de seu significado ocorre de forma indutiva, dedutiva ou abdutiva. Liu explica, "resumidamente, a abdução cria, a dedução explica, e a indução verifica" (Liu, 2000). A indução constrói regras (*affordances*) com base em fatos observados (Liu, 2000), por exemplo:

- Fato: Tomei os remédios indicados para tratamento precoce e não peguei Covid-19;
- Conclusão: Tomei os remédios e não fiquei doente;
- Regra: O tratamento precoce **funciona**;

Na comunicação de enganação deliberada, a parte mal intencionada busca induzir os indivíduos a chegarem a uma conclusão que ele mesmo construiu previamente (Li e Santos Jr, 2012). O enganador empurra uma conclusão sobre algum aspecto observado para uma audiência. Aspectos do campo de informação social, podem ser explorado por uma parte mal intencionada para induzir a conclusão de uma inferência - como no exemplo do semáforo.

A inferência de significado por dedução encontra consequências lógicas de premissas, tirando conclusões baseadas num fato observado pela aplicação de uma regra (Liu, 2000), por exemplo:

- Regra: Hidroxicloroquina é um remédio antiviral;

---

<sup>13</sup>Spearphishing é um tipo de ataque cibernético onde o atacante coleta dados do alvo e prepara uma isca para ele com base nesses dados.

- Fato: Covid-19 é um vírus;
- Conclusão: Hidroxicloroquina **funciona** contra a Covid-19;

Da mesma forma que a indução, a dedução também é impactada pelo campo de informações. Como a ideia de que a hidroxicloroquina é um antiviral, e só por isso, combate o vírus da Covid-19 (Uol, 2021a). No caso do semáforo, o próprio semáforo e o código de trânsito amparam a decisão do observador de seguir quando a luz verde está acesa.

Uma abdução procura um padrão em um fenômeno e sugere uma hipótese, por exemplo:

- Regra: Hidroxicloroquina é um antiviral;
- Fato: Covid-19 é um vírus;
- Conclusão: Hidroxicloroquina **pode** funcionar contra a Covid-19;

As atividades de inferência sofrem impacto do campo de informação da organização. As partes interessadas, os sub-sistemas de informação (Cebola Organizacional<sup>14</sup>) informal, formal e técnico, com normas, consensos e mecanismos, são aproveitados por quem cria a desinformação deliberada com um propósito bem definido em sua mente. Assim, é plausível considerar que a desinformação deliberada *atua na semiose do observador, como uma tentativa de direcionar ou atrapalhar o processo de inferência de um significado, deturpando o uso de um signo e seu significante, com o propósito de enganar*, por exemplo, se aproveitando das *affordances*. Essa conjectura é discutida na biologia, em comportamentos de animais (El-Hani et al., 2010), e em perspectivas de Humberto Eco sobre a semiose (Eco, 1979).

No ambiente digital, a atividade de enganar através da desinformação deliberada aumenta em complexidade, explorando também a distância entre a audiência e o objeto/estímulo da realidade que é representado pelo signo no meio digital. O meio digital molda o formato da desinformação deliberada e suas propriedades na plataforma tecnológica, o que na comunicação, indica uma representação de outra representação. Dessa forma, os signos no ambiente digital exigem inferências sobre as propriedades que representam, não sendo uma representação direta do que o agente experimenta, mas uma representação de uma representação (recursividade) (McCabe, 1982). Assim, como processo de inferência pode ser explorado de várias formas, por exemplo através dos vieses cognitivos (Aïmeur et al., 2019; Kovic e Laissue, 2016), a comunicação de desinformação digital se transforma em um vetor potencial para manipulação.

Um outro exemplo relacionando a *affordances* (experiências e objetos dinâmicos), semiose, percepção e interpretação, e normas (hábitos e consensos). Mesmo na plataforma tecnológica, se uma regra for codificada (software) para parecer correta, e mostra resultados que parecem corretos, ele pode estar desinformando e manipulando a percepção do usuário por ele ter hábitos e consensos de uso da tecnologia e de programas. Por exemplo, quando você vê seu saldo no aplicativo do banco, você acredita estar vendo uma representação do dinheiro que está na sua conta no banco. No entanto, essa representação é um signo que indica determinado estado para sua percepção, sua *affordance* que virou norma internalizada no seu sistema de informação humano lhe induz a isso. Na realidade, você pode estar sofrendo um ataque de Homeno-no-Meio, caracterizado pela interceptação e manipulação das informações que seu celular envia para o servidor do banco por um ator mal intencionado. Ele pode alterar os dados sem que você ou o banco saibam, fazendo você acreditar que os dados são íntegros e a comunicação é confiável. Ou então, algum funcionário mal intencionado do banco que alterou o sistema, ou ainda, o próprio aplicativo do banco pode parecer legítimo mas se tratar de uma falsificação. Considerando essa formalização de semiose, normas e desinformação deliberada, as possibilidades são inúmeras.

<sup>14</sup>Apend:cebola-semi



Nesse sentido, um artefato tecnológico como um *smartphone* oferece uma representação da realidade que oculta processos digitais de manipulação de informação do seu usuário.

As características sociais de um grupo revelam seus consensos, seus hábitos de julgamento e ação. Estes atributos caracterizam tanto situação de comunicação accidental de desinformação quanto deliberada. No primeiro caso, desentendimentos, ambiguidades e normas sociais diferentes, se relacionam com desinformação accidental (Bessi et al., 2015). No segundo caso, estas características tornam as partes interessadas suscetíveis a desinformação deliberada, sendo exploradas ativamente (Bay, 2018). Assim, crenças, valores, experiências, expectativas, necessidades, e consensos configuram aspectos do contexto de uma organização que são explorados por uma parte interessada mal-intencionada através do uso da desinformação. Nesse sentido, além dos aspectos sociotécnicos relacionados a desinformação accidental, também há aspectos sociotécnicos de influência deliberados usados no enquadramento que a mensagem é construída, como a relação entre o tema e o ambiente da audiência para a qual a desinformação é enviada, a ideologia do grupo (Starbird et al., 2019), as referências a fontes confiáveis o grupo (Doty, 2015), entre outros.

É nesse sentido de suscetibilidade que a tese entende a ideia de "Vulnerabilidades Humanas- como sendo invariantes-comportamentais que podem ser identificadas e exploradas ativamente por um agente mal intencionado para ganho próprio. Elas podem existir no Sistema de Informação Humana, como normas de comportamentos sociais aceitáveis ou padrões de procedimentos que ocorrem informalmente, formalmente e tecnicamente em uma organização. Por exemplo, os procedimentos, protocolos, regras de negócios, normativas éticas, são aspectos que regem como as interações e comunicações devem ser realizadas ou entendidas. Nesse sentido, aspectos como autoridades e valores influenciam no posicionamento e na interpretação de desinformações pela audiência (Jansen van Vuuren et al., 2012). Também existem como *affordances* de avaliação e julgamento, como vieses cognitivos que induzem a significações equivocadas de estímulos percebidos, e podem ser induzidas por aspectos em uma mensagem, como as falácias lógicas. Os vieses cognitivos são diferentes fatores humanos que impactam na interpretação e percepção de uma situação pelos sujeitos envolvidos na situação de comunicação (Nussbaumer et al., 2016). As falácias lógicas são conclusões derivadas de um raciocínio errado que parece verdadeiro (Nuruddin Hidayat et al., 2020). A Dismiss buscou incorporar estes elementos como eixos complementares para sustentar a análise de casos de desinformação, apoiando a identificação de aspectos sociotécnicos de influência na suscetibilidade do consumo de desinformação.

## A.2.2 Uma classificação para estratégias de mitigação da desinformação digital

### A.2.2.1 As estratégias de mitigação

A mitigação envolve as atividades relacionadas ao combate da desinformativo e intervenções para reduzir seus efeitos. As estratégias de mitigação podem ser executadas de forma automática, híbrida ou inteiramente manual pelas partes interessadas (Parikh e Atrey, 2018). A mitigação automática envolve artefatos técnicos que mecanizam ações de controle e responsabilização, como mecanismos de detecção automática, ou rótulos que marcam uma determinada desinformação. A mitigação manual inclui estratégias que apoiam as partes interessadas na tarefa de identificar uma desinformação, como soluções integradas ao browser que facilitam a verificação de um mensagem (Pourghomi et al., 2017). As estratégias híbridas unem os aspectos técnicos e as funções do sistema de informações humanas para aprimorar a eficiência e eficácia de detecções e correções, como a detecção automática de mensagens suspeitas para serem verificadas por agentes humanos (Volkova et al., 2017).



Como mostrado no capítulo do mapeamento, não encontramos uma categorização para as estratégias de mitigação de desinformação. Nesse sentido, esta tese entende que a mitigação da desinformação deliberada digital é uma atividade relacionada com a segurança da informação. Assim, as estratégias de mitigação da desinformação foram caracterizadas com base na taxonomia de contramedidas de segurança da informação (Avizienis et al., 2004).

As atividades de mitigação são agrupadas em quatro categorias de contramedidas: prevenção, previsão, remoção e tolerância. As estratégias de Prevenção objetivam evitar a "contaminação" das partes interessadas por objetos desinformativos, prevenindo seu consumo e criação, e barrando sua divulgação inicial. As estratégias de Previsão intendem estimar a existência de aspectos que configuram vulnerabilidades passíveis de exploração pelo uso de desinformação deliberada, como a identificação de tendências, a avaliação de limitações em protocolos de operação, e análises de suscetibilidade (Wilder e Vorobeychik, 2018). Ademais, estes aspectos também indicam fragilidades que podem levar a ocorrências acidentais de desinformação. As estratégias de Remoção consistem na propagação de correções, na correção de vulnerabilidades conhecidas em uma organização, como atualização de algoritmos de detecção e treinamento das partes interessadas. As estratégias de Tolerância almejam a resiliência da organização na presença de desinformação, como o uso de mecanismos de detecção automática, e campanhas de conscientização e treinamentos. O Quadro A.3 apresenta exemplos das estratégias e atividades de mitigação em cada categoria.

Tabela A.3: Estratégias de mitigação de desinformação digital.

| Tipo       | Descrição                                                                                                                                                                      | Atividades                                                              | Exemplos                                                                                                                                                                        |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Prevenção  | Evitar a contaminação da organização por objetos desinformativos conhecidos, prevenindo seu consumo e criação, e barrando sua divulgação inicial;                              | Deteção; Bloqueio; Treinamento; Conscientização; Normatização;          | Deteção de spam; deteção de fraudes <i>phishing</i> ; treinamento contra engenharia social; conscientização sobre consequências; Legislação para coibir fraude, difamação, etc. |
| Previsão   | Intendem estimar a existência de vulnerabilidades passíveis de exploração com desinformação em uma organização e limitações que possam ocasionar desinformação na comunicação; | Avaliação de risco; Identificação de vulnerabilidades; Análise forense; | Modelagem de comportamento, avaliação de suscetibilidade, avaliação de confiabilidade e credibilidade;                                                                          |
| Remoção    | Consistem em corrigir desinformações e mitigar vulnerabilidades conhecidas que possam ocasionar desinformação na comunicação;                                                  | Manutenção; Treinamento; Conscientização;                               | Atualização de detectores automáticos, remoção de objetos desinformativos, modelagem de objetos desinformativos, identificação de padrões em interações, censura de contas;     |
| Tolerância | Almejam a resiliência da organização na presença de desinformação;                                                                                                             | Deteção, Bloqueio, Descontaminação; Correção; Treinamento; Normatização | Monitoramento de difusão, identificação de padrões em conteúdo, controle de difusão, treinamento e conscientização, etc;                                                        |

A mitigação por deteção engloba a coleta objetos desinformativos (Dang et al., 2016b), reconhecimento de padrões (Wai et al., 2018), verificação de veracidade (Devi e Karthika, 2018), uso de ferramentas digitais para auxiliar na identificação de objetos desinformativos,

como anotações visuais nas mensagens (Gao et al., 2018; Wood et al., 2018) e mecanismos manuais de detecção (Pourghomi et al., 2017), e ferramentas automáticas de detecção de objetos desinformativos e contas (Wilder e Vorobeychik, 2018). O bloqueio engloba o corte de links (Ruan et al., 2015), bloqueio de contas (Khaled et al., 2018), remoção de objetos desinformativos (Krishnamurthy e Hamdi, 2013), controle de difusão, limitações de compartilhamento (Melo et al., 2019) etc.

Treinamentos e conscientização incluem o incentivo de comportamentos responsáveis (Wang et al., 2014), e de verificação de conteúdo (Torres et al., 2018a), aprimorar a percepção de risco das partes interessadas (Almaliki, 2019a; Caputo et al., 2013), e o ensino de técnicas de verificação (Ireland, 2018). Atividades de avaliação de riscos consideram a descoberta de suscetibilidade (Jansen van Vuuren et al., 2012), a inferência de credibilidade e confiabilidade de conteúdo e das origens (Dongo et al., 2019; Cota et al., 2019), a previsão de rumores futuros (Qin et al., 2018), simulações de eventos desinformativos (Bossetta, 2018), modelos de comunicação de desinformação (Traylor et al., 2017; Zhou e Zhang, 2007) etc. A mitigação também trata de atividades formais como legislação, normas e estatutos. Informações falsas no ciberespaço requerem novas leis (Rowe e Rothstein, 2004; Rowe, 2006), e instrumentos legislativos para controle e responsabilização (Watney, 2018).

A manutenção inclui a criação e atualização de bases de dados (Rubin e Conroy, 2012), o aprimoramento de mecanismos de detecção automática (Kapusta e Obonya, 2020; Poddar et al., 2019), modelos de objetos desinformativos (Cho et al., 2017; Mack et al., 2007), identificação de motivações de compartilhamento (Wason et al., 2019; Pal et al., 2018; Chen e Sin, 2013; Zhou e Zhang, 2007).

As correções de desinformações podem ser feitas por qualquer parte interessada, de organizações especializadas como checadores de fatos, a grupos amadores, até *bots* em redes sociais. Para maior efetividade deste tipo de mitigação, é importante que as correções sejam divulgadas no meio de comunicação que o usuário está habituado a usar (Fernandez e Alani, 2018). Além das correções, divulgar as motivações da desinformação, as suas consequências, e origens, impactam no consumo de correções pelas partes interessadas (Fernandez e Alani, 2018). Ainda, outro aspecto importante neste tipo de mitigação é a relevância da parte interessada (autoridade, influenciadores, especialistas, etc.) que divulga as correções para o grupo (Fernandez e Alani, 2018).

### A.3 APÊNDICE 3 - PROTOCOLO DO MAPEAMENTO SISTEMÁTICO

O propósito desse mapeamento foi obter um panorama abrangente das pesquisas sobre a desinformação digital. Dessa forma, o objetivo foi definido de acordo com o paradigma *Goal-Question-Metric* (GQM) (Basili e Rombach, 1988):

*Analisar a literatura científica sobre o fenômeno da desinformação digital para identificar e categorizar os desafios e oportunidades de pesquisa pela perspectiva da área de Ciências da Computação no contexto das bibliotecas digitais da ACM, Scopus e IEEE.*

Com relação ao objetivo, foram definidas as seguintes Questões de Pesquisa (QP):

Tabela A.4: Questões de pesquisa

| QP     | Descrição                                                              | Objetivo                                                                                                                                                                    |
|--------|------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| QP01.a | O que a literatura em Ciência da Computação define como desinformação? | Identificar os conceitos para os termos que referenciam desinformação digital e investigar os problemas no entendimento da taxonomia.                                       |
| QP01.b | Quais são os tipos de desinformação estudados e os temas?              | Identificar os tipos de desinformação digital e os seus atributos (sátira, notícia falsa, fraude, etc.).                                                                    |
| QP02   | Quais e como são abordadas as partes interessadas?                     | Identificar as partes interessadas abordadas pelos estudos e a metodologia usada para caracterizá-las.                                                                      |
| QP03   | Como as pesquisas abordam os aspectos sociotécnicos?                   | Identificar metodologias e a abrangência multidisciplinar dos modelos e caracterizações considerados pelas soluções, e como os aspectos sociotécnicos são abordados.        |
| QP04   | Como as pesquisas abordam o ciclo de vida da desinformação?            | Categorizar os estágios do ciclo de vida da desinformação considerados pelos pesquisadores sobre o fenômeno da informação falsa para identificar oportunidades de pesquisa. |

O Quadro A.5 exibe a chave de pesquisa. A chave foi definida após uma busca informal pelos termos "misinformation", "disinformation", "fake news", "false information" e "rumor" nas bases da IEEE, ACM e SCOPUS. As palavras-chave das revisões e mapeamentos encontrados foram incorporadas na chave de pesquisa objetivando maior abrangência do mapeamento. O termo "rumor" foi considerado por ter definições que variam de informações não verificadas (Tan et al., 2019; Jin et al., 2017) até desinformação deliberada (Zhu e Liu, 2018; Tong et al., 2017).

Tabela A.5: Chave-base de pesquisa

("disinformation" OR "misinformation" OR "fake news" OR "false news"  
OR "fabricated news" OR "hoax" OR "rumor" OR "false information"  
OR "fake information" OR "fabricated information")

O processo de seleção foi realizado por uma equipe de 3 pesquisadores<sup>15</sup> do Laboratório de Pesquisa em Interação Humano-Computador da UFPR. A metodologia adotada para a seleção dos artigos foi, com um processo de filtragem dividido em três etapas e considerando critérios de inclusão e exclusão. O Quadro A.6 apresenta os critérios de seleção e exclusão<sup>16</sup>. Os artigos foram divididos em três conjuntos de tamanhos iguais designados para cada pesquisador. Os pesquisadores analisaram individualmente seus conjuntos. Ao final de cada etapa cada um revisou o trabalho do outro e discutiram os resultados em busca de consenso. Os dados foram extraídos empregando os formulários de extração apresentado nos Apêndices A.7. No primeiro filtro, foram revisados o título e o resumo de todos os artigos. Em seguida, as seções de introdução e conclusão dos artigos selecionados foram revisados. A terceira e última etapa consistiu na leitura completa dos artigos selecionados para extração de dados.

<sup>15</sup>1 estudante de graduação, 1 mestrando e 1 doutorando.

<sup>16</sup>Os critérios CI01 e CI02 foram definidos com base nas revisões encontradas na busca informal (julho 2018).

Tabela A.6: Critérios para seleção e exclusão dos artigos

| Inclusão | Exclusão | Descrição                                                                                                                                                      |
|----------|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CI01     |          | Publicações científicas em inglês que apresentem soluções computacionais ou tecnológicas, experimentos e avaliações sobre o fenômeno da desinformação digital. |
| CI02     |          | Publicações científicas em inglês que apresentem métodos de investigação, modelagens e caracterizações do fenômeno da desinformação digital.                   |
| CI03     |          | Publicações científicas em inglês que apresentem outros tipos de estudos relacionados ao fenômeno de desinformação.                                            |
|          | CE01     | Campo: Outro que não Ciência da Computação                                                                                                                     |
|          | CE02     | O artigo é duplicado;                                                                                                                                          |
|          | CE03     | Literatura cinzenta: exceto periódicos, revistas, conferências, artigos e livros.                                                                              |
|          | CE04     | Não acessível no âmbito do Portal de Publicações da Capes.                                                                                                     |
|          | CE05     | Mais antigo que 1993 <sup>17</sup> .                                                                                                                           |
|          | CE06     | Não relacionado a desinformação digital ou algum de seus tipos.                                                                                                |

Os dados extraídos para cada questão de pesquisa foram analisados usando o método de análise de conteúdo (Lazar et al., 2017) juntamente com o Arcabouço Semiótico (AS)<sup>18</sup>. O AS divide os aspectos do sistema de informação em subdisciplinas e foi usado para categorizar os estudos em relação à multidisciplinaridade do fenômeno da desinformação. Os artigos foram categorizado com base no seguinte esquema:

- Social: Abordagens que buscam compreender os aspectos sociais do fenômeno, como valores, a cultura, os impactos da desinformação na sociedade, e efeitos de vieses cognitivos de grupo (efeito manada, etc.);
- Pragmático: Abordagens que buscam compreender as funções interpretativas do ato de comunicação, como as motivações de compartilhamento, aspectos de deliberação, difamação, infâmia, e a finalidade dos tipo de desinformação, etc;
- Semântico: Abordagens que buscam compreender aspectos do significado/entendimento das informações, como a subjetividade da veracidade, análise sentimental, análise de posicionamento, ambiguidades e a relação do tema com o ambiente;
- Sintático: Abordagens que buscam compreender aspectos estruturais do fenômeno (conteúdo, difusão, organização, etc;), como a diagramação, elementos de ênfase, agrupamentos em redes virtuais, identificação de gargalos na difusão;
- Empírico: Abordagens que buscam compreender aspectos da interação e dos meios comunicação, como as *affordances*<sup>19</sup> tecnológicas, sincronicidade, direcionalidade, tempo de leitura, ruído, interferência, entre outros;
- Físico: Abordagens que buscam compreender aspectos físicos do fenômeno, como limitações de hardware, o custos de difusão, defeitos em dispositivos, e localização geográfica dos participantes;

Os dados foram coletados utilizando o formulário de extração apresentado no Quadro A.7. Para cada publicação selecionada foram extraídas as seguintes informações:

- Identificação de artigo

Esta é uma string de identificação única do artigo.

<sup>17</sup>O ano de nascimento do NCSA Mosaic que impulsionou a popularização da Web (Schatz e Hardin, 1994).

<sup>18</sup>Artefato da Semiótica Organizacional usado na análise de Sistemas de Informação (Stamper e Liu, 1994).

<sup>19</sup>O termo "affordances" significa "aquilo que, na interação com o sujeito, o ambiente potencialmente proporciona, oferece ou convida a agir em razão das propriedades ou características de reconhecimento que possui, ou deveria possuir, para satisfazer certas necessidades de uso" (McCabe, 1982).

Tabela A.7: Formulário de extração de dados

| Dado                             | Valor                                               | RQ  |
|----------------------------------|-----------------------------------------------------|-----|
| Id da Pesquisa                   | Chave que identifica uma pesquisa                   | *   |
| Definição da informação falsa    | Definição do tipo de informação falsa               | QP1 |
| Classes de informação falsa      | Tipos (desinformação, notícia falsa, sátira, etc.); | QP2 |
| Propriedades da informação falsa | Estrutura, tamanho, composição, efeitos, etc;       | QP2 |
| Problema                         | Problema investigado                                | QP3 |
| Classificação do problema        | Classificação na Escada Semiótica                   | QP3 |
| Tipo de Solução                  | Ferramenta, treinamento, automática, etc.           | QP3 |
| Veiculo de publicação            | Evento, revista, etc.                               | *   |
| Autores                          | Nome e grupos de pesquisa                           | *   |
| Universidade                     | Nome da Universidade                                | *   |

- Definições de informações falsas

Identificar as definições dos tipos de informações falsas considerados nos estudos. O campo deve ser preenchido com a definição de desinformação, a referência para definição ou nenhum (sem definição direta).

- Tipos de informação falsa

Identificar as categorias e tipos de desinformação abordados nas pesquisas. Tipos de desinformação, como notícias falsas, sátiras, teorias da conspiração, críticas falsas, etc.

- Propriedades de informações falsas

Identificar as principais características dos tipos de desinformação. Ex: recursos de aparência, estrutura de texto, conteúdo de texto, recursos de rede, recursos de engano, recursos de usuário, recursos de engano: semelhança de URL para enganar ou uso de perfis falsos; Recursos de rede: Capacidade de propagação como velocidade, alcance, duração, etc; Características do texto, como tamanho da mensagem, quantidade de informações sobre o título, qualidade, etc; Conteúdo do texto como construção linguística, assuntos interessantes, uso de elucidação, uso de linguagem cultural, etc; Características do usuário: autenticidade, intenção, credibilidade, etc.

- Problemas

Identificar os problemas investigados e seus desafios.

- Classificações de problemas

Categorizar os aspectos sociotécnicos dos problemas com a Escada Semiótica.

- Tipo de Solução

Identificar categorias e tipos de soluções que estão sendo geradas. Ex: estratégias de mitigação de informações falsas, detecção de contas falsas, validação de conteúdo, etc.

- Veículo de publicação

Identificar os veículos de publicações das pesquisas.

- Autores

Mapear os grupos de pesquisa e autores interessados na área da desinformação.

- Universidade

Mapear as universidades que desenvolvem estudos na área da desinformação.

### A.3.1 Resultados e Discussões

A chave de pesquisa retornou 4591 artigos. Destes, foram selecionados 788 estudos para responder as QP's. Os dados extraídos foram organizados em planilhas do Google Sheets para as fases de filtragem<sup>20</sup> e a extração de dados<sup>21</sup>. A codificação para filtragem considerou as camadas do Arcabouço Semiótico, apresentado em detalhes nas planilhas de extração de dados. A Tabela A.8 resume os resultados do processo de seleção.

Tabela A.8: Estatísticas do mapeamento sistemático

| Estatísticas                              | Processado | Revisão | Total      |
|-------------------------------------------|------------|---------|------------|
| Todas as artigos                          |            |         | 4591       |
| 1º filtro excluído (duplicados incluídos) | -1776      | -1733   | -3059      |
| 2º filtro excluído (duplicados incluídos) | -275       | -19     | -294       |
| Dados extraídos                           |            |         | <b>788</b> |

### A.3.2 Relatório: Mapeamentos e Revisões sistemáticas da literatura

Na Ciência da Computação, encontramos 39 revisões de literatura de 2011 a 2020 relacionadas a pelo menos uma forma de informação falsa<sup>22</sup>. O mapeamento revelou uma revisão da literatura sobre os avanços sociotécnicos do fenômeno (Fernandez e Alani, 2018). O Quadro A.9 apresenta as revisões encontradas.

Fernandez e Alani (2018) apresentam o estado da arte de soluções sociotécnicas para combater a desinformação digital. Os resultados apontam obstáculos para cada grande área. Trabalhos sobre *detecção* fornecem alertas sem justificativa, geralmente desvinculados do contexto do usuário e seu papel na comunicação (consumidores ativos, co-criadores, etc.). Além disso, soluções de detecção podem ser potencialmente enganadas, em especial por desinformação deliberada. Trabalhos sobre dinâmica, falham ao não analisar a topologia da rede em profundidade, além de não considerarem o comportamento do usuário, somente dados demográficos (idade, sexo, localização geográfica) em seus modelos.

Nos trabalho de validação, os autores citam a dificuldade de checadores de fatos e abordagens baseadas em multidões de lidar com o volume de dados crescente. Além disso, as correções acontecem desconectadas do local onde o usuário consome informações, sendo pouco acessadas ou difundidas. Por fim, correções dificilmente conseguem sozinhas mudar as crenças

<sup>20</sup>[https://docs.google.com/spreadsheets/d/1QqHWVMC\\_CMMwLePvYGf8miM7NlM-CxYz6Z1ttDi494Q/](https://docs.google.com/spreadsheets/d/1QqHWVMC_CMMwLePvYGf8miM7NlM-CxYz6Z1ttDi494Q/)

<sup>21</sup><https://docs.google.com/spreadsheets/d/1lZORTRirkr-kKYzZXC9gV8lip3d3lfsVG7X8nY7RLbg/>

<sup>22</sup>Retornar para o estado da arte Tabela 3.1



Tabela A.9: Revisões Sistemáticas sobre desinformação

| Tópico                               | Descrição                                                                                                                                                        | Artigos.                                                                                                                                                                                                                                                                                          |
|--------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Avaliação da qualidade da informação | Revisões sobre métodos, técnicas, critérios e métricas para avaliação da qualidade de conteúdos de modo automático, manual e híbrido.                            | (Pal e Chua, 2016; Arroyo et al., 2018; Gupta e Han, 2011; Chomphoosang et al., 2012; Viviani e Pasi, 2017; Devi e Karthika, 2018; Lozano et al., 2020; Kaliyar e Singh, 2019; Habib et al., 2019)                                                                                                |
| Detecção de desinformação            | Revisões sobre métodos, técnicas, critérios e métricas para identificação de indícios da veracidade de conteúdos de modo automático, manual e híbrido.           | (Yuliani et al., 2018; Assiroj et al., 2018; Figueira e Oliveira, 2017; Özgöbek e Gulla, 2017; Shu et al., 2017; Ahmed et al., 2019; Gaonkar et al., 2019; Bondielli e Marcelloni, 2019; Devi et al., 2019a; Sharma e Sharma, 2019; Saquete et al., 2020; Devi et al., 2019b; Chen et al., 2015c) |
| Detecção de rumores                  | Revisões sobre métodos, técnicas, critérios e métricas para identificação de indícios de informações não validadas de modo automático, manual e híbrido.         | (Al-Sarem et al., 2019; Ahsan et al., 2019; Reuter et al., 2017)                                                                                                                                                                                                                                  |
| Estratégias de enganação             | Revisões sobre métodos e técnicas para identificação de indícios de tentativas de manipulação e persuasão de modo automático, manual e híbrido.                  | (Ahmad et al., 2019; Varol e Uluturk, 2018);                                                                                                                                                                                                                                                      |
| Estado da arte                       | Revisões sobre os avanços da academia no entendimento do fenômeno, desafios e oportunidades de pesquisa.                                                         | (Zhang e Ghorbani, 2020; Fraga-Lamas e Fernández-Caramés, 2020; Pierri e Ceri, 2019; Figueira et al., 2019; Sharma et al., 2019; Fard e Cunningham, 2019; Meel e Vishwakarma, 2019; Fernandez e Alani, 2018)                                                                                      |
| Controle de difusão                  | Revisões sobre métodos, técnicas, critérios, métricas e modelagens relacionadas ao comportamento de difusão de desinformações para identificação e rastreamento. | (Rana et al., 2019; Almaliki, 2019b; Shelke e Attar, 2019)                                                                                                                                                                                                                                        |
| Detecção em multimídia               | Revisões sobre métodos e técnicas para identificação de indícios relacionados a desinformação em conteúdos multimídia.                                           | (Albahar e Almalki, 2019);                                                                                                                                                                                                                                                                        |

dos usuários, sendo importante considerar não apenas quais informações corretivas devem ser fornecidas, mas quando, como e a quem devem ser fornecidas. Sobre os trabalhos relacionados ao gerenciamento, os autores afirmam que as estratégias não vão além da geração de fatos e da detecção precoce de contas maliciosas, e tendem a se concentrar nos aspectos técnicos e não nos humanos do problema. Os autores afirmam a necessidade de mais abordagens sociotécnicas para aprimorar a capacidade de mitigação, sobretudo relacionadas ao contexto e as relações humanas.

Fard e Cunningham (2019) investigam informações falsas em uma extensa revisão da literatura. Os autores usaram um método estatístico para avaliar 21.000 publicações no banco de dados da *Web of Science* relacionadas a informações falsas e não verificadas de 1900 a 2019 em todas as áreas de pesquisa disponíveis. Seus resultados discutem as tendências de crescimento, coerência e impacto, que estão aumentando, e a tendência de diminuir as novidades sobre soluções. Além disso, os autores discutem a falta de consenso taxonômico, com conceitos controversos para os mesmos termos. Ainda, o relatório o mapeamento aponta uma comunidade científica pouco organizada, com conferência esparsas, sem revistas dedicadas ao tema. Os autores concluem seu trabalho afirmando a necessidade de uma comunidade científica mais organizada e ativa, propondo uma estratégia para desenvolver um campo de pesquisa independente. Em um escopo semelhante, Pal e Chua (2016) revisaram os objetivos e métodos de pesquisa que estudam as ameaças à qualidade do conteúdo gerado pelo usuário por uma variedade de disciplinas além da ciência da computação (e.g. engenharia Civil, economia e medicina). Eles analisaram 453

publicações até o ano de 2016. Os autores propõe uma categorização com 5 grupos: detecção, dinâmica (difusão), detecção da fonte, percepção dos usuários e mitigação. Os resultados revelam poucos esforços na categoria de mitigação, apontam a dificuldade de parar a criação de rumores pela falta de curadoria, e afirmam a necessidade de mais estudos multidisciplinares e estudos qualitativos (dos 453 artigos, apenas 90 usavam uma metodologia de análise qualitativa).

Al-Sarem et al. (2019) apresentam uma revisão sistemática da literatura para detecção de rumores usando abordagens de redes neurais profundas. Os autores analisaram 25 artigos que passaram pelos critérios de inclusão. Os resultados mostram a caracterização de bases de dados, um conjunto de propriedades empregadas na avaliação automática da veracidade de rumores, e os desafios para detecção automática de desinformação. Os autores ressaltam a necessidade de pesquisas na qualidade técnica das bases dados, como redução do viés e a representatividade, foram indicados como necessárias. Assiroj et al. (2018) revisam a literatura sobre detecção automática de farsas (*hoaxes*) em mídias sociais. Os autores selecionaram 38 artigos publicados entre 2016 e 2018 buscando identificar as técnicas mais usadas e as redes sociais estudadas. Os resultados apontam técnicas de *support vector machine*, como as mais promissoras e a rede social Twitter com mais estudos. Não foram apresentados problemas ou oportunidades de investigação futuras.

Almaliki (2019b) revisou 101 estudos primários sobre a desinformação na computação entre os anos de 2011 e 2018. O autor classifica os trabalhos de acordo com o tipo, contribuição e foco de pesquisa. Quatro grandes categorias englobam todos os tipos de pesquisa analisados: Detecção, dinâmica (difusão), validação (checagem) e gerenciamento (mitigação). O autor relata que a maioria dos estudos focam na detecção automática (60%), e que os estudos cresceram em número a partir de 2016. Ademais, o autor afirma a necessidade de estudos que apresentem novas ideias de mitigação, considerando o contexto dos usuários e fomentando sua colaboração. O autor apura a necessidade de conferências dedicadas ao tema.

Habib et al. (2019) revisam a literatura sobre detecção automática de diferentes tipos de informações falsas. Os resultados apontam a necessidade de estender o entendimento sobre os tipos de desinformações, com caracterizações e pesquisas qualitativas sobre o impacto delas nos usuários. Ainda, sendo os modelos de deep learning os mais promissores, os autores apontam a necessidade de ampliar os estudos, desenvolvendo conjuntos de dados de temáticas diversas em saúde, educação e política. Por fim, há muito poucas pesquisas sobre a detecção de farsas deliberadas, sendo uma área de pesquisa que carece de atenção.

Lozano et al. (2020) considera 107 trabalhos (entre 2013 e 2017) que usam técnicas de avaliação automática de veracidade para o conteúdo em mídias sociais. Seus achados relatam a maior parte dos artigos usando técnicas ultrapassadas para avaliação automática, e critérios individuais e pouco representativos do contextos dos problemas. Os autores criticam a falta de consenso do termo "análise de veracidade", a qualidade dos artigos excluídos da análise (maior parte dos artigos carece de detalhes metodológicos), e a restrição de acesso aos dados das pesquisas, o que dificulta a reprodutibilidade dos resultados.

Reuter et al. (2017) revisam a literatura atrás de abordagens e critérios para a avaliação da credibilidade das informações geradas pelo usuário. Os resultados mostram um conjunto de critérios para avaliação, como fontes, popularidade e alcance. Os autores apontam oportunidades de pesquisa em plataformas de comunicação diferentes, com abordagens e critérios para avaliação cruzando dados de mídias diferentes. Os autores destacam obstáculos a serem superados, como encontrar critérios de credibilidade para indivíduos, que são dependentes do contexto, do proposito, das pessoas e dos papeis desempenhados por elas.

Saquete et al. (2020) realizaram uma revisão da literatura sobre detecção automática de notícias falsas usando processamento de linguagem natural. Os resultados reúnem um compilado

de características para detecção de notícias falsas de modo precoce e tardio. Os autores destacam a falta de métricas confiáveis, de bases de dados balanceadas, a necessidade de desenvolvimento de modelos detalhados de notícias falsas, a incorporação de padrões e pistas linguísticas, a seleção de tecnologias adequadas adaptadas a cenários particulares e, a elaboração de técnicas de avaliação de emoções associadas. Bondielli e Marcelloni (2019) resumem a literatura que investiga a detecção automática de notícias falsas e rumores. Os autores apresentam um compilado de características para detecção de notícias falsas e rumores com base no conteúdo e no contexto. Os resultados mostram a necessidade de aprimoramento das técnicas e bases de dados, além de um destaque para a eficiências das técnicas baseadas em *deep learning*.

### A.3.2.1 Discussão

De modo geral a literatura classifica as iniciativas para lidar com a desinformação em quatro linhas de pesquisa: detecção, dinâmica, validação e gerenciamento. As pesquisas mais comuns são relacionadas a detecção, apresentando um foco em aspectos técnicos e buscando aprimorar a eficácia de estratégias de detecção automática. Existe a necessidade de mais trabalhos multidisciplinares capazes de lidar com a complexidade do fenômeno, sobretudo abordagens que considerem aspectos técnicos, humanos, sociais e organizacionais de maneira integrada.

Encontramos revisões sistemáticas sobrepostas no mesmo tempo e escopo, com poucas diferenças. Por exemplo, os trabalhos de detecção de notícias falsas consideram praticamente o mesmo período de tempo em que encontraram os mesmos documentos relevantes (Gaonkar et al., 2019; Sharma et al., 2019). Uma possível causa desse problema pode ser a falta de revistas e conferências dedicadas ao tema, o que pode ocasionar dificuldades de acesso a trabalhos relevantes por pesquisadores em diferentes países. Ainda, os resultados sobre os problemas com a taxonomia da desinformação e a dificuldade para determinar termos representativos para os trabalhos (Lozano et al., 2020; Fard e Cunningham, 2019), foram recorrentes em diferentes mapeamentos. Os trabalhos apresentam um escopo restrito as áreas específicas de interesse dos autores, com pouca exploração de linhas de pesquisa diferentes de detecção automática ou difusão da desinformação. Estas revisões representam um ponto de partida para pesquisadores interessados nessas áreas, mas não uma visão abrangente e detalhada do campo de pesquisa.

Além disso, a falta de conferências e revistas dedicadas ao estudo da desinformação limita a capacidade de colaboração e o desenvolvimento de uma estrutura acadêmica capaz de interagir e desenvolver um campo de pesquisa *per se* (Fard e Cunningham, 2019). Assim, acreditamos ser necessária uma comunidade de pesquisa mais organizada, capaz de lidar com os desafios complexos do fenômeno de forma efetiva, eficiente, rigorosa e responsável, que busque a integração do conhecimento, que unifiquem taxonomias, discutam metodologias para investigação multidisciplinar e fomentem a colaboração.

Os trabalhos relacionados mostram a natureza complexa do fenômeno da desinformação digital, representada pela diversidade de fatores em diferentes níveis de abstração que são abordados pela literatura. A análise revelou uma predominância de estudos que buscam elaborar soluções para um contexto específico, com uma tendência em adotar uma perspectiva técnica do fenômeno, abordando a detecção e validação de desinformação em sua maioria. Parece haver um desafio metodológico para abordagens abrangentes do fenômeno que limita o avanço do conhecimento a respeito de sua natureza fundamentalmente sociotécnica e impacta na capacidade de resolução de problemas das intervenções propostas. Não encontramos nas revisões abordagens que instrumentalizam a análise de desinformações e apoiem o entendimento de diferentes elementos a serem considerados ao se pensar em intervenções. Assim, acreditamos que antes de desenvolver soluções especializadas é preciso entender os elementos que constituem o problema

tratado, reconhecer os fatores relevantes em cada contexto e organizar o conhecimento gerado para habilitar o desenvolvimento de intervenções efetivas e eficientes.

Parece haver a necessidade de investigações que entendam como ocorre a desinformação pela perspectiva de diferentes partes interessadas, olhando aspectos subjetivos do fenômeno de modo sistêmico. A tendência reportada nas revisões em abordar o fenômeno de modo fragmentado leva a perda de detalhes inter-relacionados, e ao desenvolvimento de intervenções desintegradas, segmentadas em diferentes dimensões do problema complexo, sem sinergia. Por exemplo, quem constrói uma desinformação deliberada constrói ela para um público-alvo, indicando uma relação entre a criação e o consumo de uma desinformação. Nesse sentido, parece haver a necessidade de compreender a natureza sistêmica do fenômeno da desinformação para investigar intervenções e integrar estratégias em diferentes estágios considerando diferentes níveis de abstração. Ainda, não encontramos uma visão geral ou categorizações para os tipos de estratégias de mitigação, ou como, quando e onde são utilizadas. Isto pode facilitar a integração de estratégias em diferentes níveis de abstração, o que possibilitaria melhores condições de mitigar a desinformação e seus efeitos prejudiciais.

Como as revisões reportam, divergências nos conceitos para o termo "desinformação" prejudicam a obtenção de um panorama conciso das contribuições e desafios de pesquisa (Lozano et al., 2020). Além disso, a maioria das revisões e mapeamentos focam em tarefas específicas para mitigação (como a detecção e validação) ou consideram um escopo temporal que limita a abrangência dos trabalhos selecionados e restringem o entendimento dos avanços no conhecimento sobre o fenômeno. Por isso, é necessário realizar um mapeamento abrangente da literatura para obter uma visão complementar das contribuições e desafios do fenômeno na área de Ciências de Computação, considerando as limitações das revisões analisadas. Por este motivo, realizamos um mapeamento sistemático da literatura para compreender em profundidade como aspectos sociotécnicos são abordados pelas pesquisas, quais são os tipos de desinformação digital estudados e suas características, e como as partes interessadas têm sido abordadas<sup>23</sup>.

### A.3.3 Relatório: QP01.a: O que a literatura define como *misinformation* e *disinformation*?

Um dos primeiros desafios encontrados para a caracterização das pesquisas da Ciência da Computação sobre desinformação digital é a falta de consenso taxonômico, como reportado por outros estudos (Fard e Cunningham, 2019; Jiang e Wilson, 2018)<sup>24</sup>. Uma taxonomia define conceitos e categorias do repertório de termos usados por um grupo organizado em prol de estruturar as atividades cooperativas (Van Heusden e Jorna, 2001). O consenso taxonômico é essencial para a prosperidade de atividades organizadas (Van Heusden e Jorna, 2001), como a pesquisa e elaboração de soluções para mitigar as consequências prejudiciais da desinformação.

De 788 artigos selecionados em nossa revisão, 64 estudos definiram o que entendiam por *Misinformation* e/ou *Disinformation*. Um total de 486 artigos indicavam de forma implícita o seu entendimento sobre o tipo de informação falsa com que trabalhavam, referenciando outros estudos. Outros 238 artigos expressaram explicitamente uma definição para pelo menos uma forma de informação falsa, como sátiras e notícias falsificadas. Os resultados do mapeamento mostram a inconsistência na terminologia, com ambiguidade nas definições, o que atrapalha o entendimento do objeto de estudo dos artigos, e torna a revisão da literatura morosa e improdutiva.

<sup>23</sup>O protocolo do mapeamento pode ser encontrado nos Apêndices A.3.

<sup>24</sup>Voltar ao Cap. de estado da arte: Início 3 ou QP.02 3.1.0.1

### A.3.3.1 Taxonomias

O mapeamento encontrou 6 artigos dedicados ao estudo da taxonomia e tipologia de informações falsas (Fard e Cunningham, 2019; Lemieux e Smith, 2018; Shu et al., 2017; Kumar et al., 2016b; Zhou e Zhang, 2004; Tudjman e Mikelic, 2003). Estas pesquisas propõe uma diversidade de critérios para categorizar os tipos de desinformação. Ao todo encontramos 7 critérios usados para categorizar a desinformação que são apresentados no Quadro A.10.

Tabela A.10: Critérios de categorização encontrados na literatura.

| Critério         | Descrição                                                        |
|------------------|------------------------------------------------------------------|
| Evidência        | se é ou não baseado em evidências                                |
| Significância    | é sobre o tópico de alguma urgência ou a importância percebida   |
| Sobre indivíduos | se é sobre uma pessoa ou não                                     |
| Intenção         | propósito da desinformação, sendo com ou sem intenção de enganar |
| Veracidade       | se é verdadeira ou falsa                                         |
| Função           | o objetivo da mensagem (magoar, explicar, manipular, etc.)       |
| Motivação        | ganho monetário, pessoal, político, ou outro tipo de ganho;      |

Os critérios mais usados nos estudos mapeados são (artigos podem ter considerado mais de um critério): evidência (352), intencionalidade (263) ganho/motivação (41). A Figura A.5 apresenta os critérios, tipos e exemplos. A categorização de desinformação quanto a Evidência diferencia a desinformação em dois tipos, aquelas com base em fatos e aqueles baseados em opiniões. Estas são avaliadas com base na qualidade da mensagem em relação a atributos como veracidade (Kumar e Geethakumari, 2013), autenticidade (Shu et al., 2018), e confiabilidade (Fard e Cunningham, 2019) e significância (Fard e Cunningham, 2019). Quanto ao critério Intencionalidade, a desinformação pode ser acidental (*misinformation*)- quando oriunda de erro - ou deliberada (*disinformation*) - construída/usada com a intenção de enganar (Kumar e Geethakumari, 2013; Shu et al., 2018). Quanto ao Ganho, a literatura reporta o lucro financeiro (Shu et al., 2019; Buchanan e Benson, 2019; Seo et al., 2019), ganho político (Moreno e Bressan, 2019; Shu et al., 2019; Varol e Uluturk, 2018), e ganho pessoal (Buchanan e Benson, 2019; Seo et al., 2019).

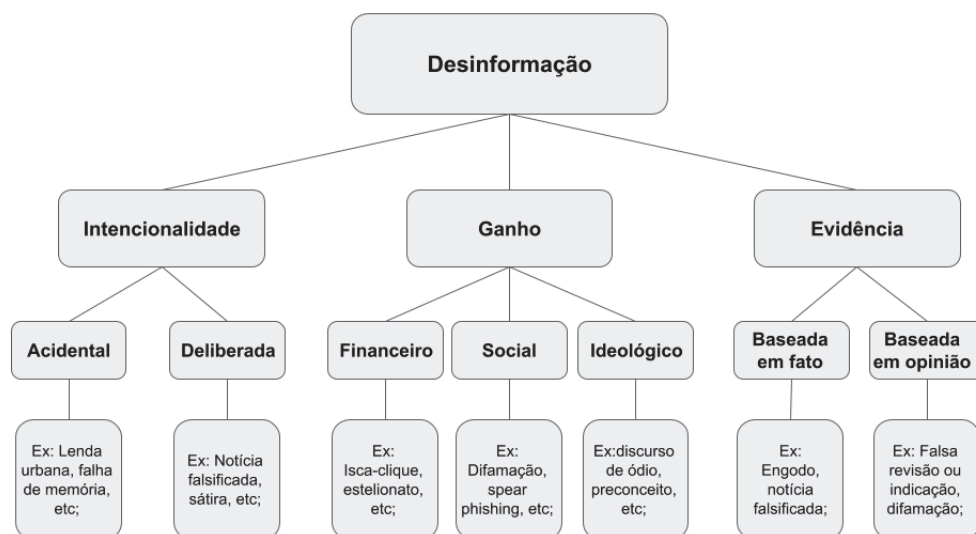


Figura A.5: Principais critérios para categorização de desinformação encontrados na literatura revisada.



Dentre as pesquisas que propõe taxonomias, Fard e Cunningham (2019) é o artigo mais abrangente, considerando 6 critérios (Evidência, Significância, Sobre indivíduos, Intenção, Veracidade e Função). Os autores revisam extensivamente a literatura de diversas áreas além da Computação (Psicologia, Sociologia, etc.), partindo de 1900 até 2019. Os resultados apontam a diferença de critérios considerados por diferentes taxonomias como um **fator responsável pela falta de consenso** da literatura e propõe uma taxonomia para clarificar a tipologia. Entretanto, o significado de alguns critérios são pouco claros. Por exemplo, o critério Função, que apresenta vários tipos de valores (magoar, explicar a realidade, explicar um sentimento, manipular, etc.), mas que não são discutidos, conceituados ou apresentados em sua totalidade. Assim, considerar o critério Função pode levar a ambiguidade na categorização dos tipos de desinformação oriunda da interpretação subjetiva de quem utilizar a taxonomia proposta. Especialmente quando consideramos que a Função de um tipo de desinformação depende de fatores como o significado de ações, atitudes, propósitos e motivações, para diferentes partes interessadas que se comunicam. Por exemplo, o estado mental de quem expressa um determinado tipo de desinformação, como quando a Intencionalidade (parte interessada) difere da Função (mensagem), e o estado mental da audiência, como quando a Intencionalidade percebida (parte interessada) é a mesma da Função percebida (mensagem). Isso indica que as taxonomias são limitadas ao considerar a essência informal do fenômeno da desinformação e da subjetividade das partes interessadas que interagem com ela, sejam os pesquisadores e projetistas de intervenções, seja o agente que desinforma e sua audiência. Ainda, se o problema se encontra na quantidade de taxonomias, mais uma acrescenta complexidade ao problema e o potencial de confundir ao invés de clarificar a terminologia.

#### A.3.3.2 Definições de desinformação

Em geral, as pesquisas conceituam *Disinformation* e *Misinformation* de duas formas. Há pesquisadores que consideram *Misinformation* e *Disinformation* como as principais categorias de informações falsas, que agrupam todos os tipos de falsidade, como notícias falsas, sátiras, e conspirações (Shao et al., 2018; Watney, 2018; Torres et al., 2018b; Kumar e Geethakumari, 2014). Existem pesquisadores que consideram três categorias: *Misinformation*, *Disinformation* e as sátiras (Charles e de Oliveira Sampaio, 2018). Outros artigos consideram apenas (*misinformation*) como uma categoria genérica que engloba todas as outras formas de informações falsas (Jiang e Wilson, 2018; Introne et al., 2017). As diferentes formas de entender a desinformação representam a falta de um padrão para a terminologia e categorização do objeto de estudo de cada trabalho, o que dificulta a organização das pesquisas e torna o processo de mapeamento moroso.

As definições *disinformation* são integras quanto a critério Intencionalidade. Alguns consideram *disinformation* como a introdução intencional de viés (Babaei et al., 2019; Watney, 2018; Cho et al., 2017), quando o distribuidor está ciente da mentira (Campan et al., 2017; Kshetri e Voas, 2017; Kumar e Geethakumari, 2014; Watney, 2018), informações falsa criadas para enganar (Guacho et al., 2018; Cho et al., 2016), e que são resultantes da alteração, adição ou remoção de partes de uma mensagem (Raynal e Gaspard, 2010; Karduni et al., 2018). Os resultados mostram a consistência da comunidade em relação ao termo *disinformation* e sua relação com a intenção de quem cria ou compartilha uma falsidade.

Além disso, encontramos artigos que usam os termos *misinformation* e *disinformation* intermitentemente no texto, sem uma definição (Ferrara, 2015; Resnik, 1998; Popat et al., 2017). Os resultados indicam a complexidade de caracterizar a tipologia do fenômeno de desinformação, representada pela ambiguidade nos entendimentos dos termos para desinformação.

Quanto a categorização, em consonância com Fard e Cunningham (2019), há pesquisadores que consideram *misinformation* e *disinformation* como as principais categorias de informações falsas, que agrupam todos os tipos de falsidade, como notícias falsas, sátiras,



e conspirações (Shao et al., 2018; Watney, 2018; Torres et al., 2018b; Kumar e Geethakumari, 2014), outros consideram três categorias: *misinformation*, *disinformation* e as sátiras de modo separado (Charles e de Oliveira Sampaio, 2018). Ainda, outra perspectiva adota *misinformation* como uma categoria geral que engloba todas as outras formas de informações falsas (Jiang e Wilson, 2018; Introne et al., 2017). As diferentes formas de entender a desinformação representam a falta de um padrão para a terminologia e categorização do objeto de estudo de cada trabalho, o que dificulta a organização das pesquisas e torna o processo de mapeamento moroso.

#### A.3.3.3 Discussão

A falta de consenso na terminologia torna o esforço para mitigar informações falsas uma tarefa contraproducente. A ambiguidade confunde a categorização dos tipos de informações falsas, como categorizar uma notícia falsa como uma desinformação acidental (*misinformation*). Isso faz com que pesquisadores da área de Ciência da Computação que pretendem trabalhar com informações falsas online não tenham um panorama integral e conciso das publicações relacionadas a um tipo de desinformação, e acabem desenvolvendo um conhecimento que já existe. Além disso, a falta de consenso torna a revisão da literatura um esforço demorado, necessitando uma revisão cuidadosa das definições dos termos nos artigos. Ainda, o problema causa confusão na comunidade acadêmica investigando a desinformação e força os pesquisadores a uma escolha ambígua da terminologia usada nas suas pesquisas (Pinheiro et al., 2018; Ahmad et al., 2019; Ferrara, 2015; Søre, 2018; Popat et al., 2017).

Ao adotar um termo "guarda-chuva", como *misinformation*, perde-se a granularidade de propriedades que diferem um tipo de desinformação de outro, e generalizam soluções que não são adequadas às necessidades do fenômeno em diferentes contextos. Por exemplo, a intencionalidade é uma propriedade relevante para a análise de desinformação, em especial quando se caracteriza as partes interessadas envolvidas na criação da desinformação. Nesse sentido, um termo genérico acaba sendo prejudicial para o desenvolvimento do conhecimento sobre o fenômeno.

A desinformação digital tem sido uma preocupação, pelo menos, desde o desenvolvimento e popularização da Web (Luciano, 1996). Embora algumas discussões sobre informações falsas digitais datem de 1998 (Resnik, 1998), e estudos definindo uma taxonomia datem do início dos anos 2000 (Zhou e Zhang, 2004), não há acordo da comunidade acadêmica da Ciência da Computação sobre a terminologia de informações falsas e seus tipos (Pang e Ng, 2017; Jiang e Wilson, 2018). Encontramos duas origens para o problema da falta de consenso: diversidade de taxonomias e os fundamentos teóricos das pesquisas. Fard e Cunningham (2019) apontam a diversidade de critérios empregados por diferentes taxonomias propostas na área de Ciência da Computação como um fator que leva a dissonância nas definições. Por exemplo, o uso do termo *misinformation* pode significar tanto uma desinformação deliberada quanto acidental, dependendo da taxonomia adotada. Essa falta de consenso torna o esforço para mitigar informações falsas uma tarefa contraproducente, pois força os pesquisadores a uma escolha ambígua da terminologia usada nas suas pesquisas (Pinheiro et al., 2018; Ahmad et al., 2019; Ferrara, 2015; Søre, 2018; Popat et al., 2017). Além disso, a ambiguidade faz com que não se tenha um panorama integral e conciso das publicações em um determinado tipo de desinformação, além do potencial de enviesar o escopo das pesquisas a serem desenvolvidas. A outra origem para o problema reside nos estudos que não adotam uma das taxonomias existentes. Nestes estudos, a literatura usada pelos pesquisadores para fundamentar suas pesquisas constitui o problema. Por ser um fenômeno multidisciplinar, os conceitos de desinformação encontrados em alguns estudos têm origem em diferentes disciplinas, como Sociologia (Chen e Sin, 2013), Psicologia (Allport e Postman, 1947) e Economia (Che et al., 2018). Cada área conceitua a desinformação seguindo

linhas de pesquisa próprias e que, presumidamente, resultam em concisão dentro dessas áreas. Quando a computação empresta os termos de outras áreas, essa concisão perde o sentido.

Um exemplo dessa ocorrência: o trabalho de Dang et al. (2016b) que referência o conceito de rumores (Rosnow, 1991) - rumores são declarações não verificadas - e Jin et al. (2017) que referencia (Allport e Postman, 1947) - rumores são falsidades deliberadas ou declarações não verificadas. Um conceito de rumor de 1947, proveniente da Psicologia, é adotado por uma pesquisa de 2017; no outro caso, um conceito de rumor de 1991, também proveniente da Psicologia, é adotado por uma pesquisa de 2016. Assim, o emprego dos mesmos termos, mas com contextos e conceitos diferentes, sem um referencial teórico consistente e uma taxonomia adequada para a computação, acarreta consequências para a caracterização do objeto de estudo e para a comunicação dos avanços do conhecimento no campo da desinformação digital. Esse aspecto é recorrente na literatura analisada, pela reutilização de termos por estudos que referenciam outros estudos, sem uma revisão cuidadosa de conceitos e entendimentos.

A falta de consenso também reflete uma perspectiva limitada em relação a caracterização do contexto estudado, e no potencial de gerar soluções de mitigação pouco eficiente e responsáveis (Fernandez e Alani, 2018). A tendência de abordar a desinformação a partir do critério Evidências representa uma perspectiva que aborda a realidade objetiva do fenômeno da desinformação. Entretanto, a desinformação digital existe a partir da experiência das pessoas, sendo um fenômeno social que abrange a subjetividade das interpretações dos participantes da comunicação. Assim, o que uma pessoa entende como verdade não é necessariamente o que outra pessoa entende como verdade, e algumas vezes não importando as evidências da realidade objetiva (por exemplo, o terra-planismo). Dessa forma, o critério Evidência pode enviesar o estudo de fatores de impacto no fenômeno, se restringindo a perspectiva objetiva, deixando em segundo plano como as pessoas se relacionam com a desinformação<sup>25</sup>.

Este problema se assemelha com o que Ronald Stamper encontrou em 1973, quando escrevia sobre informação em sistemas administrativos (Stamper, 1973). Stamper critica o entendimento do termo "informação" como sendo insuficiente para tratar a complexidade daquilo que ele representa, em especial sobre a interpretação das partes interessadas que projetam sistemas computacionais (Stamper, 1973). Para Stamper, informação é uma entidade subjetiva que pode ter diferentes interpretações dependendo das circunstâncias (Stamper, 1973). Da mesma forma, desinformação é um conceito complexo, que necessita uma perspectiva que considere a subjetividade de diferentes pontos de vista. Por esse motivo, esta pesquisa adota as lentes da Semiótica Organizacional de Stamper (Stamper e Liu, 1994) para definir desinformação. O Quadro A.11 apresenta nossa visão do entendimento de desinformação considerado nesta tese<sup>26</sup>.

<sup>25</sup>Voltar ao Cap. de estado da arte: Início 3 ou QP.02 3.1.0.1

<sup>26</sup>Mais detalhes desse entendimento podem ser encontrados no Capítulo 2 (Fundamentação teórica)

Tabela A.11: Descrição das categorias de desinformação

| Desinformação | Descrição                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Deliberada    | <p>Configura o uso <b>intencional</b> de signos<sup>27</sup> que não correspondem a um estado real das coisas ou dos consensos de uma determinada configuração social, com um propósito de manipular a audiência para obter algum tipo de ganho.</p> <ul style="list-style-type: none"> <li>• Empregam técnicas de persuasão para manipulação das partes interessadas.</li> <li>• Exploram aspectos técnicos (introduz ruído, interferência, etc.), humanos (vieses cognitivos, percepção, emoções, etc.), sociais (proximidade afetiva, intencionalidade, autoridade, consensos etc.), e organizacionais (procedimentos, protocolos, regras de negócios, etc.).</li> <li>• As consequências são conhecidas, podendo ser ponderadas, ou negligenciadas pela parte mal intencionada.</li> </ul> |
| Acidental     | <p>Configura o uso <b>não intencional</b> de signos que não correspondem a um estado real das coisas ou dos consensos de um grupo.</p> <ul style="list-style-type: none"> <li>• Distorções técnicas acidentais, como o ruído na comunicação, uma tela defeituosa ou um erro de digitação, que pode causar falhas de representação ou interpretação.</li> <li>• Distorções perceptivas acidentais, como deficiências de sensoriamento visual ou auditivo.</li> <li>• Distorções cognitivas acidentais, como erros de memória, vieses de interpretação, e falácias lógicas.</li> <li>• As consequências não são conhecidas ou devidamente ponderadas pela parte interessada que comunica.</li> </ul>                                                                                             |

#### A.3.4 Relatório: QP01.b: Quais os tipos de desinformação estudados e os temas?

O mapeamento sobre os tipos de desinformações revelou 26 termos usados para diferenciar falsidades com pequenas particularidades. Tabela A.12 apresenta um resumo desses resultados, reunindo os tipos que aparecem em mais de um trabalho.

Tabela A.12: Quantidade de artigos por tipo de desinformação

| Tipo                  | Quantidade  |
|-----------------------|-------------|
| Notícias falsificadas | 269 artigos |
| Rumores               | 69 artigos  |
| Farsas                | 27 artigos  |
| Isca-clique           | 17 artigos  |
| Sátiras/Parodias      | 8 artigos   |
| Conspirações          | 8 artigos   |
| Revisões falsificadas | 7 artigos   |
| Spam                  | 6 artigos   |

Os resultados indicam termos iguais usados com entendimentos diferentes, além de definições ambíguas. Por exemplo, em (Khan et al., 2019) notícias falsificadas são sátiras, e em (Ishida e Kuraya, 2018) são farsas. O Quadro A.13 apresenta um resumo dos problemas encontrados e exemplos para os principais tipos de desinformação.

<sup>27</sup>Estímulos da realidade que representam algo para alguém em alguma configuração social (Stamper, 1973).

Tabela A.13: Problemas de concisão no entendimento dos tipos de desinformação

| Tipo                  | Entendimento                                                                                      | Exemplos                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|-----------------------|---------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Rumores               | Ambiguidade de entendimentos quanto aos critérios<br>Intencionalidade e Evidência.                | <b>Informação não verificada</b> (Habib et al., 2019; Lee e Choi, 2018; Patel et al., 2017; Metaxas et al., 2015); <b>Que pode ser verdadeira ou falsa</b> (Rana et al., 2019; Buntain e Golbeck, 2017); <b>Informação falsa</b> (Qin et al., 2018; Chen et al., 2017; Zhang et al., 2015; Wang et al., 2014; Qazvinian et al., 2011); <b>Propaganda falsa</b> (Tan et al., 2019).                                                                                                                                                             |
| Farsas                | Concisão quanto a Intencionalidade.                                                               | <b>Informação enganosa e maliciosa</b> usada para enganar e manipular pessoas (Yuliani et al., 2018; Hui et al., 2018; Prasetyo et al., 2017); <b>Conteúdo anti-social</b> intencional, como difamação e <i>bullying</i> (Yu et al., 2018);                                                                                                                                                                                                                                                                                                    |
| Notícias Falsificadas | Concisão de entendimentos quanto aos critérios<br>Intencionalidade e Evidência.                   | <b>Qualquer forma de informação falsa</b> , de farsas à sátiras (Karduni et al., 2018); Conteúdo <b>intencionalmente criado para enganar</b> (Habib et al., 2019; Della Vedova et al., 2018; Armano et al., 2018; Al-Ash e Wibowo, 2018; Figueira e Oliveira, 2017); Histórias postadas como <b>falso fato aceito como genuíno</b> (Parikh e Atrey, 2018; Gilda, 2017); <b>Declarações tendenciosas</b> (Murungi et al., 2018); <b>Fatos alternativos</b> sem base na realidade (Purnomo et al., 2017). <b>Sátiras</b> (Chandra et al., 2017). |
| Sátiras               | Ambiguidade de entendimentos quanto aos critérios<br>Intencionalidade e o teor cômico.            | Sub-tipos: paródias, piadas e pegadinhas (Khan et al., 2019; Popat et al., 2017); <b>Desinformação acidental</b> (Cybenko e Cybenko, 2018; Ishida e Kuraya, 2018); <b>Desinformação deliberada</b> (Bedard e Schoenthaler, 2018); Histórias para <b>entretenimento</b> (Karduni et al., 2018; Campan et al., 2017).                                                                                                                                                                                                                            |
| Teoria da conspiração | Ambiguidade de entendimentos quanto ao critério<br>Intencionalidade. Concisão quanto a Evidência. | São <b>fabricações intencionais</b> (Tacchini et al., 2017); <b>informações não confiáveis</b> para explicar eventos ou circunstâncias (Glenski et al., 2018); <i>misinformation</i> como informações falsas, tanto deliberadas quanto acidentais, que reduzem a complexidade de eventos sociais (Bessi et al., 2015);                                                                                                                                                                                                                         |

Sobre os temas abordados pelos estudos, a Figura A.6 apresenta um resumo dos resultados.

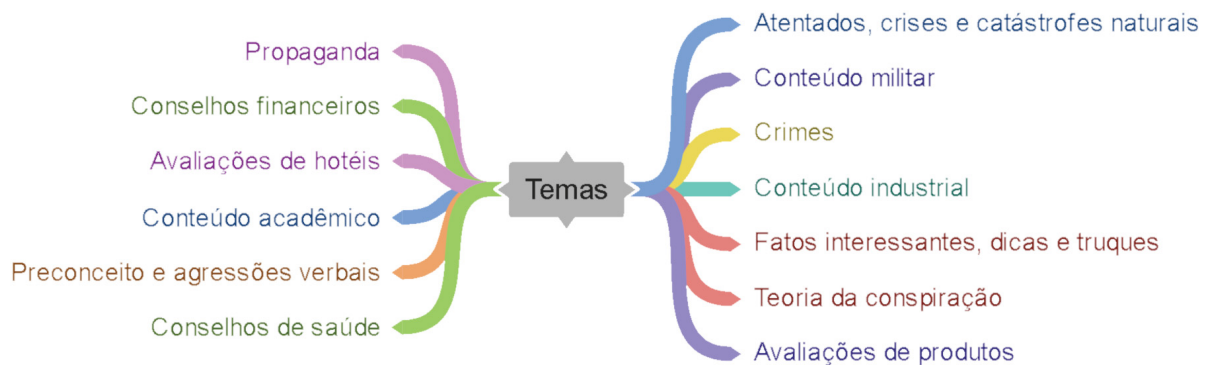


Figura A.6: Temas abordados pelos estudos.

#### A.3.4.1 Sobre rumores, boatos, notícias falsas, sátiras e isca de clique

O termo rumor apresenta uma grande variação de significado. Alguns artigos definem rumor como informação não verificada. (Patel et al., 2017; Devi e Karthika, 2018; Lee e Choi, 2018; Metaxas et al., 2015; Chua e Banerjee, 2017; Habib et al., 2019; Friggeri et al., 2014; Boididou et al., 2018) que pode mais tarde ser confirmado como verdadeira ou falsa (Rana et al., 2019; Buntain e Golbeck, 2017). Outros definiram o rumor como uma declaração polêmica (Chen et al., 2016b). Para alguns, o rumor é uma informação falsa (Qin et al., 2018; Chen et al., 2017; Zhang et al., 2015; Liu et al., 2015; Wang et al., 2014; Qazvinian et al., 2011) ou também propaganda falsa (Tan et al., 2019).

Em relação à categorização de rumores quanto a *misinformation* ou *disinformation*, alguns consideram os rumores um tipo particular de *misinformation* (Charles e

de Oliveira Sampaio, 2018; Al-Sarem et al., 2019; Pinheiro et al., 2018; Tong et al., 2017; Gupta et al., 2013; Tong et al., 2018; Akhtar et al., 2018; Dhar et al., 2016; Tan et al., 2019). Também há artigos que o consideram como *disinformation* (Søe, 2018). Outra interpretação considera o rumor como um recipiente que pode conter *misinformation* ou *disinformation* (Cai et al., 2014; Metaxas et al., 2015; Indu e Thampi, 2019). Boididou et al. (2018) categorizou o rumor em três grupos principais:

- conteúdo de um evento passado compartilhado como novo, adicionando viés na percepção do tempo;
- conteúdo que é deliberadamente manipulado;
- conteúdo multimídia publicado juntamente com uma falsa alegação sobre o evento.

Charles e de Oliveira Sampaio (2018) apresentam outra classificação para rumores. Eles apresentam uma revisão da literatura sobre rumores, agrupando os resultados em diferentes categorias:

- Tabloides, que produzem notícias, que podem ser falsas, sobre escândalos, famosos ou crimes, com o público-alvo.
- Engôdos, que são rumores elaborados nas redes sociais, utilizando fatos reais como validação imprópria com o intuito de aumentar a confiança sobre a mensagem.
- Notícias de humor são sátiras.

Encontramos uma definição relativamente estável de engôdos (*hoaxes*). Alguns trabalhos o consideram como informação enganosa e maliciosa usada para enganar e manipular pessoas (Hernandez et al., 2002; Prasetijo et al., 2017; Yuliani et al., 2018; Santoso et al., 2017; Hui et al., 2018). Outros definem engôdos como mensagens usadas para entreter as pessoas, mas apenas se for claro para o intérprete que a informação é falsa (Santoso et al., 2017). Pode consistir em conteúdo de áudio, texto e multimídia (Habib et al., 2019). Alguns consideram as notícias falsas como engôdos (Kesumawati e Thalib, 2018; Yu et al., 2018; Pratiwi et al., 2017; Wijaya e Santoso, 2018) e conteúdo anti-social, como difamação e *bullying* (Yu et al., 2018). Alguns pesquisadores definem engôdo como um tipo de *disinformation* (Baeth e Aktas, 2019; Pinheiro et al., 2018; Jiang e Wilson, 2018; Hui et al., 2018; Volkova e Jang, 2018). Outros consideram a engôdos como *misinformation* (Volkova et al., 2017; Kesumawati e Thalib, 2018). Santoso et al. (2017) apresentam uma revisão classificando o *hoax* em seis categorias:

- Erros de relatos não intencionais;
- Boatos que não se originam de uma notícia;
- Teorias da conspiração;
- Sátira;
- Declarações falsas de políticos.
- Relatórios tendenciosos ou enganosos;



Em relação ao termo notícias falsas, encontramos mais concisão. Algumas pesquisas definem notícia falsa como conteúdo intencionalmente criado para enganar (Habib et al., 2019; Della Vedova et al., 2018; Armano et al., 2018; Murungi et al., 2018; Al-Ash e Wibowo, 2018; Özgöbek e Gulla, 2017; Figueira e Oliveira, 2017; Che et al., 2018; Aldwairi e Alwahedi, 2018; Della Vedova et al., 2018; Jang et al., 2018; Alianto et al., 2017; Srivastava et al., 2018; Al-Ash e Wibowo, 2018; Parikh e Atrey, 2018). Outras consideram como histórias postadas como falso fato que é aceito como genuíno (Parikh e Atrey, 2018; Gilda, 2017; Srivastava et al., 2018; DiFranzo e Gloria-Garcia, 2017). Alguns comparam notícias falsas a *spins*, que são declarações tendenciosas (Murungi et al., 2018). Alguns trabalhos consideram as notícias falsas como fatos alternativos sem base na realidade (Purnomo et al., 2017). Outros consideram as notícias falsas como sátiras sobre assuntos públicos (Chandra et al., 2017).

Sobre a caracterização de notícias falsas em *misinformation* ou *disinformation*, alguns trabalhos a consideram como desinformação intencional (*disinformation*) (Armano et al., 2018; Jang et al., 2018; Wilder e Vorobeychik, 2018; Samonte, 2018; Shu et al., 2017; Korshunov e Marcel, 2018; Kim et al., 2018; Traylor et al., 2017; Jiang e Wilson, 2018; Baeth e Aktas, 2019; Pinheiro et al., 2018; Shu et al., 2018; Kotteti et al., 2018). Outros classificam as notícias falsas como *misinformation* (Varol e Uluturk, 2018; Granik e Mesyura, 2017; Kshetri e Voas, 2017). Alguns usam o termo notícia falsa para representar qualquer forma de informação falsa, desde *disinformation*, *misinformation*, engêdos, histórias de conspiração, isca de clique e sátiras (Karduni et al., 2018). Ainda, encontramos trabalhos que classificam notícias falsas como *misinformation* se o distribuidor acredita que a notícia é verdadeira ou em *disinformation* se o distribuidor tem conhecimento dos fatos fabricados (Campan et al., 2017; Pal e Loke, 2019).

Em relação a categorização para notícias falsas, Ishida e Kuraya (2018) apontam os objetivos da informação, categorizando as notícias falsas em dois tipos: notícias subjetivas e notícias objetivas. A primeira categoria representa o conteúdo em que o valor da verdade é baseado nos pontos de vista ou posições políticas do leitor. A notícia objetiva é onde o valor da credibilidade depende dos fatos objetivos do conteúdo. O trabalho de (Manzoor e Nikita, 2019) categoriza as notícias falsas em cinco tipos:

- Baseado na visão - uso de gráficos, imagens transformadas, vídeo adulterado, etc. Outros autores também consideram notícias falsas na forma de imagens Huh et al. (2018), *deep-fakes*, *tempered-videos* (Korshunov e Marcel, 2018; Wang et al., 2018) e transformação de voz (Wang et al., 2018).
- Baseado no usuário - gerado por contas falsas e direcionado a um público específico (faixa etária, gênero, cultura, afiliações políticas).
- Baseado em conhecimento - explicações pseudo-científicas para as mesmas questões não resolvidas.
- Baseado no estilo - Notícias fingindo e copiando o estilo de alguns jornais credenciados.
- Baseada na postura: conteúdo tendencioso, representando afirmações verdadeiras com mudança de significado e finalidade.

Em uma perspectiva enganosa de notícias falsas, Rubin et al. (2015) categorizam as notícias falsas em três tipos:

- Fabricações graves - descobertas na mídia tradicional ou participante, imprensa amarela ou tabloides;



- Hoaxes em grande escala - Hoaxes que são criativos, exclusivos e muitas vezes multiplataforma;
- Falsificações de humor - sátira de notícias, paródia, programas de jogos;

Khan et al. (2019) apresentam uma revisão da literatura sobre categorias de notícias falsas, estendendo o trabalho de (Campan et al., 2017):

- Conteúdo de clique com isca - conteúdo que chama a atenção do usuário e induz o clique. Ele usa gráficos e manchetes enganosos, sensacionalistas e atraentes.
- Sátira e Paródia - Conteúdo com a finalidade de divertir, é divertido e divertido. Esse tipo de conteúdo usa sarcasmo, histórias falsas e humor.
- Notícias baseadas em propaganda - Propaganda criada para enganar, promover um ponto de vista tendencioso ou uma causa ou agenda política específica. Normalmente usado por governos.
- Notícias descuidadas - histórias publicadas em casas de mídia com informações não confiáveis ou sem qualquer verificação de fatos.
- Notícias tendenciosas / tendenciosas - interpretações ideológicas de fatos unilaterais.

Em relação às definições de "sátira", há uma paisagem homogênea. Sátira são histórias para entretenimento (Campan et al., 2017; Chandra et al., 2017; Karduni et al., 2018). Seus tipos são paródias, piadas e pegadinhas (Khan et al., 2019; Popat et al., 2017). Também houve referências diretas a este tipo como *misinformation* (Cybenko e Cybenko, 2018; Ishida e Kuraya, 2018).

O termo 'rumor' apresenta uma grande variação de significado. Alguns artigos definem rumor como informação não verificada. (Patel et al., 2017; Devi e Karthika, 2018; Lee e Choi, 2018; Metaxas et al., 2015; Chua e Banerjee, 2017; Habib et al., 2019; Friggeri et al., 2014; Boididou et al., 2018) que pode mais tarde ser confirmado como verdadeira ou falsa (Rana et al., 2019; Buntain e Golbeck, 2017). Outros definiram o rumor como uma informação falsa (Qin et al., 2018; Chen et al., 2017; Zhang et al., 2015; Liu et al., 2015; Wang et al., 2014; Qazvinian et al., 2011) ou também propaganda falsa (Tan et al., 2019). O termo rumor apresenta ambiguidades que configuram intencionalidades diferentes no seu uso.

A definição de farsas (*hoaxes*) apresenta concisão, com a maior parte dos trabalhos conceituando farsas como desinformação deliberada. Alguns trabalhos o consideram como informação enganosa e maliciosa usada para enganar e manipular pessoas (Hernandez et al., 2002; Prasetijo et al., 2017; Yuliani et al., 2018; Santoso et al., 2017; Hui et al., 2018). Outros definem farsas como mensagens usadas para entreter as pessoas, mas apenas se for claro para o intérprete que a informação é falsa (Santoso et al., 2017). Pode consistir em conteúdo de áudio, texto e multimídia (Habib et al., 2019). Alguns consideram as notícias falsificadas como engêdos (Kesumawati e Thalib, 2018; Yu et al., 2018; Pratiwi et al., 2017; Wijaya e Santoso, 2018) e conteúdo anti-social, como difamação e *bullying* (Yu et al., 2018). Alguns pesquisadores definem farsas como um tipo de *disinformation* (Baeth e Aktas, 2019; Pinheiro et al., 2018; Jiang e Wilson, 2018; Hui et al., 2018; Volkova e Jang, 2018).

Em relação ao termo notícias falsificadas, encontramos mais concisão. Algumas pesquisas definem notícia falsa como conteúdo intencionalmente criado para enganar (Habib et al., 2019; Della Vedova et al., 2018; Armano et al., 2018; Murungi et al., 2018; Al-Ash e Wibowo, 2018; Özgöbek e Gulla, 2017; Figueira e Oliveira, 2017; Che et al., 2018; Aldwairi e

Alwahedi, 2018; Della Vedova et al., 2018; Jang et al., 2018; Alianto et al., 2017; Srivastava et al., 2018; Al-Ash e Wibowo, 2018; Parikh e Atrey, 2018; Shu et al., 2017; Sirajudeen et al., 2017; Jiang e Wilson, 2018; Garcia-Recuero et al., 2018; Badawy et al., 2018; Parikh e Atrey, 2018). Outras consideram como histórias postadas como falso fato que é aceito como genuíno (Parikh e Atrey, 2018; Gilda, 2017; Srivastava et al., 2018; DiFranzo e Gloria-Garcia, 2017). Alguns comparam notícias falsificadas a *spins*, que são declarações tendenciosas (Murungi et al., 2018). Alguns trabalhos consideram as notícias falsificadas como fatos alternativos sem base na realidade (Purnomo et al., 2017). Outros consideram as notícias falsificadas como sátiras sobre assuntos públicos (Chandra et al., 2017).

Sátiras são histórias para entretenimento (Campan et al., 2017; Chandra et al., 2017; Karduni et al., 2018). Seus tipos são paródias, piadas e pegadinhas (Khan et al., 2019; Popat et al., 2017). Também houve referências diretas a este tipo como *misinformation* (Cybenko e Cybenko, 2018; Ishida e Kuraya, 2018).

Sobre os tópicos dos conteúdos, encontramos pesquisas sobre teorias da conspiração política (Flintham et al., 2018; Volkova et al., 2017), propaganda política (Al-Rawi et al., 2019; Bedard e Schoenthaler, 2018; Karadzhov et al., 2018; Wang, 2017; Volkova et al., 2017), pessoas famosas (Al-Rawi et al., 2019; Bedard e Schoenthaler, 2018); conteúdo militar, industrial e acadêmico, (Sethi, 2017; Granik e Mesyura, 2017; Volkova et al., 2017), conselhos financeiros (Bedard e Schoenthaler, 2018), crimes (Volkova et al., 2017), preconceito e agressões verbais (Bessi et al., 2015; Ferrara, 2015), fatos interessantes, dicas e truques (Karadzhov et al., 2018), propagandas, avaliações de hotéis (Sandifer et al., 2017), avaliações de produtos (Yao et al., 2017; Ratkiewicz et al., 2011), atentados, crises e catástrofes naturais (Rajdev e Lee, 2015; Oh et al., 2010; Gupta et al., 2013), e conselhos de saúde (Chomphosang et al., 2012; Hailun et al., 2014). Os resultados revelam contextos ainda pouco explorados como esportes, apostas, artes, e religião.

#### A.3.4.2 Discussão

Os resultados mostram uma diversidade de definições para os tipos de desinformação. Observamos os efeitos da falta de consenso para os termos *misinformation* e *disinformation*. Por exemplo, "sátira" sendo desinformação intencional (Khan et al., 2019) e acidental (Cybenko e Cybenko, 2018). Nesse exemplo, os autores referenciam estudos que definem *misinformation* com diferentes conotações. Isso confunde os pesquisadores na escolha dos termos e definições (Pinheiro et al., 2018; Ahmad et al., 2019; Ferrara, 2015; Sør, 2018; Popat et al., 2017), e generaliza o entendimento do objeto estudado, sendo limitante para a compreensão da complexidade do fenômeno. O Quadro A.14 apresenta o entendimento adotado nesta pesquisa para os tipos desinformação com base no conceito de signo<sup>28</sup> oriundo da SO (Stamper, 1973).

<sup>28</sup>Estímulos da realidade que representam algo para alguém em alguma configuração social (Stamper, 1973).

Tabela A.14: Tipos de desinformação

| Tipo                     | Intencionalidade          | Descrição                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|--------------------------|---------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Rumores                  | Informação não verificada | São signos que podem ou não ser uma desinformação (Dang et al., 2016b). Fica implícito para os envolvidos na comunicação que a informação não é verificada pelo interlocutor. Geralmente referencia terceiros (pessoas, organizações, etc.) como autoridade certificadora da informação. Pode ser revelada como verdade ou falsidade posteriormente (Dang et al., 2016a).                                                                                                                                                                                         |
| Sátiras                  | Deliberada                | São signos usados deliberadamente para enganar, sendo informações falsas de conotação cômica empregadas para entretenimento (Campan et al., 2017). A intenção de enganar tem o objetivo implícito de entreter comicamente, onde a audiência está consciente disso ou tem "pistas" que indicam o teor cômico. Podem ser paródias, piadas e pegadinhas (Khan et al., 2019; Popat et al., 2017).                                                                                                                                                                     |
| Farsas ( <i>hoaxes</i> ) | Deliberada                | São signos usados deliberadamente para enganar as partes interessadas. Geralmente exploraram crenças, expectativas, e intenções. Podem ser histórias complexas, como teorias da conspiração (Recuero et al., 2020), bem contextualizadas, como fraudes (Bossetta, 2018), ou um ato oportunístico, como o <i>cyberbullying</i> (Ferrara, 2015).                                                                                                                                                                                                                    |
| Notícias Falsificadas    | Deliberada                | São um subtipo de farsas fabricadas para enganar a audiência, manipular crenças e atitudes. São similares às notícias veiculadas por autoridades certificadas, como a diagramação de jornais e portais de notícias (Al-Rawi et al., 2019). A audiência não está ciente da enganação. Podem ser totalmente fabricadas ou introduzir viés (da Silva Gomes e Dourado, 2019).                                                                                                                                                                                         |
| Teorias da Conspiração   | Deliberada/<br>Acidental  | São um subtipo de farsas que tendem a reduzir a complexidade da realidade, considerando aspectos sociais como sendo tramas maquiavélicas, muitas vezes concebidas por indivíduos ou organizações poderosas (Bessi et al., 2015). Também podem ser oriundas de distorções no conhecimento de uma comunidade, relacionadas a confiança cega no julgamento de uma autoridade (Acemoglu e Ozdaglar, 2011). Formam comunidades em torno do tema, com indivíduos que podem não ter ciência da farsa (Bandeli e Agarwal, 2018). Difícil de verificar (Shu et al., 2017). |
| Isca-clique              | Deliberada                | São signos usados com a intenção de manipular a atenção da audiência e obter acessos (Zannettou et al., 2019). Frequentemente empregam linguagem sensacionalista e apelo emocional (Glenski et al., 2018). Geralmente o conteúdo da página diverge do apresentado no signo (Chen et al., 2015c).                                                                                                                                                                                                                                                                  |

Em relação aos tipos e plataformas de comunicação, predominam artigos sobre o conteúdo gerado por usuários em redes sociais (528 artigos). O tipo de desinformação mais estudado são "notícias falsas" (269 artigos dedicados). Assim, há necessidade de pesquisas em outros tipos de desinformações e outros meios como e-mails, fóruns e apps de mensagens.

#### A.3.5 Relatório: QP03 - Como as pesquisas abordam os aspectos sociotécnicos?

Em relação a multidisciplinaridade e os aspectos sociotécnicos considerados, a Tabela A.15 resume os resultados.

##### A.3.5.1 Mundo Social & Pragmática

Nós categorizamos 57 artigos abordando aspectos do mundo social. Esta categoria inclui problemas de pesquisa associados a características de grupos ou comportamento de usuários em grupos. Foram 26 trabalhos diretamente envolvidos na investigação do comportamento social. Categorizamos oito trabalhos sobre comportamento de difusão. Outras sete pesquisas sobre estratégias de intervenção baseadas na conscientização e alfabetização do grupo de usuários. A Tabela A.16 apresenta os resultados.

Tabela A.15: Distribuição dos estudos por sub-disciplina

| Dimensão     | Quantidade |
|--------------|------------|
| Mundo Social | 57         |
| Pragmática   | 117        |
| Semântica    | 418        |
| Sintática    | 186        |
| Empírica     | 6          |
| Mundo Físico | 2          |

Tabela A.16: Classificação da dimensão Mundo Social

| Tipo                       | Descrição                                                                                                                                                                                                                                                                                                                                       | Quantidade |
|----------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| Análise do Comportamento   | Pesquisa os padrões de consumo dos usuários em grupos, perfis de usuários, com base em teorias sociológicas, como veracidade coletiva e métricas sociais quantitativas, como classificações e avaliações. Também trabalha com estratégias de verificação, informação correção, percepções do usuário sobre credibilidade, capacidade cognitiva. | 16         |
| Modelo de informação falsa | Pesquisa modelando aspectos de informações falsas relacionados à dimensão social, como crenças, valores e traços culturais.                                                                                                                                                                                                                     | 26         |
| Dinâmica de difusão        | Pesquisa baseada em aspectos da dimensão social relacionados ao comportamento de difusão de informações falsas, como práticas de compartilhamento em grupos, vulnerabilidades da audiência, etc.                                                                                                                                                | 08         |
| Estratégias de intervenção | Estratégias de intervenção testadas em grupos, como aumento da conscientização e alfabetização digital.                                                                                                                                                                                                                                         | 07         |

Exemplos desta categoria são estudos sobre as características de perfil do usuário que os tornam propensos a consumir ou descartar notícias falsas (Shu et al., 2018; Bay, 2018; Chen et al., 2016a, 2015b), seus impactos na formação de crenças (Jameel et al., 2019; Dimo, 2019; Introne et al., 2017; Theng et al., 2013; Chen e Sin, 2013; Acemoglu e Ozdaglar, 2011), as transformações causadas na tecnologia de comunicação (Miletskiy et al., 2019), aspectos sociodemográficos relacionados ao comportamento de consumo e compartilhamento (Bedard e Schoenthaler, 2018; Chandra et al., 2017; Boshmaf et al., 2013), em valores sociais como privacidade (Cho et al., 2016), em relevantes aspectos de grupos de mídia social, consciência do usuário, percepção de credibilidade e verificadores de fatos para difusão de desinformação (Wang e Fussell, 2020; Zhang et al., 2019c; Shao et al., 2018; Murungi et al., 2018; Safieddine et al., 2017; Hannak et al., 2014).

Foram categorizados 117 problemas na dimensão Pragmática. Esta categoria inclui problemas de pesquisa sobre as intenções e seus impactos no consumo do usuário, diferenciando os papéis da desinformação e da desinformação na criação de informações falsas (McCarthy, 2018), comportamento de compartilhamento e campanhas de informações falsas (Wang et al., 2018). O Quadro A.17 resume as categorias.

A dimensão pragmática inclui problemas de pesquisa sobre o comportamento do usuário diante de informações falsas, evidenciando a intencionalidade e seus efeitos no consumo do usuário. Por exemplo, trabalha na diferenciação de papéis de desinformação e desinformação na criação de informações falsas (McCarthy, 2018), compartilhando comportamento em campanhas de informações falsas (Wang et al., 2018). Investigações em comportamentos de compartilhamento de desinformação, análise de estratégias de intervenção, como dicas visuais e ferramentas com fins de verificação, (Nekmat, 2020; Karduni et al., 2018; Metaxas et al., 2015; Hailun et al., 2014; Wang et al., 2014; Chen et al., 2015a) e checagem de fatos baseada em multidão (Pinto

Tabela A.17: Classificação da dimensão Pragmática

| Tipo                         | Descrição                                                                                                                                                  | Quantidade |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| Avaliação de vulnerabilidade | Pesquisas investigando vulnerabilidades do usuário, como suscetibilidade, modelos de ataque baseados em manipulação de percepção, golpes financeiros, etc. | 34         |
| Dinâmica da difusão          | Aspectos relacionados a padrões de difusão de informações falsas com base em práticas de compartilhamento de usuários, como propagação de atenção.         | 38         |
| Estratégias de Intervenção   | Estratégias de intervenção relacionadas à conscientização e alfabetização testadas com os usuários ou baseadas em métricas de adoção.                      | 45         |

et al., 2019). Outra categoria eram problemas de avaliação de vulnerabilidade, como modelos de ataques de manipulação (Raman et al., 2019; Shrestha e Spezzano, 2019; Jansen van Vuuren et al., 2012; Campbell, 2001), explorando falhas cognitivas e de memória (Raskin, 2011; de Jongab, 2009), análise de nós influentes, aspectos, estratégias de limitação de influência (Bargar et al., 2019; Campan et al., 2017; Liao e Yang, 2012; Narahari et al., 2012) e discussão sobre a legislação para a responsabilidade da mídia social e a prestação de contas de campanhas orquestradas.

Uma trajetória de pesquisa na dimensão Mundo Social investiga os efeitos de informações falsas na cultura digital. Este tipo de pesquisa ajuda a compreender a cultura digital e seus grupos, observando alterações de comportamentos, como a forma como costumam compartilhar conteúdos, a novas ferramentas de verificação de informação, que beneficiam o desenho de estratégias e ferramentas para lidar com informações falsas sobre grupos distintos. Por exemplo, a transformação da comunicação política (Miletskiy et al., 2019) e mudanças na forma como o jornalismo (e os jornalistas) se comportam online (Starbird et al., 2018).

Trabalhos investigando as características cognitivas e psicológicas do usuário que afetam o consumo de informações falsas, como a suscetibilidade de consumir informações falsas (Jansen van Vuuren et al., 2012), os efeitos dos mecanismos de autenticidade no comportamento do usuário (Liao e Yang, 2012), ou mesmo as motivações para os usuários compartilharem informações falsas (Chen et al., 2015b). Por exemplo, pesquisas sobre a manipulação de sistemas de classificação que parecem ser insignificantes em janelas de tempo de longo prazo (Zhang et al., 2019b), mas que podem ser afetadas por redes de *bots* sociais e avatares, aumentando o consenso da comunidade em janelas de tempo de curto prazo (Wang et al., 2018; Yu e Han, 2018; Ross et al., 2019). Além disso, trabalhos que apresentam aspectos cognitivos do usuário exploráveis ("vulnerabilidades do usuário") (Gunawan e Suwandi, 2020) e particularidades das crenças do usuário, como ideologia política (Ross et al., 2019). Embora entender o comportamento do usuário seja essencial para desenvolver uma mitigação adequada estratégias, há um lado obscuro de tal conhecimento quando empregado sem a devida responsabilidade. Ele pode ser usado com más intenções de manipulação em campanhas de marketing direcionadas (Ahmad et al., 2019; Beskow e Carley, 2019; Miletskiy et al., 2019; Bevensee e Ross, 2018; Bandeli e Agarwal, 2018).

Em relação à responsabilidade, os resultados da pesquisa Pragmática apontam diferentes usos para a desinformação, como notícias falsas e falsificações profundas, como ataques de engenharia social sofisticados (Gunawan e Suwandi, 2020; Fraga-Lamas e Fernández-Caramés, 2020). Discussão sobre campanhas de desinformação orquestradas (Bandeli e Agarwal, 2018) e guerra de informação (Nestoras, 2018; Loui e Hope, 2017) projetadas para enganar o usuário com base em *hacks* semânticos e cognitivos (Cybenko et al., 2002; Thompson, 2003) destaca a necessidade de responsabilidade legal. Watney (2018) discute a dimensão legal das notícias falsas, a posição legal das mídias sociais e a necessidade de regulamentação estatal. Ele delibera



sobre a insuficiência das medidas de autorregulação das mídias sociais e os esforços contínuos dos governos legislando para definir as responsabilidades dos guardiões da informação.

#### A.3.5.2 *Discussão das dimensões Mundo Social & Pragmática*

A pesquisa sobre a dimensão do Mundo Social considera aspectos culturais e sociais sobre fenômeno da desinformação. Compreender as particularidades da interação de grupos, da modelagem do processo de raciocínio (Wang e Fussell, 2020) aos efeitos nos sistemas de crenças (Dimo, 2019; Jameel et al., 2019), são fundamentais para ampliar o entendimento da desinformação digital como um fenômeno. A intencionalidade e os objetivos das pessoas envolvidas em um evento desinformativo, seja a audiência ou os criadores, afetam a comunicação como um todo. Primeiro, é difícil entender as intenções enganosas por trás da desinformação. Pode incluir razões sociais e benevolentes, como mentir sobre uma festa surpresa ou mostrar pertencer a uma comunidade (Karlova e Fisher, 2013). Pode ter motivos pessoais antagônicos, como vender um dispositivo quebrado na Internet ou arruinar a reputação de alguém (Karlova e Fisher, 2013). Ainda, a desinformação deliberada pode ser criada por pura diversão (McCarthy, 2018). Nesse sentido, a caracterização das partes interessadas constitui a essência da comunicação de desinformação, sendo parte da estrutura do sistema de informação que contextualiza uma ocorrência de desinformação.

Embora crítica, a pesquisa na dimensão Mundo Social também levanta a preocupação com a responsabilidade. Em que medida as ferramentas feitas para ajudar os usuários estão ajudando? Algumas soluções promovem tecnologia potencialmente prejudicial e persuasiva. Por exemplo, mecanismos que podem promover mudanças no comportamento de consumo por meio de *nudges* (Nekmat, 2020; Bhuiyan et al., 2018), avisos e tecnologia automatizada para mitigar informações falsas (Goindani e Neville, 2020; Ookita e Fujita, 2017; Chen et al., 2015a; Wang et al., 2014). A interação humano-computador, especialmente os mecanismos de reforço intermitente positivo, tem o potencial de prejudicar os usuários, aumentando a ansiedade do usuário, perturbando a vida do usuário em outras dimensões (Kugler, 2018; Buffardi e Edwards, 2014; Bailey et al., 2001). Por exemplo, uma solução baseada em *feedback* que motiva os usuários a compartilhar a verdade, estimando as respostas que uma determinada postagem pode receber (Goindani e Neville, 2020). É possível conjecturar o cenário de autocensura de um usuário. Ferramentas que auxiliam o usuário de forma invasiva devem ter um desenvolvimento cuidadoso. Até que ponto a construção de "soluções persuasivas" está melhorando a experiência do usuário? Acreditamos que mais discussão sobre design ético e implementação de tais medidas são cruciais para construir um mundo social digital melhor com responsabilidade.

A educação também desempenha um papel na pesquisa do Mundo Social e Pragmática como uma arma para combater as informações falsas online, melhorando o comportamento de consumo de informações do usuário. Ensinar maneiras de melhorar o pensamento crítico em grupos distintos de usuários, como salas de aula ou usuários de redes sociais (Tanaka et al., 2013; Pollalis et al., 2018), é essencial para desenvolver o usuário a detectar e se envolver em atividades que esclarecem fatos e promovem um comportamento responsável.

#### A.3.5.3 *Semântica & Sintática*

Encontramos 418 artigos sobre a dimensão Semântica. Esta dimensão inclui pesquisas sobre significado, parcialidade e análise de sentimento para detecção de informações falsas. O Quadro A.18 resume as categorias.

Os exemplos são análise de características semânticas para caracterização de informações falsas (Vereshchaka et al., 2020; Oehmichen et al., 2019; Zhou e Zhang, 2004), avaliação de



Tabela A.18: Classificação da dimensão Semântica

| Tipo                              | Descrição                                                                                                                                                                 | Quantidade |
|-----------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| Detecção de informações falsas    | Trabalhos dedicados à caracterizar e identificar falsas alegações, usando aprendizado de máquina.                                                                         | 217        |
| Revisões da literatura            | Revisões relacionadas com modelos de detecção e técnicas de avaliação da qualidade da informação.                                                                         | 39         |
| Verificação automatizada de fatos | Pesquisa dedicada a verificar a veracidade das afirmações, com base em critérios semânticos como proximidade semântica entre textos, e uso de linguagem sentimental.      | 106        |
| Modelos de informação falsa       | Pesquisa Caracterizando um tipo de informação falsa, como aspectos semânticos que afetam o consumo. Exemplos são vieses políticos, discurso retórico, e falácias lógicas. | 15         |
| Modelos de Difusão                | Pesquisa Caracterizando padrões de difusão.                                                                                                                               | 21         |
| Conjuntos de dados                | Conjunto de dados rotulados, notícias falsas e multimídia desmascaradas usado para fins de <i>benchmark</i> .                                                             | 20         |

credibilidade (Hassan, 2018), avaliação de veracidade (Lal et al., 2018), avaliação de viés (Patankar et al., 2019) e (Yasser et al., 2018). Existem pesquisas sobre padrões de difusão considerando a análise de conteúdo (Budak, 2019; Zhang et al., 2019c), e pesquisas sobre análise de sentimento (Del Vicario et al., 2017; Dang et al., 2016b) e análise de significado (Huang et al., 2019; Broniatowski e Reyna, 2019; Tschitschek et al., 2018). Todas as revisões de literatura estão nesta categoria.

Encontramos 177 trabalhos na dimensão "Sintáticos". Nesse grupo estão trabalhos relacionados a estruturas formais de difusão e conteúdo de mensagens. Houveram 96 artigos sobre estratégias de detecção automatizadas com foco em alteração de imagem ou análise de nível de sintaxe. Encontramos 65 trabalhos sobre dinâmica e padrões de difusão. Existem 25 avaliações de estratégias e ferramentas de intervenção. A Tabela A.19 resume as categorias.

Tabela A.19: Classificação da dimensão Sintática

| Tipo                           | Descrição                                                                                                                                                                                                                     | Quantidade |
|--------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| Detecção de informações falsas | Modelos de detecção de informações falsas, detecção precoce, métodos baseados em sintaxe, como distância Levenshtein e vetores multidimensionais, e avaliação de credibilidade da fonte.                                      | 96         |
| Dinâmica de difusão            | Pesquisas sobre dinâmica e controle de difusão, aspectos estruturais da rede social, onde cortar links, quais nós intervir, características de campanhas concorrentes, etc.                                                   | 65         |
| Estratégias de intervenção     | Pesquisa sobre usabilidade e caracterização de estratégias e ferramentas de intervenção, como recursos de sites de verificação de fatos, fraqueza das estratégias de intervenção, minimização de lucros, <i>nudges</i> , etc. | 25         |

Neste grupo, encontramos trabalhos conduzidos em nível estrutural de mensagens, como caracterização de informações falsas (Karimi et al., 2018; Rehm, 2017), investigando critérios para avaliação de credibilidade da fonte (Pacheco et al., 2019; Fernandez-Luque et al., 2011) e detecção de contas falsas (Santia et al., 2019; Khaled et al., 2018; Kumar et al., 2014), manipulações de áudio / vídeo (Huh et al., 2018; Chen et al., 2018), bancos de dados de informações falsas (Rubin e Conroy, 2012; Kareem e Awan, 2019; Kapusta e Obonya, 2020) e modelos de infraestrutura (Rehm, 2017).

Trabalhos que entrelaçam os níveis de Semântica e Sintática investigam características de informações falsas. São 39 trabalhos modelando um tipo de informação falsa. A partir de abordagens baseadas em ontologia, encontramos modelos de desinformações que apresenta atributos de desinformação, como tipo, motivação, origem/ destino, canal de comunicação, data / hora de início, evidência e confiança (Zhou e Zhang, 2007). Outros trabalhos modelam propriedades de informações falsas de conteúdo da web, como relevância (Lin et al., 2009), ansiedade e certeza informativa (Oh et al., 2010), confiança (Mendoza et al., 2010; Cholvy, 2014), enganação (Hussain et al., 2018; Traylor et al., 2017; Rubin et al., 2015), e comportamentos e padrões de difusão (Broniatowski e Reyna, 2019; Kumar et al., 2020).

Em estratégias de detecção, categorizamos os trabalhos em dois caminhos principais: detecção precoce e verificação tardia. Os métodos de detecção precoce são aqueles projetados para detectar informações falsas o mais rápido possível para evitar que se espalhem para uma grande parte da rede social ou para impedir que infectem outros grupos (Kumar et al., 2020; Budak, 2019; Del Vicario et al., 2017). Eles consideram aspectos multidimensionais, como avaliação de relevância (Yasser et al., 2018), padrões de difusão (Kumar et al., 2020; Budak, 2019) e a pontuação de influência dos nós da rede social (Yang et al., 2019). Exemplos de estratégias de mitigação são cortar links de propagação (Ruan et al., 2015), contenção da difusão em nós influentes (Yang et al., 2019), e melhorar cascatas de informação positiva (Farajtabar et al., 2017). Os métodos de verificação mais recentes são aqueles construídos para verificar as informações que já estão se espalhando. Eles contam com a semelhança de aspectos das informações falsas mais antigas já identificadas para detectar a disseminação de novas informações falsas (Sethi e Rangaraju, 2018; Pourghomi et al., 2017; Leblay et al., 2017). Ambos consideram aspectos multidimensionais, como emoção de conteúdo (Sethi e Rangaraju, 2018), classificação de postura (Xuan e Xia, 2019), e credibilidade de perfil e conteúdo (Nilforoshan e Shah, 2019).

Também encontramos 20 conjuntos de dados sobre conteúdo de informações falsas. A diversidade de idiomas e propósitos vai do tcheco (Přibáň et al., 2019), polonês (Přibáň et al., 2019), eslovaco (Přibáň et al., 2019; Kapusta e Obonya, 2020), paquistanês (Kareem e Awan, 2019), português (Brasil) (Silva et al., 2020; Moreno e Bressan, 2019) e Arab (Alkhair et al., 2019) para Multimídia (Kopev et al., 2019), e Image Forgery Detection Dataset (Rahman et al., 2019). Havia 12 outros conjuntos de dados no idioma inglês.

#### *A.3.5.4 Discussão das dimensões Semântica & Sintáctica*

A pesquisa nas dimensões Semântica e Sintática investiga aspectos relacionados a características de conteúdo de informação falsa, das estruturas da organização virtual de grupos de usuários, padrões de difusão, e aspectos de influência para o entendimento e avaliação de veracidade. As caracterizações, geralmente provenientes dos métodos de detecção automatizados, apresentam limitações relacionados a complexidade multidisciplinar dos aspectos envolvidos na ocorrência de desinformação na comunicação. Nesse sentido, os estudos buscam determinar um conjunto adequado de características capazes de indicar uma informação falsa, a fim de melhorar a precisão de classificadores (Katsaros et al., 2019). Os métodos mais proeminentes consideram critérios multidimensionais, analisando características do perfil do usuário, como se um perfil é verificado, características de difusão, como as conexões sociais ou número de compartilhamentos em uma janela de tempo, e aspectos de conteúdo, como apelo sentimental (CORDEIRO et al., 2019). Entretanto, estes conjuntos pre-determinados de aspectos são limitadores para o entendimento da complexidade multidisciplinar, em especial dos aspectos informais da comunicação. Valores, cultura, e crenças, são aspectos informais que ficam ocultos a estas abordagens, o que acaba refletindo a tendencia de abordar aspectos técnicos que caracterizam um perfil, um conteúdo ou a difusão, mas não consideram as relações com o ambiente da

audiência. Nesse sentido, classificadores automáticos são pouco utilizados e apresentam um potencial de causar consequências prejudiciais, como alienação de bolhas de filtro e câmaras de eco ideológico (Kumar et al., 2020).

Outra consideração sobre os métodos de detecção automatizados é sua adaptabilidade a formas alternativas de conteúdo falso. Se a desinformação é intencionalmente criada para manipular, é razoável supor que as técnicas dos projetistas de desinformação se desenvolvem para evitar estratégias de detecção conhecidas (Gray e Terp, 2020). Considerando que a detecção automática utiliza modelos de informações falsas para analisar o conteúdo falso, havendo algumas limitações quanto às características contextuais, além das canônicas de desempenho e precisão (Zhou et al., 2019). Lidar com as comunicações humanas requer a compreensão do comportamento humano e das nuances informais, como cultura, valores e crenças, que moldam a forma como cada comunidade humana pensa e se comunica. Além disso, a classificação simples na dicotomia verdadeira ou falsa (verificadores de fatos e estratégias automatizadas) falha, quando os fatos por si só não convencem as pessoas, mas a retórica do apelo persuasivo ligada às estruturas de crenças de um determinado grupo social sim (Murungi et al., 2018).

Algumas pesquisas sobre a dimensão Semântica e Sintática entrelaçadas com particularidades dos modelos do Mundo Social do comportamento do usuário impactam nos usos e processos de informações falsas, como a difusão. Pesquisa investigando o impacto da verificação de fatos na propagação (Tambuscio et al., 2015), aspecto psicométrico da propagação de informação (Kumar et al., 2016a), apresentando ferramentas para rastrear a difusão de informações falsas (Metaxas et al., 2015). Embora as estratégias de mitigação de difusão apresentem um caminho promissor para lidar com informações falsas, elas precisam de mais subavaliação das dimensões da Pragmática e do Mundo Social, conforme demonstrado por de Freitas Melo et al. (2019) em um exame da insuficiência da estratégia de mitigação de mensagens no WhatsApp. Outro fator é o controle intrusivo na rede social dos usuários, como limitar a propagação cortando links (Ruan et al., 2015) e promovendo cascatas positivas de informações por usuários influentes (Wang et al., 2020) pode nem sempre ser a preocupação do usuário. Esta é uma prática semelhante à censura direcionada e a curadoria do fluxo de informações para o público.

Existe a necessidade de conjuntos de dados em idiomas alternativos ao inglês. Os resultados mostram que os conjuntos de dados em inglês são os mais proeminentes, com 12 artigos relatando sobre eles. A eficácia da estratégia de mitigação depende de um conjunto de recursos do contexto da informação falsa, como conteúdo, usuário, social e rede. A diversidade das características das dimensões Sintática e Semântica de modelos de informações falsas em linguagens alternativas aumentam as avaliações de *benchmark* de métodos de classificação automática. Além disso, as nuances dos fenômenos de informação falsa das dimensões “Pragmático” e “Mundo Social” de diferentes comunidades ampliam as ações de estratégias e ferramentas de mitigação, introduzindo perspectivas alternativas para mitigar os problemas.

#### A.3.5.5 *Empírica & Mundo Físico*

Foram 8 trabalhos na dimensão Mundo Empírico e Físico. A Tabela A.20 resume as categorias.

A categoria Empírica inclui problemas relacionados à estrutura dos signos. Avaliação de desempenho de algoritmos de detecção de última geração, como (Jeong e Park, 2019) e detecção baseada em frequência cardíaca de vídeo de multimídia falsa (Fernandes et al., 2019). Lá, há trabalhos sobre os aspectos de conexões lógicas e restrições de software para design e difusão de informações falsas (Zhu et al., 2018; Elkasrawi et al., 2016; Ruan et al., 2015). Além disso, trabalhos sobre desafios de criptografia se enquadram nesta categoria (Tyagi et al.,

Tabela A.20: Classificação das dimensões Empírica e Mundo Físico

| Tipo                                | Descrição                                                                                                                      | Quantidade |
|-------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|------------|
| Comparativo de modelos de detecção  | Pesquisa dedicada exclusivamente à comparação de desempenho de modelos de detecção automática de informações falsas.           | 6          |
| Propriedades estruturais de difusão | Estudos sobre a economia da difusão de informações falsas, como patrocinadores, e a infra-estrutura para mitigação da difusão. | 2          |

2019; Melo et al., 2019). A pesquisa do Mundo Físico é o nível mais bruto de um sistema de informação, abrangendo pesquisas sobre os aspectos de compressão de imagem com base na detecção (Nikoukhah et al., 2018).

#### A.3.5.6 *Discussão das dimensões Empírica e Mundo Físico*

Para as dimensões Empírica e Mundo Físico, há considerações a respeito do acesso de conteúdo privado em aplicativos de comunicação pessoal, como WhatsApp (de Freitas Melo et al., 2019; Tyagi et al., 2019; Melo et al., 2019). Este tipo de software de comunicação tem uma natureza diferente de uma plataforma de comunicação aberta, onde o conteúdo gerado pelo usuário possui configurações de privacidade personalizadas. As mensagens nesses aplicativos são privadas e acessá-los configuram uma invasão de privacidade, mesmo que uma máquina o faça. É do interesse do responsável pela estrutura de difusão mitigar qualquer conteúdo que possa causar danos aos usuários. Por exemplo, o Facebook e o Twitter regulam o conteúdo em suas redes sociais para evitar conselhos prejudiciais à saúde (). O WhatsApp tem diferentes requisitos para lidar com os problemas de informações falsas, como a troca de monitor de privacidade (). Além disso, os provedores de serviços de rede podem ser afetados e impactar os fenômenos de informações enganosas ().

Estudos sobre difusão e criação de conteúdo apresentam uma perspectiva das funções técnicas que afetam o fenômeno da informação falsa. Por exemplo, a criptografia torna a detecção e monitoramento da disseminação uma tarefa desafiadora (Tyagi et al., 2019; Melo et al., 2019). Além disso, o hardware e as propriedades do software afetam o processo de comunicação do usuário. Por exemplo, uma conexão de rede defeituosa pode afetar as informações recebidas e a maneira como o usuário as consome, introduzindo vieses. É necessário entender as restrições e aprimoramentos das propriedades tecnológicas nos processos de consumo e difusão.

#### A.3.5.7 *Discussão: a inter-relação entre aspectos sociotécnicos*

O Escada Semiótica (ES) permite uma visão segmentada da comunicação. Com base nessa perspectiva, podemos categorizar os aspectos sociotécnicos da desinformação em cada nível semiótico da ES. Os resultados guiam a discussão da inter-relação entre aspectos sociotécnicos em ocorrências de desinformação na comunicação, indicando possíveis caminhos para avançar no entendimento do fenômeno. Podemos notar a importância de se entender a desinformação de forma abrangente - como um fenômeno sociotécnico de modo sistêmico.

A dimensão de Mundo Social revela fatores relevantes do contexto social de uma ocorrência de desinformação na comunicação. Uma vez que aspectos sociodemográficos como desigualdade econômica, alfabetização digital ou acesso à Internet são partes estruturais do fenômeno, é preciso entender os aspectos contextuais que levam a uma ocorrência de desinformação na comunicação. Esses aspectos representam limites para o uso de tecnologias de comunicação e ao acesso à informação, e refletem a forma como uma mensagem é

interpretada/representada. Aspectos contextuais informais, como valores, expectativas, e crenças auxiliam a entender o enquadramento que uma mensagem utiliza para comunicar uma ideia e o enquadramento usado pela audiência para compreendê-la. Nesse sentido, uma ocorrência de desinformação pode ser acidental - quando não existe a intenção de enganar, mas ocorre um desvio da interpretação desejada pela parte interessada que profere a mensagem -, ou deliberada - quando a intenção é de manipular a interpretação de uma determinada situação/mensagem.

Assim, a relação da dimensão Mundo Social com a Pragmática é importante para diferenciar tipos de desinformação (Tudjman e Mikelic, 2003). Ambos os tipos podem ser compostos apenas por fatos verdadeiros, mas criados para introduzir preconceitos e manipular a percepção e a consciência do usuário em relação a uma situação (narrativas/enquadramentos) (Aigner et al., 2017; Karlova e Fisher, 2013). Algumas estratégias de propaganda apelam à vieses cognitivos que desencadeiam comportamento sociais (Volkova et al., 2017; McCombs e Shaw, 1972) e são indícios da intencionalidade da comunicação. Outros exemplos exploram a dissonância cognitiva (Larson et al., 2019; Bai et al., 2019), o uso de armas de persuasão (Varol e Uluturk, 2018), a engenharia social semântica (Java et al., 2019), *hacks* cognitivos (Cybenko et al., 2002; Java et al., 2019; Bai et al., 2019), e *hacks* semântico (Thompson, 2003). Nessa perspectiva, as informações falsas deliberada podem ser vistas como parte de um esforço coordenado para enganar, persuadir e manipular. Assim, a má intenção está implícita à aquele que cria e difunde deliberadamente uma informação falsa (da Silva Gomes e Dourado, 2019). A pragmática é uma dimensão crucial na discussão sobre responsabilização legal, de medidas de autorregulação das mídias sociais aos esforços contínuos dos governos legislando para definir as responsabilidades dos portadores da desinformação.

A relação entre as dimensões Mundo Social, Pragmática, Semântica e Sintática, os aspectos estruturais da mensagem de desinformação são particularmente importantes. Por exemplo, as características de verossimilhança e plausibilidade do conteúdo falso relacionado a fatos reais e eventos em andamento podem aumentar a credibilidade do tópico do conteúdo (Halbach et al., 2019; Rath et al., 2017) e manipular a percepção de credibilidade dos usuários (Aigner et al., 2017). Nesse sentido, a relação da temática com o ambiente da audiência é um aspecto que precisa ser considerado quando se analisa uma ocorrência de desinformação. Murungi et al. (2018) atesta que é provável que o perigo da desinformação deliberada não esteja em seu desvio do fato, mas em seu apelo persuasivo de acordo com as crenças pré-existentes de um determinado grupo social. Assim, os aspectos informais do ambiente social em que um desinformação se difunde precisam ser considerados no estudo do fenômeno, em especial para análise de aspectos semânticos como a interpretação de uma desinformação como verdade, e pragmáticos como a intenção de compartilhar uma mentira conscientemente.

Nos níveis físico e empírico estão as tecnologias utilizadas para a transmissão das mensagens. Estas dimensões inter-relacionadas com Mundo Social, Pragmática, Semântica e Sintática, revela diversos aspectos de impacto. Por exemplo, a hiper-conectividade, induzida pela computação pervasiva móvel, possibilita que as pessoas recebam informações falsas a qualquer hora e em qualquer lugar (Berghel, 2017b), em especial durante eventos significativos como atentados terroristas, acidentes e catástrofes naturais (Indu e Thampi, 2019). A comunicação móvel ou o acesso à Internet de banda larga, a forma de transmissão, seja unidirecional ou bidirecional, seja simultânea ou não, impactam na dinâmica de comunicação de informações e nas interações da audiência. A forma como as organizações de redes sociais digitais ocorrem, como uma rede móvel *ad-hoc* oportunista ou estática (Wang et al., 2008), impõe limites para difusão e consumo de desinformação. Por exemplo, o WhatsApp usa a infraestrutura de serviço móvel para interconectar nós conhecidos pelos seus *smartphones* ou um aplicativo de comunicação oportunista que constrói uma rede social quando os nós estão próximos fisicamente (Tyagi et al.,



2019). Outro aspecto é o uso de recursos digitais automatizados, como campanhas de marketing micro-direcionadas para modelagem de opinião (Zhang e Ghorbani, 2020; Lee, 2019) e o uso de bots para compartilhamento (Bastos e Mercea, 2019), que desempenham um papel no alcance e velocidade de difusão da desinformação. Nesse sentido, as propriedades dos mecanismos digitais e as mecânicas da Internet contribuem para difusão e o consumo da desinformação (Basu, 2019). Além disso, as cascatas de informação - propagação de informações em padrões de difusão encadeados pelo compartilhamento orgânico (Jameel et al., 2019) - estão ligadas a aspectos informais do contexto social e humano (Burbach et al., 2019), e as *affordances* tecnológicas dos artefatos e meios de comunicação (Starbird et al., 2019).

Aspectos físicos da tecnologia moldam o processo de comunicação, impondo limites para velocidade e alcance de difusão. Ainda, a forma de transmissão, seja unidirecional ou bidirecional, simultânea ou não, molda a interação do usuário com a fonte difusora de desinformação (Luciano, 1996). Por exemplo, um portal jornalístico ou uma página web funcionam como o rádio e a TV, comunicando a sua audiência que só pode receber informações. Por outro lado, o WhatsApp permite a troca de informações entre o interlocutor e a audiência, permitindo contestações e perguntas diretas para esclarecer pontos duvidosos do discurso. A criptografia também representa um aspecto relevante, considerando que as mensagens trocadas digitalmente podem ser criptografadas (Tyagi et al., 2019). Isso dificulta o uso de ferramentas automáticas de detecção ou de controle de difusão, além de caracterizar invasão de privacidade quando a comunicação é confidencial, como no WhatsApp (Puska et al., 2020).

Além disso, os aspectos econômicos são outra variável importante no fenômeno da desinformação online (Kshetri e Voas, 2017). As características do dispositivo físico usado para o processo de comunicação e a forma como as pessoas interagem com ele. Por exemplo, a teoria dos efeitos de longo prazo da ciência da comunicação (Wolf, 2003), preconiza a estratégia de construção gradual de significados de algum tipo de conteúdo a ser consumido pelo público, inundando uma mensagem propagandeada em diferentes formas e meios. Da mesma forma, o usuário pode receber o mesmo conteúdo desinformativo em diferentes aplicativos no mesmo ou em diferentes dispositivos de sua propriedade (Bandeli e Agarwal, 2018). A restrição ao acesso à informação jornalística ou corretiva dificulta as atividades de verificação e mitigação da desinformação (Kshetri e Voas, 2017). Por outro lado, o usuário pode estar preso a planos de acesso restritivos, dispositivos de baixo custo que tornam sua experiência de uso muito lenta, tediosa e cansativa, além de uma diversidade de aplicativos que possuem regras próprias para bloquear e distribuir conteúdo. Ainda, indivíduos influenciadores, com milhares de seguidores, divulgam desinformação recebida/discutida em outras mídias em seus próprios perfis, monetizando as visualizações que seu perfil recebe. A dimensão física engloba aspectos importantes para melhor compreender os fenômenos e os pontos cegos onde a tecnologia e o design falham.

As redes sociais reúnem informações de perfis e comportamentos dos indivíduos e as vendem a terceiros com propósitos de marketing (Bastos e Mercea, 2019). Além disso, métodos automatizados para coletar informações de redes sociais online, como *web crawlers*, podem ser usados para construir perfis de usuários e identificar grupos de interesse (Erlandsson et al., 2015). Até mesmo recursos técnicos de interface e recursos de interação às vezes revelam mais do que o necessário, expondo informações críticas do usuário a terceiros (Puska et al., 2020). Murungi et al. (2018) atesta que é provável que o perigo da desinformação deliberada não esteja em seu desvio do "fato", mas em seu apelo persuasivo de acordo com as crenças pré-existentes de um determinado grupo social. Assim, os dados pessoais e comportamentais dos usuários configuram outro aspecto contextual a ser considerado na análise de eventos desinformativos por serem impactantes na persuasão, influência e manipulação humana.



Nesse sentido, é preciso apoiar o estudo do fenômeno em relação a inter-relação de aspectos sociotécnicos e seus impactos na comunicação, em especial nas ocorrências de desinformação deliberadas ou acidentais.

#### *A.3.5.8 Relatório: QP02. Como as pesquisas abordam as partes interessadas?*

A caracterização das partes interessadas é um aspecto importante do estudo do fenômeno da desinformação que tem sido abordado de modo limitado pelas linhas de pesquisa. As pesquisas apresentam uma tendência a caracterizar as partes interessadas de modo genérico (136 artigos), com classificações como receptor e emissor, interlocutor e audiência, produtor e consumidor, ou uma referência ao papel de "usuário de rede social" que interage com as falsidades, porém com pouca profundidade (Resende et al., 2018; Traylor et al., 2017; Hassan et al., 2017). De modo geral, são duas formas com que a literatura aborda as partes interessadas:

- Modelos teóricos: incluem estudos que mapeiam caracterizações e modelos de fatores humanos, atributos e papéis das partes interessadas de outras disciplinas, como sociologia, psicologia, e antropologia (Saquete et al., 2020);
- Evidências empíricas: inclui estudos que investigam fatores humanos do fenômeno da desinformação para caracterizar as partes interessadas (Almaliki, 2019a). Por exemplo, estudos que buscam entender os critérios que determinados grupos de pessoas usam para avaliar a credibilidade de informações/fontes de informações em diferentes meios de comunicação (Saquete et al., 2020).

Modelos teóricos, Zhou e Zhang (2007) apresenta um modelo ontológico da comunicação de desinformação deliberada, e considera algumas atividades para as partes interessadas: contribuinte, curador, leitor, administrador, e analista. Entretanto, não apresenta a caracterização de aspectos informais ou a origem destes papéis. Zannettou et al. (2019) propõem uma tipologia do ecossistema de desinformações, onde descrevem diversas partes interessadas: bots, criminosos/terroristas, ativistas, governos, trolls, jornalistas, idiotas e "indivíduos que lucram com a desinformação". Embora indiquem partes interessadas, os papéis não são discutidos ou caracterizados.

Como exemplos de Evidências empíricas, estudos que buscam entender os critérios que determinados grupos de pessoas usam para avaliar a credibilidade de informações/fontes de informações em diferentes meios de comunicação (Saquete et al., 2020). Nestes estudos predomina o uso de entrevistas e questionários abordando aspectos sociodemográficos, como sexo, idade, e escolaridade (Zhang et al., 2018; Torres et al., 2018b; Fernandez e Alani, 2018), valores (Piccolo et al., 2020), motivações de compartilhamento (Chin e Zanuddin, 2019), entre outros. Outros estudos empregam técnicas etnográficas para analisar os comportamentos como, ações de compartilhamento, verificação e consumo de desinformação em grupos específicos, como estudantes e idosos, em contextos culturais diferentes (Wang e Fussell, 2020; Chin e Zanuddin, 2019; Wandoko et al., 2019; Wason et al., 2019; Torres et al., 2018b), ou investigando os impactos da homofilia e proximidade afetiva (Wu et al., 2018). Ainda, algumas abordagens empregam métodos e técnicas da Psicologia, como investigações dos padrões de consumo de desinformação baseadas na personalidade (Halbach et al., 2019) e, diferenças no comportamento de avaliação de mensagens objetivas e com linguagem emocional (Flintham et al., 2018).

Os estudos sobre desinformações deliberadas apresentam os modelos mais detalhados. Por exemplo, as partes interessadas que criam a desinformação deliberada possuem intenção de enganar, motivação para obter ganhos, recursos e conhecimento, podendo ser uma pessoa, um grupo de pessoas ou um robô (Starbird et al., 2019). A criação deste tipo de desinformação pode

incluir atividades como, planejar uma campanha de desinformação, coletar dados do público alvo, observar uma comunidade particular, preparar a desinformação para a audiência intendida, e planejar a difusão (Chen et al., 2016a; Alexander e Smith, 2010; Madhusudan, 2003).

As partes interessadas envolvidas na comunicação de desinformação possuem diferentes níveis de controle sobre a comunicação (Bro e Wallberg, 2015). "Árbitros" são partes interessadas capazes de agir como curadores de informações em diferentes meios de comunicação e para diferentes grupos de usuários (Pourghomi et al., 2017). Os árbitros podem ser sintéticos e automáticos, como algoritmos de distribuição de conteúdo, ou orgânicos e manuais, como checadores de fatos e influenciadores (Kshetri e Voas, 2017), agindo tanto para difundir quanto para mitigar. A literatura classifica três classes de árbitros: sociais, econômicos e legais (Kshetri e Voas, 2017). Árbitros sociais pertencem ao grupo social, podem ser jornais, checadores de fatos, especialistas, checadores amadores, influenciadores etc. Árbitros legais são os reguladores que reforçam as leis e definem normativas, como as organizações governamentais. Árbitros econômicos são as organizações comerciais, como microblogs (Facebook, Twitter). Ainda, árbitros populares (influenciadores) podem ser usados para ampliar o alcance e velocidade de difusão em campanhas organizadas (Starbird et al., 2019).

Quanto aos critérios considerados para caracterizar as partes interessadas, Zhang e Ghorbani (2020) apresentam três categorias como resultado de uma revisão da literatura: perfil, comportamento, e credibilidade. As características de perfil incluem nome, geolocalização, dados de registro, auto-descrições etc. As características de credibilidade incluem quantidade de seguidores, quantidade de respostas recebidas, quantidade de reações, etc. As características de comportamento englobam aspectos do contexto social, como quantidade de compartilhamentos e postagens em uma janela de tempo, câmaras ideológicas de ecô etc.

#### A.3.6 Estudos Relacionados

Estudos que abordam o fenômeno de desinformação com uma perspectiva sociotécnica são escassos na literatura revisada. Encontramos 3 estudos dedicados (explicitamente dizem abordar o fenômeno como sociotécnico) a estudar o fenômeno por esta perspectiva: uma análise da natureza colaborativa da desinformação, uma ferramenta para promover comportamentos responsáveis e coerentes em redes sociais, e a análise de vulnerabilidades e intervenções para um sistema sociotécnico. Starbird et al. (2019) analisa a natureza colaborativa de operações de desinformação orquestradas usando as lentes teóricas da estruturação (Giddens, 1986), socio-materialidade (Orlikowski, 2007) e cognição distribuída (Hutchins, 1995). Os autores empregam a metodologia *Grounded Theory* em conjunto com artefatos de análise interpretativa de conteúdo para investigar ocorrências de desinformação e construir um entendimento de *affordances* tecnológicas, estruturas sociais e ações humanas que constituem o fenômeno. Os dados são analisados com uma metodologia baseada na etnografia e os resultados caracterizam a comunicação de desinformação como operações estratégicas considerando os fatores tecnológicos, humanos e sociais. Embora a metodologia utilizada para as análises auxilie no estudo do fenômeno de modo sociotécnico, ela não foi formalizada para auxiliar na identificação de aspectos sociotécnicos, e exige experiência nas teorias e métodos de coleta e análise utilizados. Além disso, a investigação aborda as *affordances* de colaboração nos estágios de difusão, consumo e reiteração da comunicação de desinformação, deixando a criação em segundo plano.

Wang et al. (2014) propõe uma ferramenta para auxiliar o usuário a verificar uma informação, desenvolvida com base na Translucidez Social (Erickson e Kellogg, 2000). O objetivo do trabalho consiste em auxiliar o usuário a ter um comportamento social aceitável e responsável. Influenciar o comportamento do usuário é uma prática invasiva que tem um potencial danoso (Kugler, 2018; Buffardi e Edwards, 2014; Bailey et al., 2001). Além disso, a

abordagem não foi proposta para auxiliar o estudo do fenômeno de modo sociotécnico, apenas emprega esta perspectiva na elaboração de uma artefato tecnológico.

(Caulfield et al., 2019) estudam o sistema eleitoral inglês, sob a ótica dos ataques de interferência estrangeira, como o caso Brexit. O objetivo é aconselhar em possíveis intervenções e estratégias de mitigação de desinformação para sistemas sócio-técnicos. Os autores revisam a literatura de cibersegurança e elencam possíveis intervenções para proteger este tipo sistema. Os autores defendem estratégias que integrem leis, medidas anti-abuso em plataformas de comunicação, detecção e mitigação, bloqueio de financiamento de campanhas, e educação da população. Apesar dos apontamentos relevantes, o escopo da investigação é limitado a interferências estrangeiras em eleições, e não é apresentada uma metodologia que ampare a análise sistêmica do sistema sociotécnico em busca das vulnerabilidades discutidas, sendo dependente da experiência do analista.

Quanto aos estudos que abordam a caracterização da comunicação de desinformação, Traylor et al. (2017) modelam a comunicação de notícias falsas em um ciclo de vida. Os autores usam a abordagem *Grounded Theory*, analisando casos de desinformação em redes sociais, para derivar um modelo do ciclo de vida de notícias falsificadas com 4 fases: ideação, gancho (ajustes para efetividade), replicação e disseminação. Baseado nesse modelo, os autores derivam duas estratégias de enganação usadas nos casos analisados: finta e ofuscação. Ofuscação tem o objetivo de sobrecarregar o ambiente de comunicação com informações para desestimular conclusões cognitivas. Fintas são projetadas para focar o leitor em uma conclusão específica ou induzir um comportamento específico. No entanto, o modelo do ciclo não é abrangente para outros tipos de desinformação, e adota somente a perspectiva do enganador.

Shin et al. (2018) caracterizam o ciclo de vida de rumores com base em padrões de difusão: surgimento, decaimento e ressurgimento. A caracterização foca nos aspectos técnicos do fenômeno, como tempo de atividade de um perfil em uma determinada rede social e quantidade de compartilhamentos que um perfil faz, ou sociais, como a quantidade de ligações de um perfil. No entanto, a caracterização não abrange aspectos sociais informais, como crenças e vieses ideológicos de grupo. Ainda, o modelo conceitual do ciclo de vida adota a perspectiva da difusão, não abordando a criação de desinformação em detalhes.

Zhou e Zhang (2007) propõe uma ontologia que auxilia no estudo da enganação no meio digital e no desenvolvimento de soluções. Os autores consideram aspectos do canal de comunicação, da interação de partes interessadas, e da cognição humana, em um modelo que auxilia no entendimento de como a enganação pode explorar características técnicas e humanas. Embora a ontologia considere aspectos técnicos, humanos e relações sociais, o modelo desconsidera aspectos organizacionais e culturais, como o impacto de autoridades e influenciadores, e não permite abordar os usuários como consumidores ativos da desinformação. A pesquisa foi descontinuada, não sendo atualizada desde 2007.

Introne et al. (2017) emprega a perspectiva de *storytelling* para analisar os mecanismos e dinâmicas sociais que permitem uma desinformação persistir e evoluir ao longo do tempo. Os autores caracterizam partes interessadas, fatores humanos e aspectos sociais. Esta abordagem busca identificar narrativas, onde o conteúdo das mensagens segue um tema comum e se transforma enquanto é consumido. Os aspectos técnicos, humanos e sociais são identificados com base na Teoria da Retórica e Teoria da Enganação Interpessoal. Entretanto, o estudo não aborda a comunicação de desinformação de modo sistêmico, focando na difusão e consumo.

#### A.4 APÊNDICE 4 - DISSMISS FALSE INFORMATION: A VALUE MATTER

Esta seção apresenta o artigo publicado no INTERACT XVII - International Conference on Human-Computer Interaction.

## DisMiss False Information: A Value Matter

**Abstract**—The popularization of social media and the increasing consumption and dissemination of information online rise concerns on the possible impacts of disinformation on a global scale. Although relevant progress to tackle online disinformation has been made, the problem seems to be still growing in space and complexity affecting different aspects of the society, from personal relationships to entire democratic systems. In this paper, we argue for the need to understand and approach disinformation and misinformation as a sociotechnical phenomena in cultures mediated by information and communication technology, in which both universal and specific values influence the way people experience the problem. A sociotechnical perspective aware of the cultural influence can inform technical developments of user interfaces and algorithms, as well as the preparation of educational content in a systemic and socially responsible way.

**Keywords:** Disinformation, Misinformation, Social media, Human-computer Interaction, Values, Culture.

### I. DISINFORMATION AND MISINFORMATION ON SOCIAL MEDIA

Mobile computing, social media and social application have connected people and made all sort of information available as never before. By boosting the access to information and social interactions, applications such as Facebook and WhatsApp have become key to foster digital literacy and education, including in countries where social and economic inequalities are still prominent. However, if on the one hand social media and social applications have contributed to democratize the access to information and communication technology (ICT), on the other hand they are favouring misinformation and disinformation spread, catalyzing social and cultural changes, sometimes with hard and undesirable consequences.

Disinformation can be characterized as false information with the intention to mislead, to generate decision errors, manipulate individuals belief systems and for user deception [1]. The intentional aspect is mandatory and characterize the use of disinformation in different contexts on the Internet for human deception, the so called cognitive hacking [2], using human-factors and social engineering to persuade or mislead individuals. For example, it is used for spear phishing [3], luring to install malware [4] or “fake news”, hoaxes and rumors [5]. Misinformation, in turn, can be defined as misrepresented information that causes confusion and are not always intentionally created [6]. Despite the differences, both disinformation and misinformation are harmful and challenging. As a matter of simplification, in the remainder of this paper the term disinformation will refer to both.

Disinformation can be used as a weapon of social cyber attacks [7], [8], and “fake news”, hoaxes and rumors can be observed as ways to achieve it. These are frequently disseminated through social media, fraudulent sites and communication apps [5] with an impressive widespread capability.

It can be designed to deceive and shape people perception and the behaviors of great populations, used in large scales as an attempt to control, induce and manipulate the masses [7], [9], as the attempts in Brexit [10] and Cambridge Analytica cases [11]. Recently, the spread of disinformation against vaccines have reduced the results of public health programs to immunize citizens, and some diseases such as yellow fever, measles and poliomyelitis, which have been under control for decades, started infecting people again and causing deaths [12]. In Brazil, India and Mexico, for example, social applications have been used to disseminate vicious rumours that led to barbaric deaths of innocent victims [13], [14] and to influence elections [14], [15].

Although some progress has been made to stop disinformation spreading recently, the impact of disinformation seems to be reaching extreme levels, affecting relationships, freedom of expression in personal and professional domains, and threatening democratic systems. In this paper, we argue that disciplinary approaches are not enough to deal with the complexity of disinformation spread and impact, and point out to the need for understanding and approaching disinformation as a sociotechnical phenomena in cultures mediated by ICTs. In ICT cultural contexts, both universal and specific values influence the way people perceive disinformation, react to it (or not), consume and spread it. Such a sociotechnical and cultural perspective can help us to deal with the problem in a systemic way, being able to design technical solutions and promote information literacy in a more socially responsible way.

### II. VALUES AND CULTURE MATTER!

People in different countries, regions or even social groups may have their own ways to establish communication, which includes the communication process meta-factors and how they perceive and appropriate received messages [16]. Social media mediated communication is not apart from the sociocultural influence. Human values, social and economic conditions, educational level and different cultural traits seem to both affect and be affected by the way people use social applications to communicate with each other, produce, share and consume information, also online. Therefore, we argue that any initiative to tackle the disinformation problem must understand their consumption and dissemination as a sociotechnical phenomena. By a sociotechnical phenomena we mean that technical systems design, human processing capabilities, socioeconomic conditions, human values and cultural aspects must be understood and considered as intertwined to each other.

A recent socio-political event in Brazil related to formal education illustrates our view. Recently, Brazilian universities



are suffering from funding cuts and political persecution. Political instability has caused cuts of \$2.1 billion and the retain of thousands of masters and Ph.D. scholarships. These policy measures were accompanied by a strong production movement and sharing of disinformation in social media, apparently with the intent of undermining social support to universities. Since Brazilian government announced funding cuts, a wave of disinformation against public universities has been flagged. According to “Aos Fatos”, a monitoring tool developed at UFMG (Federal University of Minas Gerais), the sharing of nude student image in 350 WhatsApp groups monitored by the tool grew by at least 950% in 24 hours after the cuts [17]. Most of the shared content used images and recordings with a tone that demoralized Brazilian universities, showing naked students in random contexts, years, and situations as if they were a current and common routine in universities [18], [19]. In addition, title of dissertations and theses on gender and sexuality were shared without situating the content, using photos from other contexts and public, including other countries [18], [19]. Even old news about corruption scandals at universities that have been already solved were shared with appealing headlines, exploring confirmation bias effect on readers and promoting biased hate discussions.

The same piece of information can be used to build perspectives of two opposite ideas that appeal to the subjectiveness of individual interpretation [?], which can trigger different perceptions and consequent behaviors on judging claims, as indicated in [20], [21], [22], [23], [24]. By omitting the original context in which the images were captured and by omitting or modifying details about the research, the content have the potential to deceive the reader by distorting or ignoring the intended meaning and purpose of images and content of researches. This kind of action put in check the social media democratizing potential, promoting skepticism and outrage behaviors [13], [14], [25], [26], [23], [24].

The university founding cuts example can be used to evidences three different dimensions that must be addressed when dealing with disinformation. The first dimension is related to people's *culture and values*: the content explored nudity going against religious values, as they are very appealing to a considerable part of the Brazilian society. In Brazil, WhatsApp groups are popular among families to get in touch and share news about their lives, jokes, etc., and have become a source for disinformation as family members, usually the older and less educated, tend to disseminate any appealing content as if it was ultimately true [15]. When the information touches something valuable for people, they seem more prone to assume the information is true as if the information carry authority per se and is beyond justification [27].

The second dimension is related to *formal education and the perception of authority*. Public universities usually have very competitive and difficult admission process for students and are not able to cope with the demand of a country with more than 200 million people. Most Brazilians have never had the opportunity to experience the environment of a university. Therefore, universities are a distant reality for

the majority of Brazilians, especially the least privileged ones. Thus, people are easily influenced to believe lots of public money is being wasted by universities inefficient researches and immoral events instead of being applied for high quality education. As a consequence of the post-truth era, where objective facts influence less on the formation of individuals opinions than emotional calls and personal beliefs [28], there is a lack of trust in formal authorities to whom people may resort to get verified information from. And influenced by cognitive elements that trigger biased views, as people seem to give credibility to controversial content or to content that agrees with their previous beliefs [29], [20], [27], social platforms (e.g., WhatsApp) end up becoming a source of authority that certificates the information quality itself.

The third dimension is related to *technical issues* that range from software requirements to user interface and legal norms. If we consider the software as the media where information travels, each different program has different interaction characteristics with its own specific challenges. WhatsApp is an example of a widespread application digitally including a considerable part of population in Brazil. The application can be used in simple and cheap smartphones and is accessible to most people, even for those still not very familiar with technology. It is very easy to disseminate any content through WhatsApp groups. However, the application does not support any sort of verification of the information quality, as the authorship of content or history of the content in the platform. Another example comes as we consider the legal aspects and the security property of confidentiality. The messages on WhatsApp are personal and their content must not be accessed by any unauthorized entity. So, the access control mechanism that protects the content from unauthorized visualization creates new for manual and automatic disinformation detection tools. Besides that, for marketing purposes, mobile carriers usually do not restrict data consumption for social media applications such as WhatsApp and Facebook, but they do charge the access to other services online. Hence, people receive unlimited content in their social media groups having no means to check the content in another source, search for additional information, or even report the content to other organization than the application itself. The soil for spreading disinformation becomes fertile!

The three dimensions mentioned above: culture and values, formal and authority, and technical issues are *contextually* dependent. The cultural characteristics of a people, the values they share and the beliefs they hold heavily influence the way they understand and use technological tools for communication, like social media apps. The socioeconomic conditions of people and the formal system that regulates and guides their life also play an important role. Finally, the technical system used by people, its interface, underlying structure and algorithms will favors certain behaviors while inhibiting others depending on the cultural context in which they are being applied.

### III. TOWARDS A RESEARCH AGENDA

Existing literature in Computer Science has focused mainly on the technical dimension of the problem paying little attention to the cultural and formal ones. Although there is no arguments against the need for technical advancements to reduce disinformation spreading, starting by focusing on technical aspects in isolation may not be the most effective way. Technical advancements must be made grounded on the knowledge we gather from understanding the sociocultural context where people live, from understanding their values, preferences, needs, social norms, behavioral patterns, beliefs and demands.

Before introducing technical innovations to social groups to tackle misinformation spreading, we suggest a sociocultural contextual analysis considering, for example, whether people have access to resources and knowledge to critically evaluate pieces of information, for example, whether there are external factors or socioeconomic conditions pressuring for disseminating certain positions and intentions. In particular, investigate weather and how values and culture influence people judgment of information quality and veracity[30], which are open questions on the literature. Only starting by understanding a context and the people living in it we can grasp the particularities of their problems and design good solutions for them.

### REFERENCES

- [1] M. Tudjman and N. Mikelic, "Information science: Science about information, misinformation and disinformation," *Proceedings of Informing Science+ Information Technology Education*, pp. 1513–1527, 2003.
- [2] G. Cybenko, A. Giani, and P. Thompson, "Cognitive hacking: A battle for the mind," *Computer*, vol. 35, no. 8, pp. 50–56, 2002.
- [3] D. D. Caputo, S. L. Pfleeger, J. D. Freeman, and M. E. Johnson, "Going spear phishing: Exploring embedded training and awareness," *IEEE Security Privacy*, vol. 12, no. 1, pp. 28–38, Jan 2014.
- [4] L. Nguyen-Vu, J. Park, N. Chau, and S. Jung, "Signing key leak detection in google play store," in *2016 International Conference on Information Networking (ICOIN)*, Jan 2016, pp. 13–16.
- [5] S. Bazan, "A new way to win the war," *IEEE Internet Computing*, no. 4, pp. 92–97, 2017.
- [6] M. Tudjman and N. Mikelic, "Information science: Science about information, misinformation and disinformation," *Proceedings of Informing Science+ Information Technology Education*, vol. 3, pp. 1513–1527, 2003.
- [7] R. Goolsby, L. Shanley, and A. Lovell, "On cybersecurity, crowdsourcing, and social cyber-attack," OFFICE OF NAVAL RESEARCH ARLINGTON VA, Tech. Rep., 2013.
- [8] R. Gandhi, A. Sharma, W. Mahoney, W. Sousan, Q. Zhu, and P. Laplante, "Dimensions of cyber-attacks: Cultural, social, economic, and political," *IEEE Technology and Society Magazine*, vol. 30, no. 1, pp. 28–38, 2011.
- [9] P. Wang, R. Angarita, and I. Renna, "Is this the era of misinformation yet? combining social bots and fake news to deceive the masses," in *The 2018 Web Conference Companion*, 2018.
- [10] C. Cadwalladr, "The great british brexit robbery: how our democracy was hijacked," *The Guardian*, vol. 7, 2017.
- [11] N. Persily, "The 2016 us election: Can democracy survive the internet?" *Journal of democracy*, vol. 28, no. 2, pp. 63–76, 2017.
- [12] "Bbc," <https://www.bbc.com/news/health-46387167>, note = Last accessed May 2019.
- [13] "Bbc," <https://www.bbc.com/news/world-latin-america-46145986>, note = Last accessed May 2019.
- [14] "Bbc," <https://www.bbc.com/news/world-asia-india-47797151>, note = Last accessed May 2019.
- [15] "Folha," <https://www1.folha.uol.com.br/amp/poder/2019/05/2-em-cada-3-receberam-fake-news-nas-ultimas-eleicoes-aponta-pesquisa.shtml>, note = Last accessed May 2019.
- [16] E. T. Hall and T. Hall, *The silent language*. Anchor books, 1959, vol. 948.
- [17] "Aosfatos," <https://aosfatos.org/noticias/apos-cortes-no-mec-envio-de-imagens-de-estudantes-nus-cresce-950-em-grupos-de-whatsapp-em-24-horas/>, note = Last accessed May 2019.
- [18] "Aosfatos," <https://aosfatos.org/noticias/meme-que-critica-manifestacoes-pela-educacao-usa-fotos-de-protestos-antigos/>, note = Last accessed May 2019.
- [19] "Aosfatos," <https://aosfatos.org/noticias/foto-de-faixa-com-fora-bolsonaro-e-liberdade-para-lula-de-marco-nao-de-ato-pela-educacao/>, last accessed May 2019.
- [20] J. S. B. Evans, *Bias in human reasoning: Causes and consequences*. Lawrence Erlbaum Associates, Inc, 1989.
- [21] N. L. Kerr, R. J. MacCoun, and G. P. Kramer, "Bias in judgment: Comparing individuals and groups," *Psychological review*, vol. 103, no. 4, p. 687, 1996.
- [22] A. Tversky and D. Kahneman, "Judgment under uncertainty: Heuristics and biases," *science*, vol. 185, no. 4157, pp. 1124–1131, 1974.
- [23] M. Bedard and C. Schoenthaler, "Satire or fake news: Social media consumers' socio-demographics decide," in *Companion Proceedings of the The Web Conference 2018*. International World Wide Web Conferences Steering Committee, 2018, pp. 613–619.
- [24] M. Flinham, C. Karner, K. Bachour, H. Creswick, N. Gupta, and S. Moran, "Falling for fake news: investigating the consumption of news via social media," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 2018, p. 376.
- [25] T. Hwang, I. Pearce, and M. Nanis, "Socialbots: Voices from the fronts," *interactions*, vol. 19, no. 2, pp. 38–45, 2012.
- [26] J. Lenzer, "Fake vaccine campaign in pakistan could threaten plans to eradicate polio," *BMJ: British Medical Journal (Online)*, vol. 343, 2011.
- [27] R. S. Nickerson, "Confirmation bias: A ubiquitous phenomenon in many guises," *Review of general psychology*, vol. 2, no. 2, pp. 175–220, 1998.
- [28] A. Flood, "Post-truth named word of the year by oxford dictionaries," *The Guardian*, vol. 15, 2016.
- [29] R. Keyes, *The post-truth era: Dishonesty and deception in contemporary life*. Macmillan, 2004.
- [30] S. Y. Rieh, "Judgment of information quality and cognitive authority in the web," *Journal of the American society for information science and technology*, vol. 53, no. 2, pp. 145–161, 2002.

## A.5 APÊNDICE 5 - WHATSAPP AND FALSE INFORMATION: A VALUE-ORIENTED EVALUATION

Esta seção apresenta o artigo publicado no XIX Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais.

# WhatsApp and false information: a value-oriented evaluation

Alisson Puska  
Federal University of Paraná  
Curitiba, Paraná, Brazil  
alisson.puska@gmail.com

Luiz Adolpho Baroni  
Federal University of Paraná  
Curitiba, Paraná, Brazil  
labaroni@inf.ufpr.br

Maíra Codo Canal  
Federal University of Paraná  
Curitiba, Paraná, Brazil  
mairapdn@gmail.com

Lara S. G. Piccolo  
The Open University  
Milton Keynes, United Kingdom  
lara.piccolo@open.ac.uk

Roberto Pereira  
Federal University of Paraná  
Curitiba, Paraná, Brazil  
rpereira@inf.ufpr.br

## Abstract

Disseminated mainly through social media, online false information has triggered alarming negative impact on all aspects of life including politics, science, and education. In Brazil, WhatsApp has been pointed out as one of the main vehicles for disseminating false information, reaching millions of people all across the country. This paper presents an evaluation on WhatsApp with a values-oriented approach to investigate what aspects of this platform contributed to the false information spread in Brazil. The evaluation was conducted with 6 WhatsApp heavy users in a 2-hour session, identifying 16 different problems. Results revealed aspects of the technical infrastructure and user interface that influence the way different stakeholders can use the tool to promote false information. Finally, the paper presents prospective paths for improvements on the platform and possible ways to deal with the identified problems.

**CCS Concepts:** • Human-centered computing → User studies; Laboratory experiments; • Social and professional topics → Cultural characteristics.

**Keywords:** WhatsApp, False Information, Value-Oriented Evaluation, Fake News

## ACM Reference Format:

Alisson Puska, Luiz Adolpho Baroni, Maíra Codo Canal, Lara S. G. Piccolo, and Roberto Pereira. 2020. WhatsApp and false information: a value-oriented evaluation. In *XIX Brazilian Symposium*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org. *IHC '20, October 26–30, 2020, Diamantina, Brazil*

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-8172-7/20/10...\$15.00

<https://doi.org/10.1145/3424953.3426638>

on *Human Factors in Computing Systems (IHC '20)*, October 26–30, 2020, Diamantina, Brazil. ACM, New York, NY, USA, 10 pages.  
<https://doi.org/10.1145/3424953.3426638>

## 1 Introduction

Online false information can be found mainly as misinformation or disinformation. Misinformation is characterized as unintentional false information caused by mistakes, such as typos, biased information or persons' knowledge that happens to be wrong [7]. Disinformation is characterized as false information intended to deceive, disorient, generate decision errors, manipulate users' belief systems and deceive human beings [31]. The intentional aspect is mandatory and defines the use of disinformation as a form of human deception. Biased news and propaganda, click-bait, fake product reviews, fake news, and hoaxes are examples of disinformation.

In 2020, the world has faced a pandemic situation where false information has proliferated via social media and Information Communication Technology. The World Health Organization (WHO) spread an alert on the epidemic of disinformation about COVID-19 (Disinfodemic), which has contributed to worsen the situation in several countries [15]. In Brazil, false information about the COVID-19 pandemic have been extensively conveyed in WhatsApp groups [9]. WhatsApp has about 120 million users in the country and has been used increasingly as a means for transmitting and discussing news and advertisements [20].

A diversity of issues arises from the number of users and quite diverse context of use of WhatsApp. In Brazil, WhatsApp has been used to commit crimes, exploring users vulnerabilities for extortion purposes such as social attacks, and crimes such as false kidnaps, cloned cellphone numbers asking family and friends for money, fake sites, and phishing scams [13]. Another complicating factor is the dissemination of misinformation by authorities, as the omission of important information, health-damaging advice, press contradictions or even conspiracy theories [2, 11].

Out of context information has been used to build perspectives of two opposing ideas triggering different behaviors.

IHC '20, October 26–30, 2020, Diamantina, Brazil

Puska et al.

For example, distortion of the WHO communications introducing bias or with comic tone has been widely shared [25]. By omitting, highlighting or modifying details about what was really announced, the message has the potential to mislead the reader, distorting the intended meaning and purpose of what was originally said by the WHO. This type of action calls into question the potential for democratization of social media, promoting also harmful behaviors of groundless skepticism and outrage.

The literature in Computer Science has given strong focus on the technical dimension of online false information problems, paying little attention to the cultural and regulatory ones. The most popular approaches to deal with false information are based on technical characteristics, such as aspects of the message content, semantic analysis, models of diffusion through social networks, aspects related to social organizations in the social media environment, as *social bubbles* [14]. Approaches usually focus on English language or European behavior patterns and legal aspects [18].

Although we believe technical advances are mandatory to deal with the online false information phenomena, we think starting by focusing on technical aspects alone may not be a productive or responsible approach [26]. Our technical advances must be made on the basis of knowledge we gain from understanding the socio-cultural context in which people live, from understanding their values, preferences, needs, social norms, behavioral patterns and beliefs to design efficient solutions [23]. Such a perspective can help us to deal with the problem in a systemic and socially responsible way, from the development of algorithms and computer interfaces to the creation of artifacts to help designing software and information literacy material [24].

This paper presents an evaluation of WhatsApp application with in order to investigate issues that could promote or inhibit false information consumption and dissemination. We applied MOvalues [6], an inspection method designed to support the evaluation of social networks from the lenses of sociotechnical values. Results reveal different problems, including the lack of interface features, infrastructure conditions, and legal issues, that may influence false information consumption and dissemination in the Brazilian context. Furthermore, the paper presents a discussion on these problems and point out prospective paths for improvements on the application and aspects related to its use.

This paper is organized as follows, Section 2 addresses works on false information phenomena in the context of mobile social networks. Section 3 present the MOvalues method and the tools used on the evaluation. Section 4 shows the evaluation methodology, the participants profile, and the evaluation scenario. Sections 5 and 6 presents and discusses the results from the evaluation on WhatsApp, and Section 7 concludes the work.

## 2 Related Work

Although different similar applications have been used and investigated (e.g., Telegram, Messenger) and although false information problems are a worldwide challenge, understanding false information as a phenomena with both social and technical aspects requires us to investigate specific applications for different contexts. The literature has presented several different works related to the topics of our research. However, to the best of our knowledge, no existing research has reported results from WhatsApp evaluation in relation to values and false information especially in the Brazilian context. Therefore, the related works corroborate the need for studies to fulfill this gap, presenting different perspectives to investigate and mitigate false information on WhatsApp and similar applications.

Melo et al. [10] investigate how limitations on virality features, such as limits for message forwarding<sup>1</sup>, are effective to mitigate misinformation campaigns. The authors applied the epidemiological model of Susceptible-Exposed-Infected to estimate the virality of malicious messages in WhatsApp groups. They performed experiments using the model comparing the dissemination in different scenarios and concluded that limits on forwarding and broadcasting can reduce velocity of dissemination by one order of magnitude. However, depending on the content virality, those limits are not effective in preventing a message to reach the entire network.

Evangelista et al. [12] discusses WhatsApp as a weaponized social media. Closed, encrypted architecture, the relative anonymity afforded by only using telephone numbers as identifiers, and the limited number of members in a group are characteristics that allowed WhatsApp to become one of the most explored channels for false information. According to Evangelista, WhatsApp played a crucial role in the spread of misinformation during Brazilian elections in 2018. Of many problems involved in using a platform such as WhatsApp for an election campaign, the authors point out one in particular: the invisibility of actors who produce, monitor, distribute or direct contents.

Resende et al. [28] present a system for collecting, analyzing and viewing data from public groups on politics and news on WhatsApp in Brazil. The authors created a system to organize information and assist journalists in data analysis and fact-checking. The work provides a database on the main misinformation shared via WhatsApp public groups in the Brazilian context. The tool represents a first step to mitigate dissemination of disinformation on WhatsApp in Brazilian context. In another study, Resende et al. [27] also analyzed the properties of textual content and the propagation dynamics from political-oriented groups in WhatsApp, revealing chain messages characteristics. Their study also revealed a faster diffusion tendency of such messages into

<sup>1</sup><https://blog.whatsapp.com/more-changes-to-forwarding>

particular groups, although they take more time to spread outside the group.

Caetano et al. [5] also investigates the information dissemination process in WhatsApp groups in Brazilian context. They analyzed 120 WhatsApp public groups about random interests over one year to characterize attention cascades. They observed a tendency on false information cascades, reaching more users, and lasting longer in political groups.

Jones [17] analyzes algorithms and interfaces regarding their characteristics to perpetuate fake news spread on social media. Studying YouTube's interface, the author highlights epistemological limits of interfaces and how they affect human actions. Noting, on YouTube home page, the absence of information regarding the veracity of the content it presents, the author suggests that depending on how an interface works, the way it was designed and its purposes, it can (intentionally or not) influence human behavior by favoring spreading misinformation on the internet.

Results of solutions found in literature show that technologies are still far from causing a lasting change in the consumption of false information on mobile social networks. Therefore, we need a better understanding of communication tools and their role in the false information phenomena. In this paper, we evaluate WhatsApp to discover sociotechnical issues that may impact false information consumption and dissemination, favoring or inhibiting it.

### 3 Evaluation Method

Social media may be approached as extensions of users' real world. Each country, region or group has its own perceptions about the content of the messages. Human values, educational level, social and economic conditions and different cultural traits seem to affect and be affected by the way people use social applications to communicate, produce, share and consume information. Therefore, we argue that any initiative to address the problem of online false information must first understand its consumption and dissemination as a contextualized sociotechnical phenomenon. By sociotechnical, we mean that when approaching false information, the design of technical systems, human processing capabilities, socioeconomic conditions, human values and cultural aspects must be considered as interconnected. Our view is inspired by Stamper's Organizational Semiotics theory, which covers the technical, formal and informal dimensions of information systems [30], as well as on his 6 levels of signs where he approaches human functions and technical infrastructure in an integrated way [29].

In order to evaluate WhatsApp regarding its possible influence on false information consumption and dissemination, we adopted the Value-Oriented Evaluation Method for Online Social Networks (MOvalues) [6]. MOvalues is an evaluation method to investigate problems with values in the design and use of online social networks adopting a sociotechnical

perspective. The method supports the evaluators to look beyond the technical system, considering problems that may arise from formal (e.g. laws, regulations, etc.) and informal (e.g. beliefs, values, etc.) aspects of life in society. It considers informal aspects with a focus on human values and was designed to be applied in groups.

MOvalues helps to identify hidden aspects related to values that may be important to understand false information phenomena where traditional usability and accessibility evaluation methodologies are not intended for. Following suggestions from the literature [19, 23], the method allows one to focus on specific stakeholders and values for a specific context: values can be either understood and selected in advance, or they can be identified during the method application, emerging from the evaluation activity.

MOvalues method has three stages: 1. Preparation, 2. Individual Inspection and 3. Consolidation. As the Consolidation stage suggests, the method is intended to be applied with a team with at least 2 evaluators. Figure 1 illustrate its stages.

In the *Preparation* stage, the evaluators must gather all the relevant information to guide the evaluation activity, including discussing the application and its purpose, the evaluation focus and key values to look at. Supported by artifacts provided by the method, they must identify the stakeholders related to WhatsApp and the main values each stakeholder brings into scene. This stage must be carried out jointly by the evaluators to favor a shared understanding on the evaluation scope and purpose, preparing them for the second stage.



Figure 1. Value Oriented Inspection Method Stages

In the *Individual Inspection* stage, each evaluator explores the system interface to investigate problems with values from the perspective of each stakeholder. The method offers a list of values and guiding questions related to informal, formal and technical aspects the evaluator may consider. Although stakeholders and values have been identified in the first stage, the method allows the evaluators to include new stakeholders and values that may arise during individual evaluation. When a possible problem is identified, it is registered and assigned to a severity level.

In the *Consolidation* stage, results from individual inspections are discussed by the evaluation team resulting in a consolidated report presenting a list of problems and recommendations for improvements. For each value raised in the individual inspection, each evaluator cites the possible



IHC '20, October 26–30, 2020, Diamantina, Brazil

Puska et al.

problems identified for different stakeholders<sup>2</sup>. The evaluation team discusses whether the reported problem is really a problem or not, its impacts, related stakeholders and so on. Similar problems are grouped and merged when necessary. Finally, the problem severity must be agreed on: if there is disagreement, the greater suggested severity is adopted (the scale ranges from '0' to '4', with '0' being an unimportant problem and '4' being a catastrophic problem that requires mandatory correction). Finally, for each problem, the team must discuss how the problem can be solved, offering improvement suggestions ("Solutions").

#### 4 Evaluating WhatsApp

The evaluation of WhatsApp had as main objective to identify problems with values and to prospect their possible consequences and influence on the consumption and dissemination of false information. The evaluation was carried out by 6 participants guided by a facilitator<sup>3</sup> all from the Department of Informatics from the Federal University of Paraná: 2 Ph.D. students, 3 Master's students, and 1 undergraduate student. All the participants work in the same HCI Research Lab, have completed an HCI course and have previous experience with HCI inspection methods, such as heuristic evaluation [21] and usability testing. They also have experience with the WhatsApp application, using it daily as end users for communication purposes. The evaluation was executed in a computer lab, lasting approximately 2 hours.

The *preparation stage* involved MOvalues usage briefing and introduction by the facilitator, who also presented the evaluation tasks and the procedure to be followed by the evaluators. The 6 participants signed an Informed Consent Form. The facilitator presented the method and its context of use before the beginning of the experiment. Evaluation templates and forms to guide the evaluation and record the data obtained at each stage were also presented and explained to the participants.

A 10 minutes preparation activity to explore WhatsApp interface and interaction mechanisms was applied to familiarize the evaluation team with the perspective of users facing false information issues. The activity was designed to encourage the evaluators to think of values that could be important to different stakeholders regarding false information. Because, to the best of our knowledge, no universal set of values in the fake news phenomenon has been adopted in literature, the evaluators used the set suggested by the applied method as a starting point remaining and could include new values and issues that would arise during evaluation stages.

The preparation activities were divided into 4 tasks: 1) Share a false information content (fake news format) to a friend or group; 2) Check the veracity of the content; 3)

Search for additional information about the content shared within the application, and 4) Post the news in the participant's WhatsApp status. The evaluation team was oriented to share the false news in a group specially created for this purpose, containing the participants only, or to share the false information directly to another participant. The fake material was provided by the facilitator.

The participants were able to identify how different parties are involved in WhatsApp design and use, raising relevant values for each stakeholder. They indicated problems within the application design related to technical, formal and informal dimensions. Each problem identified by the evaluators was recorded on MOvalues Templates<sup>4</sup>.

#### 5 Results

The 6 evaluators identified 16 unique stakeholders that would influence or suffer influence from false information on WhatsApp. The evaluation team discussed 21 values related to the different stakeholders and decided to prioritize 9 values in the consolidated report. Examples of stakeholders and their values are: users valuing trust, press valuing meta-communication, marketing companies valuing identity, and government valuing security, showing the diversity of stakeholders with different values they care about. Considering the "<stakeholder, value>" tuple, 40 problems were identified, and from those, 16 unique problems concerning the technical, formal e informal dimensions were included into the consolidated report. There were similar problems related with the same dimension but for different "<stakeholder, value>" tuple, indicating the importance to consider individual subjectivity when evaluating a social network. Moreover, the participants were able to suggest 13 different improvements and strategies of mitigation for the 16 identified problems.

##### 5.1 Stakeholder analysis and their values

In the first evaluation stage, the evaluation team identified 16 stakeholders, indicated the relevant values for each one and discussed why each value was important for each stakeholder. When assigning relevant values to them, the participants focused on 12 stakeholders with 21 different values.

MOvalues uses an adaptation [22] of the Stakeholder Identification Diagram from Organizational Semiotics [29] where possible stakeholders are related to one or more of four categories: contribution, source, market and community and bring values with them. In the *Contribution* category there are actors directly involved in the design and use of social network, stakeholders who use it directly and those that are decisive in its existence. The *Source* category corresponds to customers and suppliers of resources and information

<sup>2</sup>Alternatively, for each stakeholder, each evaluator may cite the possible values and problems related to them

<sup>3</sup>The third author who created the MOvalues method

<sup>4</sup>The website with instructions, as well as the evaluation data, may be accessed here: <https://docs.google.com/spreadsheets/d/1UU503LYdEpsYVfOi1xLM58EcNzf9SIXSGE8UrsUfCao/edit?usp=sharing>.

WhatsApp and false information: a value-oriented evaluation

IHC '20, October 26–30, 2020, Diamantina, Brazil

**Table 1.** Stakeholders Identified

| Contribution | Source      | Market    | Community  |
|--------------|-------------|-----------|------------|
| Users        | Advertising | Telegram  | Justice    |
| Legal Ent.   | Press       | Instagram | Government |
| Analysts     |             | Messenger | Companies  |
|              |             | Skype     | Family     |
|              |             | Service   | Friends    |
|              |             | Providers | Workmates  |

for social networks, people who will benefit from the produced information. The *Market* category associates partners of WhatsApp as well as its competitors, such as other similar social applications. Finally, the *Community* category refers to the community in general, those that do not necessarily use the social network but may influence or be impacted by it, such as viewers and legislators.

For stakeholders that directly contributes to WhatsApp design and usage, the evaluators indicated "Users", "Legal Entity", and the "Analyst" as members of this group. "Advertising" (Marketing) and "Press" were identified as member of Source category, as those that produce content for the network or that get some benefit from the information produced on it. For the Market category, they indicated "telegram", "Instagram", "Messenger", "Skype", and "Mobile Phone Operators", representing applications or media that dispute the user and offer the same type of service. Finally, for the Community category, the evaluators indicated people from users' close relationship to the government. Examples are "Justice", "Government", "Workmates" and "Family". Table 1 present the stakeholders in their respective categories.

The evaluators indicated 21 important values for 13 stakeholders, presented with their motivation for inclusion. For example, the stakeholder *User* valuing *Awareness* for "feel like you're in control", and valuing *Privacy* because "Content is important only for interested parties of the conversation". *Instagram Direct* valuing *Storage* "to recover old conversation", and *Press* valuing *Reputation* as "a press vehicle is considerably reliable from its reputation", and so on. Table 2 present the amount of values by stakeholders.

Below, we present 9 values, the ones appearing in the consolidation report and their meanings according to [6]. All the 21 values selected by the evaluators are available with the complementary material, and the complete list with 35 values provided to them, a description, categorization and examples of use for each value can be found in [6].

- **Awareness:** Perception about the system, about what it is doing and who is doing what, such as notification updates. For the evaluators, this value was selected because the "User" must "feel that the application is under his/her control".

- **Availability:** The application must always be available to the user. "In Brazil, there have been several times when WhatsApp was taken offline causing several users to look for alternatives (e.g., Telegram)", the evaluators agreed.
- **Trust:** Refers to the extent to which the system, or other users using the system, behave as expected by a particular person. For the evaluators, it is an important value because "information shared via the app can build trust, towards the app, for the one doing it".
- **Emotion and Affection:** People's feelings, such as well-being, fun, engagement, annoyance, and other related aspects to the user experience. The evaluators suggested that the material produced by the "Press" may be influenced by the emotional relationship of the readers, which may influence consumption.
- **Groups:** A set of users with characteristics, situations, interests or purposes in common. For the evaluators, "friends" and "families" are groups of people with similar interests, which is a important value that may impact information consumption.
- **Innovation:** Refers to application updates, such as interfaces and features. For the evaluators, for competitive purposes, it is necessary that "developers" are always providing innovative ways to build the platform, and improvements based on the needs of the users, as tools to tackle false information.
- **Property Ownership:** The right to own an object or information, and the right to the actions that can be performed on / through that object. For the evaluators, for the "press" the authorship is extremely important because it reflects with the vehicle's identity and for accountability purposes.
- **Sharing:** Possibility for users to make objects or information in their possession available to other users. For the evaluators, the purpose of being an online social network is to share content created by users, which is important to WhatsApp and its competition.
- **Security:** It refers to how well the system protects the information it contains, whether from external attacks or possible technical failures. For the evaluators, it is important for the user to configure restrictions to avoid abuses, such as personal information leakage.
- **Informed Consent:** Situation of an individual's awareness and agreement about possible actions to be taken and the impacts of these actions. The evaluators consider "it is important for the User because it is possible to receive messages from people we don't know".
- **Privacy:** A claim, authorization or right of users to determine what information about them will be available and who has access to that information. For the evaluators, it is an important value for the "User" because content transmitted within WhatsApp conversations are of personal and private nature.

IHC '20, October 26–30, 2020, Diamantina, Brazil

Puska et al.

**Table 2.** Values by Stakeholders

| Advertising | Legal Entity | Analysts         | User        | Instag. Direct      | Press                | Friends       | Government              | Others       |
|-------------|--------------|------------------|-------------|---------------------|----------------------|---------------|-------------------------|--------------|
| Scalability | Awareness    | Scalability      | Scalability | Sharing             | Emotion & affection  | Groups        | Historic                | Availability |
| Historic    | Privacy      | Interoperability | Visibility  | Storage             | Property & ownership | Collaboration | Norms, rules & politics | Discovery    |
| Identity    | Conversation | Innovation       |             | Conversation        | Reputation           | Sharing       | Collaboration           | Trust        |
|             | Groups       |                  |             | Emotion & affection |                      | Conversation  |                         | Reputation   |
|             |              |                  |             | Trust               |                      |               |                         |              |
|             |              |                  |             | Security            |                      |               |                         |              |
|             |              |                  |             | Groups              |                      |               |                         |              |

## 5.2 Consolidation

In the preparation stage, the evaluators identified 16 different stakeholders. From those, they selected 12 stakeholders most representatives and 21 important values. Following that, the user inspection stage led the evaluators to give their opinion on possible problems related to values in WhatsApp according to the informal, formal and technical dimensions of its usage, operation and impact. These levels incite designers to think beyond technical issues, paying attention also to the social ones. From the individual inspection stage, 40 problems related to 13 values and 5 stakeholders were identified by the evaluators. For the consolidation stage, the evaluators discussed the results and generated the final report, resulting in 16 problems on 10 different values related with 4 stakeholders presented in Figure 2.

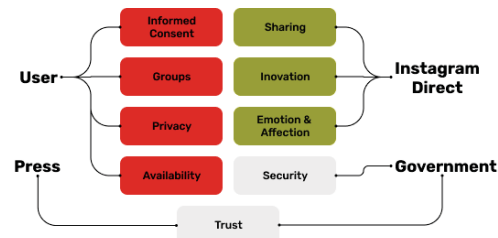
From the individual inspection stage, 18 (45%) problems were related to the informal dimension, 11 (27,5%) to the formal and 11 (27,5%) to the technical one. Examples of Informal problems raised for the value "Awareness": "Any type of content (photos of naked children, false news, bodies in accidents, etc.) external (link, image, news) can harm the user, cause discomfort when viewing images or fall into financial donation scams". Another perspective stated that "I may share false/incorrect information that affects my image (reputation) in other circles.", concerning the "Awareness" value.

For the Formal dimension, "Awareness" have problems like "The lack of legislation and rules to define the type of content allowed to the user, preventing them to violate moral and act in an unethical way.". For the Technical dimension, "The lack of warning and inspection mechanisms to help the user not to view sensitive or invalid / unverified content" was pointed as a problem.

The consolidation activity had a fixed time for execution. For this reason, the evaluators decided to focus on values related to the interested parties "User", "Instagram", "Government" and "Press". The next step was to consolidate the problems through a group discussion where each problem identified in the individual inspection was discussed and classified by their severity degree. They also suggested improvements that could mitigate them.

The evaluation report presented 16 problems resulting from the consolidation stage for the 4 selected stakeholders. From those, 1 catastrophic problem, 6 serious problems, 5 small problems and 1 cosmetic problem. From the 16 problems, 7 were related only to the user stakeholder, 5 about a WhatsApp competitor, 3 regarding the government, and 1 about the Press.

After applying the method, the evaluators discussed their conclusions about the method and the obtained results. They were asked about the difficulties and ways to assess false news and how WhatsApp could be improved to assist and protect the considered stakeholders. Table 3 presents an overview of the main results from the consolidated report.

**Figure 2.** Values by Stakeholders - Consolidation Stage

## 6 Discussion

Results from WhatsApp evaluation indicate that the application design presents several problems related with false information consumption and dissemination. These problems vary from technical characteristics, such as privacy features (problem #1), as well as formal issues, such as the lack of laws to regulate false information within the application (problem #3), to informal issues of life in communities (problem #12), as jokes and irony that may be considered as true, or about its quite different uses. More than technical problems (11 problems), the total of informal (18) and formal (11) problems suggests non-technical issues play important roles on online false information phenomena. It shows the

WhatsApp and false information: a value-oriented evaluation

IHC '20, October 26–30, 2020, Diamantina, Brazil

**Table 3.** Consolidated Report

| Prob. ID | Stakeholder | Value               | Problem                                                                                                                      | Severity | Solution                                                                                                                                                                            |
|----------|-------------|---------------------|------------------------------------------------------------------------------------------------------------------------------|----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| #1       | User        | Privacy             | "When added to a group, my number is available to other members of the group, including those I don't know".                 | 3        | "profile levels that can show certain information (name, number, photo) only to friends already added. Or not show the photo and number by default to unknown people only the name. |
| #2       | User        | Informed Consent    | "It is possible to receive messages from people we don't know".                                                              | 3        | "session request / message permission".                                                                                                                                             |
| #3       | User        | Availability        | "Legal problems can make the system unavailable".                                                                            | 3        | "clear legislation on the unavailability of the system, do not leave it unavailable for any reason. Give people time to prepare for the lack of service.                            |
| #4       | User        | Groups              | "It is not possible to prevent people I don't know from adding me to groups without my permission".                          | 2        | "show new WhatsApp group session request".                                                                                                                                          |
| #5       | User        | Groups              | "In many WhatsApp groups, there is no way to place restrictions on certain types of people".                                 | 2        | "add private group option; put administrators to check who enters".                                                                                                                 |
| #6       | User        | Groups              | "It is forbidden to post fakes news, or obscene images, but there is no option to report such attitudes in the application". | 3        | "create message reporting option / functionality".                                                                                                                                  |
| #7       | User        | Groups              | "WhatsApp groups have a lot of information, and it gets very polluted".                                                      | 1        | "group post types; add filter types (text only, video only, so-and-so)".                                                                                                            |
| #8       | Instagram   | Sharing             | "Pedophilia- sharing images of children with sexual content".                                                                | 4        | "Develop tools to detect or block this kind of content".                                                                                                                            |
| #9       | Instagram   | Innovation          | "How to interact with the app and possibilities to start a conversation - features easily run on Instagram Direct".          | 2        | "whatsapp has to be updated to not loose users".                                                                                                                                    |
| #10      | Instagram   | Sharing             | "publications can influence people's decisions".                                                                             | 3        | Literacy. Veracity check (the responsibility rests with the user).                                                                                                                  |
| #11      | Instagram   | Sharing             | "buying likes."                                                                                                              | 3        | Identify recent accounts involved in inorganic likes. Analysis of activity behavior.                                                                                                |
| #12      | Instagram   | Emotion & Affection | "Publications can 'dictate' the standard of beauty".                                                                         | 2        | "Literacy. Content label and tips".                                                                                                                                                 |
| #13      | Government  | Security            | Viral "fake news can generate crisis of violence in the real world."                                                         | 2        | No suggestion.                                                                                                                                                                      |
| #14      | Government  | Security            | "Critical information leakage by the hacking of government posts WhatsApp accounts".                                         | 3        | No suggestion.                                                                                                                                                                      |
| #15      | Government  | Confidence          | "Targeted marketing campaigns may cause political instability".                                                              | 3        | No suggestion.                                                                                                                                                                      |
| #16      | Press       | Trust               | "Readers' trust for shared data".                                                                                            | 2        | "Notification or label of verified info".                                                                                                                                           |

0 - Unimportant Problem, no correction; 1 - Cosmetic Problem, occasional correction; 2 - Small Problem, low priority correction;  
3 - Serious Problem, high priority correction; 4 - Catastrophic Problem, obligatory correction

need to consider sociotechnical aspects from the first phase of the design to produce an artifact in a socially and ethically responsible way. An example of problem related to formal issues is the *lack of legislation which could curb the creation and sharing of disinformation*. An example of problem related to informal issues is the *possibility of sending sensitive content in groups, from car accidents or child pornography to the sale of drugs, with poor accountability, promoting illegal or harmful behavior*.

The free nature of the internet allows the *creation of all sort of contents, from defamatory messages to scams that corroborates with extortion and manipulation of users*. The evaluators discussed the possibility to use autonomous detection systems to tackle false and manipulative information (problems #13 and #15). Since messages are encrypted from end to end,

automatic monitoring becomes a costly activity. Moreover, from a formal perspective, accessing the content of a private communication message is an invasion of "Privacy" covered by legislation in several countries, whether this access occurs by a machine or a human.

The evaluators discussed several privacy issues such as *forwarding a message without authorization or being added into a group conversation without permission* (problem #1). A message communicated inside a particular conversation and, then, shared outside the scope of that conversation (e.g. financial data in a group/ financial data in a particular communication) may promote gossip and misinformation. In addition to the violation of users' "Privacy", such action may implicate in issues with other values. For instance, "Property

IHC '20, October 26–30, 2020, Diamantina, Brazil

Puska et al.

and Ownership" as users have rights over their private information; "Trust", impacting users relationships with other users or influencing the way they uses WhatsApp; "Security" as users' information may be shared without their permission; and "Informed Consent" as information can be shared without user acknowledgment and approval.

Although communication within WhatsApp is usually considered personal, even if in a group where several users interact, sharing pieces of messages or being add into group conversations are common practices. These actions impact formal issues, such as Carolina Dickmann's Law[1] in Brazil that covers issues related to private information sharing, and informal issues, such as biased information that causes confusion or embarrassment. For "Security" and "Informed Consent", improving access control mechanisms, such as "asking users explicit consent before adding them into a group", or adopting "a policy that allows only known users to be added into a group" could soften the problem. Other proposition were simple technical improvements, such as "authenticate content with signatures when created", which promotes accountability, "Trust" and "Awareness", or "authorization to receive messages from unknown numbers", for a better "Informed Consent" value. All these could discourages, or even mitigate, informal abusive behavior.

Technical, formal and informal issues interconnect different values. "Privacy" of user sensitive data (Phone number) appearing in metadata of groups, for everyone, is linked with "Informed Consent", as observed in problems #1 and #2. Figure 3 and Figure 4 illustrate the problems. Thus, values can conflict with each other, as "Property and Ownership" value that is important for information credibility, and accountability purposes, can conflict with the "Privacy" value of users as anonymity must be maintained in some cases (users' address can be found out via their phone number). Moreover, at the same time, values can enhance each other: if a message has users' explicit authorization to be shared ("Informed Consent"), it may enhance the value "Property and Ownership".

These results illustrate how problems are related to both informal, formal and technical issues, sometimes in intertwined ways. For example, because any user can be included in a group (problem #4), people may have personal data unduly exposed and receive communication unwillingly, which may lead to formal issues when unasked/unwanted messages have illegal content. Informal implications are also present, for example, as users' phone number can be used outside WhatsApp on pranks, receive spam or sale calls, or even to commit crimes. Also, this problem may be the gateway for other information related attacks, such as using social engineering to build a profile of the user, or lure to install malware [8]. A rather simple aspect of the system "Privacy" can cause discomfort for the "User" in different dimensions and can trigger impacts difficult to be identified and measured.

Indeed, more than being difficult to be identified and measured, problems related to values produce impacts that can



Figure 3. Example of "Privacy" problem: Contact information visible for any user in a group.

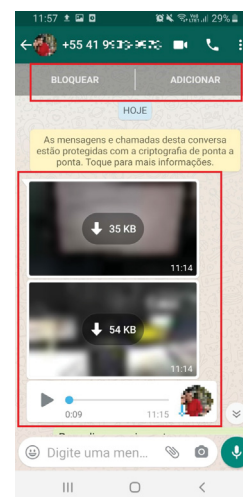


Figure 4. Example of an "Informed Consent" problem: Non-requested content received and visible.

hardly be undone or corrected. Impacts on reputation, privacy, security, safety, social relationships, personal and professional groups, emotional health cannot be undone either by a technical feature or a formal action. Indemnities and apologies do not erase the effects as well. When problems with values have occurred, there not much to do but managing the disaster. This is why paying attention to values is essential when designing and evaluating social systems.

Understanding what is valuable for the different stakeholders involved in a system design and usage is essential for a socially responsible design. Privacy as a value has different importance and scope for different interested parties. From ethical implications to necessary procedures, dealing with



trade-offs, such as the privacy-monitoring one, is challenging and requires us to understand the different forms that different interested parties value both privacy and monitoring. Such an understanding can help us to design features that allow interested parties to be aware and informed on whether, what and how monitoring is happening, and about what and how privacy can be ensured. To consider "Informal" and "Formal" problems require careful debate, pondering the consequences of extreme actions not desirable, as censorship.

The evaluators reported difficulties in investigating the authenticity and veracity of a received message. Verifying message authenticity is not easy for users as messages do not have a signature when created, and the mere act of access it may be invasive. *The only information perceived by users about the authenticity of the message is an interface symbol indicating whether the message has been forwarded.* By itself, this design decision can hardly raises users' concern about the origin or legality of the received message. Thus, they suggested technical modifications on WhatsApp interface to make available hints and meta-information of a message track, along with a regulatory legislation that may inhibit gossip and make users aware and accountable for their actions.

Furthermore, WhatsApp does not provide resources to investigate information within the platform. *When something is shared there is nothing that can be done within the application to assist users when verifying content veracity (problem #6).* The only available resource is to report a message to the application's "bugs area" — but even this is not easy to find.

The *violation of network neutrality* is also a problem. In Brazil, cellphone companies provide Internet access plans where WhatsApp traffic is not charged even when users does not have data allowance. Thus, users may receive false information but even if they are concerned about that, they are unable to access other resources to verify or seek for additional information. Considering that this kind of product is most used by people with lower economic conditions (usually the ones with lower educational levels), WhatsApp becomes a kind of one-way road for spreading false information to the most vulnerable part of the Brazilian population.

Mass messaging features available for business profiles are also problematic. There is concern that mass advertising spread misinformation with political bias or marketing campaigns. Hate speeches can encourage violent behavior and promote psychological, moral and physical attacks on users. The flashy and uninformative "headlines" can promote misinterpretation of a fact, especially when users are unable to access external links to WhatsApp. This may be a catastrophic problem when we consider that WhatsApp had influence on Brazilian elections [12].

The evaluators mentioned different ways to check the authenticity of news, from considering trust in the person who shared it to using external tools such as fact-checking sites. *An evaluator stated that trust in the person is the only way used to check the news, but the other the evaluators argued*

*this is enough and prefer to check other sources taking into account their origin and credibility.* An evaluator suggested using reverse image search to verify whether it originates from a reliable source or whether it has been manipulated or presented out of context, distorting its original meaning. Whatever the strategy, they require a certain level of digital literacy and user experience, which may not be generalized as WhatsApp users' common skills.

For WhatsApp groups, administrators should have control to define that only allowed people would enter a certain group, which could reduce the number of undesirable members sharing false information. The possibility to block a certain person in a group was also taken into account, given that certain users can share content that is considered offensive. Additionally, people must have control over what others can do with their data, deciding who can view their data and who can send them messages or invitations.

Another critical problem is the fake personification of entities where someone impersonates other people, stealing their identity and gaining the trust of their contact network [3]. This problem may have catastrophic consequences as extortion schemes exploring this vulnerability, such as fake kidnaps and jackpot scams, that may result in financial and psychological damage to the victims [16].

Another critical problem is the identity theft where someone impersonates other people, gaining the trust of their contact network [4]. This problem may have catastrophic consequences as extortion schemes exploring this vulnerability, such as fake kidnaps and jackpot scams, that may result in financial and psychological damage to the victims [16]. Even more complex attacks, as Sybil Attacks in social networks, can generate bad-mouth and opinion bias. WhatsApp mitigation strategies require user familiarity with technology, which may not always be the case.

*The evaluators also suggested solutions of literacy for the problems related to privacy and security (problems #10 and #12),* such as user manipulation or extortion, as instructed users tend to be aware of the possibilities of scams and manipulation. Literacy may also be a non-invasive way to deal with privacy value and online false information and, although literacy alone cannot solve the problem, without literacy we cannot solve it effectively either.

Problems on the false information phenomena reach technical, formal and informal issues, increasing the complexity to solve them with disciplinary isolated approaches. Problems range from the technical, such as identifying the origin of a message, passing through the formal, as legal aspects for effective accountability of disinformation authorship and diffusion, to the identification of linguistic patterns of disinformation that can support automatic solutions or the design of fair and responsible content distribution algorithms. Even solutions that consider technical and formal issues still need to improve their efficiency, effectiveness, adaptability and scalability to meet the needs on a global scale.



IHC '20, October 26–30, 2020, Diamantina, Brazil

Puska et al.

## 7 Conclusion

This paper presented an evaluation of WhatsApp investigating problems with values that may impact the consumption and dissemination of online false information. A value-oriented inspection method was applied by six evaluators in a 2-hours session. Results show WhatsApp has different problems allowing the dissemination of dis/misinformation, and may cause real damage on users and the society. A rather simple example is sharing part of a private message out of the context, causing bias and even damaging user reputation.

Our results and discussions expose the need for a perspective that considers both social and technical aspects when designing and evaluating social interaction tools. More than revealing different problems related to informal, formal and technical issues, the paper presents discussions and insights towards possible ways to deal with them.

## References

- [1] 2019. Brazilian penal code - Criminal classification of computer crimes. [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2012/lei/l12737.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12737.htm) [accessed 25. Jun. 2020].
- [2] 2020. Epidemia de fake news | Revista Pesquisa Fapesp. <https://revistapesquisa.fapesp.br/epidemia-de-fake-news> [accessed 16. Jun. 2020].
- [3] 2020. WhatsApp FAQ - Stolen accounts. <https://faq.whatsapp.com/general/account-and-profile/stolen-accounts/?lang=en>
- [4] Kayode Sakariyah Adewole, Nor Badrul Anuar, Amirrudin Kamsin, Kasturi Dewi Varathan, and Syed Abdul Razak. 2017. Malicious accounts: Dark of the social networks. *Journal of Network and Computer Applications* 79 (2017), 41–67.
- [5] Josemar Alves Caetano, Gabriel Magno, Marcos Gonçalves, Jussara Almeida, Humberto T Marques-Neto, and Virgílio Almeida. 2019. Characterizing Attention Cascades in WhatsApp Groups. In *Proceedings of the 10th ACM Conference on Web Science*. 27–36.
- [6] Maira Codo Canal and Roberto Pereira. 2020. MOvalues: A Value-Oriented Method for Evaluating Online Social Networks. In *XIX Brazilian Symposium on Human Factors in Computing Systems*.
- [7] Giovanni Luca Ciampaglia, Alexios Mantzarlis, Gregory Maus, and Filippo Menczer. 2018. Research challenges of digital misinformation: Toward a trustworthy web. *AI Magazine* 39, 1 (2018), 65–74.
- [8] George Cybenko, Annarita Giani, and Paul Thompson. 2002. Cognitive hacking: A battle for the mind. *Computer* 35, 8 (2002), 50–56.
- [9] Ministério da Saúde. 2020. Fake News. <https://www.saude.gov.br/fakenews> [accessed 16. Jun. 2020].
- [10] Philipe de Freitas Melo, Carolina Coimbra Vieira, Kiran Garimella, Pedro OS Vaz de Melo, and Fabricio Benevenuto. 2019. Can WhatsApp Counter Misinformation by Limiting Message Forwarding?. In *International Conference on Complex Networks and Their Applications*. Springer, 372–384.
- [11] Estado de Minas. 2020. Bolsonaro promete reunião com mulher que aponta enxofre como a cura da COVID-19. *Estado de Minas* (Jun 2020). [https://www.em.com.br/app/noticia/nacional/2020/06/08/interna\\_nacional,1155021/bolsonaro-promete-reuniao-com-mulher-que-tem-a-cura-da-covid-19.shtml](https://www.em.com.br/app/noticia/nacional/2020/06/08/interna_nacional,1155021/bolsonaro-promete-reuniao-com-mulher-que-tem-a-cura-da-covid-19.shtml) [accessed 26. Jun. 2020].
- [12] Rafael Evangelista and Fernanda Bruno. 2019. WhatsApp and political instability in Brazil: targeted messages and political radicalisation. *Internet Policy Review* 8, 4 (2019), 1–23.
- [13] Felipe Souza @felipe\_dess. 2019. Veja quais são os golpes mais comuns no WhatsApp e como se proteger. *BBC News Brasil* (Nov 2019). <https://www.bbc.com/portuguese/geral-50294962>
- [14] Miriam Fernandez and Harith Alani. 2018. Online misinformation: Challenges and future directions. In *Companion Proceedings of the The Web Conference 2018*. 595–602.
- [15] <https://plus.google.com/+unesco>. 2020. Combating the disinfodemic: Working for truth in the time of COVID-19. *UNESCO* (Apr 2020). <https://en.unesco.org/covid19/disinfodemic> [accessed 26. Jun. 2020].
- [16] Julia Jacobo. 2019. Authorities warn of fake kidnapping scam targeting parents in Cincinnati area. *ABC News* (Dec 2019). <https://abcnews.go.com/US/authorities-warn-fake-kidnapping-scam-targeting-parents-cincinnati/story?id=67800579> [accessed 26. Jun. 2020].
- [17] L. Corinne Jones. 2019. Listening to lies and legitimacy online: A proposal for digital rhetorical listening. *First Monday* 24, 3 (2019).
- [18] Georgi Karadzhov, Pepa Gencheva, Preslav Nakov, and Ivan Koychev. 2018. We built a fake news & click-bait filter: what happened next will blow your mind! *arXiv preprint arXiv:1803.03786* (2018).
- [19] Carla Leitão, Cristiano Maciel, Lara Schibelsky Godoy Piccolo, Luciana Salgado, Patricia C de Souza, Raquel Prates, Roberto Pereira, and Vinicius Carvalho Pereira. 2017. Human Values in HCI: a challenge for the GrandIHC-BR. In *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems*. 1–6.
- [20] Nic Newman, Richard Fletcher, Antonis Kalogeropoulos, and Rasmus Nielsen. 2019. *Reuters Institute digital news report 2019*. Vol. 2019. Reuters Institute for the Study of Journalism.
- [21] Jakob Nielsen and Rolf Molich. 1990. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 249–256.
- [22] Roberto Pereira and Maria Cecilia Calani Baranauskas. 2015. A value-oriented and culturally informed approach to the design of interactive systems. *International Journal of Human-Computer Studies* 80 (2015), 66–82.
- [23] Roberto Pereira, Maria Cecilia Calani Baranauskas, and Kecheng Liu. 2018. An essay on human values in HCI. *SBC Journal on Interactive Systems* 9, 1 (2018), 4–16.
- [24] Lara SG Piccolo and Roberto Pereira. 2019. Culture-based artefacts to inform ICT design: foundations and practice. *AI SOCIETY* 34, 3 (2019), 437–453.
- [25] Gustavo Maia e Leandro Prazeres. 2020. Bolsonaro distorce fala da OMS e volta a pregar retorno ao trabalho. *O Globo* (Mar 2020). <https://oglobo.globo.com/brasil/bolsonaro-distorce-fala-das-oms-volta-pregar-retorno-ao-trabalho-24340661>
- [26] Alisson Puska, Lara Piccolo, and Roberto Pereira. [n.d.]. DisMiss False Information: A Value Matter. *Human Computer Interaction and Emerging Technologies: Adjunct Proceedings from* ([n. d.]), 75.
- [27] Gustavo Resende, Philipe Melo, Julio CS Reis, Marisa Vasconcelos, Jussara M Almeida, and Fabricio Benevenuto. 2019. Analyzing textual (mis) information shared in WhatsApp groups. In *Proceedings of the 10th ACM Conference on Web Science*. 225–234.
- [28] Gustavo Resende, Johnnatan Messias, Márcio Silva, Jussara Almeida, Marisa Vasconcelos, and Fabricio Benevenuto. 2018. A system for monitoring public political groups in whatsapp. In *Proceedings of the 24th Brazilian Symposium on Multimedia and the Web*. 387–390.
- [29] R.K. Stamper. 1993. A semiotic theory of information and information systems. In *Proceedings of the Joint ICL/University of Newcastle Seminar on the Teaching of Computer Science, Part IX: Information*, B. Randell (Ed.). University of Newcastle, Australia, 1–33.
- [30] Ronald Stamper, Kecheng Liu, Mark Hafkamp, and Yasser Ades. 2000. Understanding the roles of signs and norms in organizations-a semiotic approach to information systems design. *Behaviour & information technology* 19, 1 (2000), 15–27.
- [31] Miroslav Tudjman and Nives Mikelic. 2003. Information science: Science about information, misinformation and disinformation. *Proceedings of Informing Science+ Information Technology Education* 3 (2003), 1513–1527.

## A.6 APÊNDICE 6 - DESENVOLVIMENTO DO MODELO CONCEITUAL DO CICLO DE VIDA DA DESINFORMAÇÃO DIGITAL

A análise e o projeto de sistemas de informação devem começar com a compreensão e modelagem da organização primeiro (Liu, 2000). Assim, com base no Sistema de Informação Humano (SIh) de uma organização, Stamper propões um artefato para auxiliar a compreensão das organizações e a modelagem de SI's. A Cebola Organizacional (CO) é um artefato da SO que possibilita uma perspectiva ampla para análises da organização. Ela divide a arquitetura da organização nas camadas do SI técnico, formal e informal, relacionando os atributos de cada camada em um processo analítico (Stamper, 1993a). A CO possibilita distinguir os limites de cada camada do SI, bem como as ligações e interdependências entre a organização e as suas normas. A Figura A.4 apresenta a CO.

A CO foi usada para caracterizar os estágios do ciclo de vida da desinformação. A caracterização engloba as atividades e os aspectos sociotécnicos do fenômeno encontrados no mapeamento sistemático da literatura e nos resultados de análises de eventos desinformativos aplicando a abordagem para análise da comunicação de desinformação.

Além da Cebola Organizacional, foram empregados a Análise Semântica e a modelagem Ontológica numa tentativa de modelar os elementos e relações que constituem casos de desinformação. A Análise Semântica foi embasada nos resultados do mapeamento e das análises de casos de desinformação, buscando construir uma descrição textual dos elementos que constituem um caso de desinformação. Ela parte da identificação de unidades semânticas (verbos, substantivos, preposições, etc.) que indicam os agentes, as propriedades e as *affordances* em um contexto de análise. Geralmente substantivos são agentes, verbos são *affordances*, e propriedades caracterizam os dois. O objetivo é identificar relações entre agentes, *affordances* e propriedades para modelar o contexto de interação. Por exemplo, a *affordances* “enganar” é dependente da existência de um meio de comunicação, e as propriedades de um agente são dependentes da existência desse agente.

Com os dados provenientes da análise semântica se constrói uma representação gráfica do contexto de interação analisado. Para isso se utiliza o Diagrama Ontológico que define símbolos para construção dessa representação. A Figura A.15 exemplifica o Diagrama Ontológico para um caso de desinformação (primeiramente chamado de Evento Desinformativo).

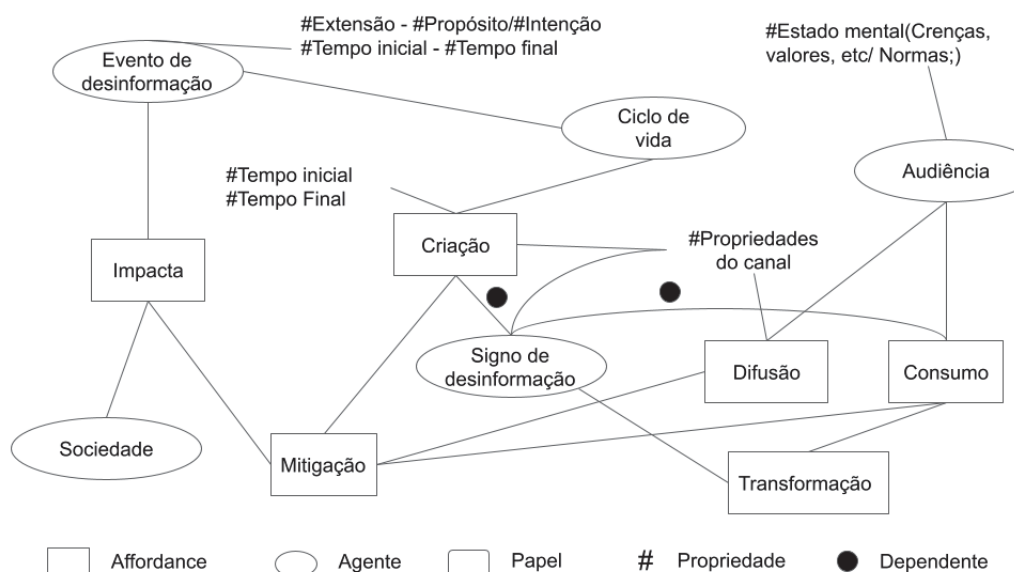


Figura A.7: Aspectos de um caso de desinformação (evento) na DO.

### A.6.0.1 O estágio de Criação

No estágio de Criação um objeto desinformativo é criado e divulgado. Este estágio engloba aspectos sociotécnicos do meio de comunicação e da organização que devem ser considerados na análise de objetos desinformativos. A Figura A.8 apresenta alguns exemplos destes aspectos.

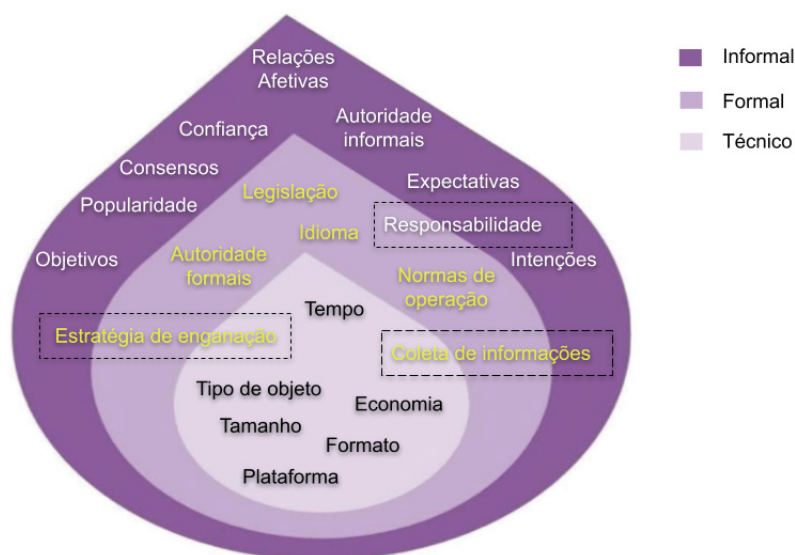


Figura A.8: Aspectos do estágio de criação na CO.

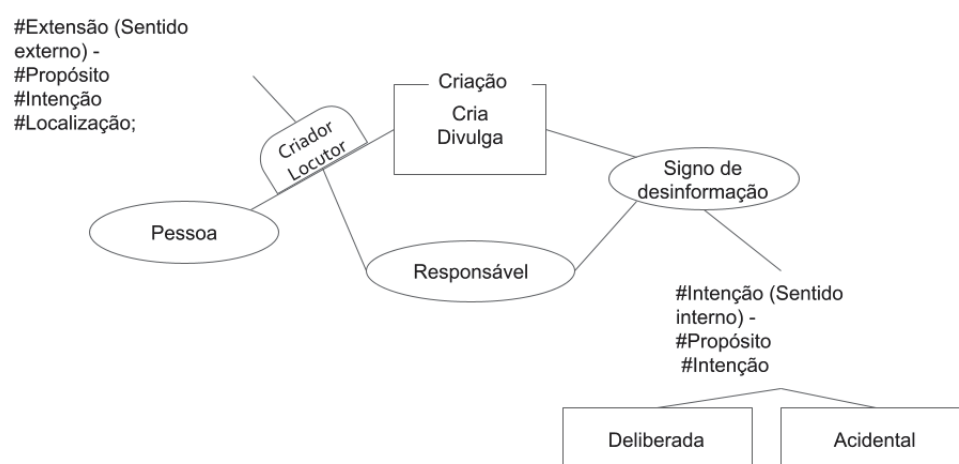


Figura A.9: Aspectos do estágio de criação na CO.

A tecnologia de comunicação molda a forma como as pessoas se expressam e em qual extensão e intensidade elas conseguem interagir (McLuhan, 1974). Esse fator indica que aspectos tecnológicos limitam a capacidade de expressão de uma ideia e restringem a percepção de aspectos da meta-comunicação, como a entonação da voz, movimentos corporais, expressões

faciais e aspectos do ambiente (distrações, sobrecarga de informações, etc. ). Nesse sentido, é importante considerar aspectos sociotécnicos que determinam os limites para se construir a mensagem e as propriedades do meio de comunicação que auxiliam no entendimento da intenção/proposito da comunicação. Exemplos desses aspectos vão desde uma propriedade física, como um artefato tecnológico que só permite representações textuais ou uma tela quebrada, e a capacidade de sensoramento do indivíduo, como dificuldade para enxergar ou digitar, até aspectos culturais que determinam significados para os termos e gírias, ou situacionais, como o estado emocional da parte interessada que cria a mensagem.

As características sociais de um grupo revelam seus consensos, seus hábitos de julgamento e ação. Estas propriedades caracterizam tanto a ocorrência de desinformação acidental quanto deliberada. No primeiro caso, desentendimentos, ambiguidades, normas sociais diferentes, são exemplos de aspectos que podem levar a um desentendimento ou interpretação indesejada de uma mensagem (Bessi et al., 2015). No segundo caso, estas características tornam as partes interessadas suscetíveis a desinformação deliberada, sendo exploradas ativamente (Bay, 2018). Assim, crenças, valores, experiências, expectativas, necessidades, e consensos configuram aspectos do contexto de uma organização que são explorados por uma parte interessada mal-intencionada através do uso da desinformação. Nesse sentido, além dos aspectos sociotécnicos relacionados a desinformação acidental, também há aspectos de influência deliberados usados no enquadramento que a mensagem é construída, como a relação entre o tema e o ambiente da audiência pretendida, a ideologia do grupo (Starbird et al., 2019), os valores defendidos (Piccolo et al., 2020), as referências a fontes confiáveis do grupo (Doty, 2015), etc.

Um objeto desinformativo pode ser fruto de uma atividade colaborativa ou um ato solitário e oportunístico (Hussain et al., 2018). As atividades colaborativas caracterizam operações estratégicas de informação (Starbird et al., 2019), onde uma organização surge do interesse comum de partes interessadas com intenções de manipulação e no ganho conjunto. Outra forma de operação estratégica envolve uma organização com funções e hierarquias de operação, como um órgão governamental (Starbird et al., 2019), onde nem todas as partes interessadas estão cientes da intencionalidade/proposito da atividade. Nesse contexto, atividades organizadas requerem algum nível de planejamento e coordenação (Hussain et al., 2018; Bazan, 2017; Loui e Hope, 2017; Traylor et al., 2017; Carlsson e Gustavsson, 2017; Li e Santos Jr, 2012). Além do planejamento, o estágio de criação do objeto desinformativo pode englobar a preparação (Oehmichen et al., 2019), elaboração (McCarthy, 2018) e divulgação do conteúdo falso (Boshmaf et al., 2013; Traylor et al., 2017; Raynal e Gaspard, 2010).

#### *A.6.0.2 O estágio de Difusão*

O estágio de Difusão aborda a propagação do objeto desinformativo. Este estágio considera os aspectos do meio comunicação e das mecânicas de difusão. A Figura A.10 apresenta alguns aspectos da difusão organizados na CO que podem ser considerados na análise de desinformação digital deliberada.

A comunicação de desinformação pode ocorrer de uma parte interessada para outra (telemarketing de máquina, golpes no WhatsApp, etc.), uma parte interessada para muitas partes interessadas (podcasts, Youtube, etc.), e muitas partes interessadas para muitas partes interessadas (Facebook, grupos no WhatsApp,). As partes interessadas podem ser humanos, bots ou ciborgues (Shahid et al., 2022). A difusão de um objeto desinformativo pode ocorrer em um único canal de comunicação ou em vários ao mesmo tempo.



Figura A.10: Aspectos da difusão na CO.

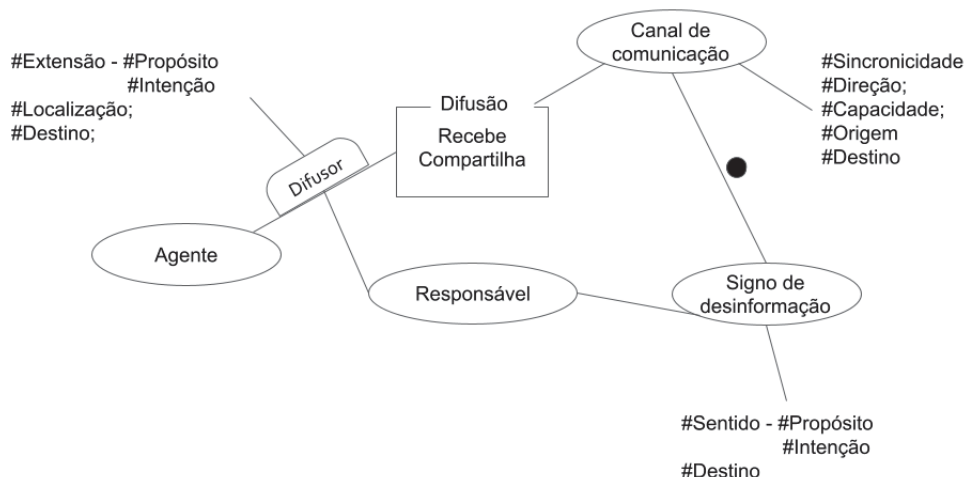


Figura A.11: Aspectos do estágio de Difusão na DO.

A comunicação pode ser síncrona ou assíncrona (Madhusudan, 2003; Zhou e Zhang, 2007). Na comunicação síncrona as partes interessadas interagem simultaneamente, como em uma conversação por aplicativos de chamada de vídeo e voz. Isso implica em um estilo de comunicação parecido com uma conversação, onde aspectos técnicos influenciam o formato da interação entre as partes interessadas. Por exemplo, o tempo de resposta em uma conversação é diferente do tempo de resposta em um e-mail, sendo uma comunicação dinâmica e veloz (Khodabakhsh et al., 2018). Isso impacta no controle sobre o fluxo de informações que a parte mal intencionada tem na interação, como a possibilidade de perceber, em tempo real, as reações da audiência e adaptar o estilo de comunicação (Li e Santos Jr, 2012). Por exemplo, usar uma linguagem formal ou informal, se apropriar de gírias/termos habituais do grupo, etc. Do mesmo modo há restrições na comunicação síncrona, como o alcance direto da desinformação, as restrições de compartilhamento do conteúdo, o tamanho da audiência, a rastreabilidade da fonte que divulga a desinformação, etc.



No formato assíncrono, a comunicação ocorre em momentos diferentes, como em blogs, e-mails e aplicativos de mensagens. Isso implica em um formato de comunicação com regras diferentes para o comportamento de interação entre interlocutor e a audiência. Nesse tipo de comunicação, a parte mal-intencionada tem controle restrito sobre o fluxo de informações na plataforma de comunicação, o que reduz a capacidade dela de coletar informações sobre a audiência enquanto se comunica, e adaptara o estilo para ser mais facilmente aceita. Além disso, o formato afeta os recursos de interação disponíveis em cada plataforma de comunicação, como em aplicativos de mensagens e e-mails há a facilidade de encaminhamento de mensagens dentro da plataforma, portais e microblogs possuem recursos de encaminhamento para outras plataformas, e fóruns e caixas de comentários em portais jornalísticos não possuem nenhum recurso desse tipo.

O canal de comunicação pode ser bidirecional e unidirecional (Neumann, 1996). Na comunicação bidirecional, todos os participantes têm o mesmo privilégio na interação, podendo enviar e receber mensagens, em um meio síncrono (videochamada) ou assíncrono (fórum). Nesse sentido, canais bidirecionais possibilitam a interação da audiência com a parte mal-intencionada, o que aumenta o controle da parte mal-intencionada sobre a comunicação. Este tipo de comunicação possibilita a resolução de conflitos (mal entendidos) da audiência, e a intervenção da parte mal-intencionada para direcionar conclusões (Li e Santos Jr, 2012).

Na comunicação unidirecional, existe um transmissor e um receptor bem definidos, como em blogs e portais informativos. Nesse tipo de comunicação não há interação direta da parte mal-intencionada com a audiência durante a comunicação (Neumann, 1996). Nesse sentido, este tipo de comunicação exige que a parte mal-intencionada planeje, prepare e construa a desinformação previamente para uma determinada audiência, e sobre um determinado tema. Ainda, a audiência está restrita a posição de receptor, o que impacta na capacidade de resolução de conflitos de entendimento, podendo levar ao consumo da desinformação (Chen e Sin, 2013).

As características do meio de comunicação se relacionam com os Princípios de Influência e seus impactos na comunicação. Por exemplo, na comunicação síncrona, as reações e características da audiência podem ser usadas pela parte interessada mal-intencionada para adaptar a desinformação, como alterar o estilo da comunicação, o fluxo de informações, e explorar vieses cognitivos com base no que ele observa.

Quadro A.21 apresenta um comparativo das propriedades dos canais de comunicação em relação ao estilo de comunicação e interação que possibilitam. Este quadro está incorporado na abordagem para análise de desinformação como parte do material de apoio. Ele foi elaborado com base nos dados do mapeamento sistemático da literatura.

#### *A.6.0.3 O estágio de Consumo*

O estágio de Consumo aborda à interação das partes interessadas com o objeto desinformativo. As interações podem ser vistas como qualquer ação feita em relação ao conteúdo falso, como ignorar, ler, reagir (gostar, odiar, etc.), comentar, compartilhar, alterar e corrigir (Bossetta, 2018). A Figura A.12 apresenta alguns aspectos do estágio de Consumo organizados na CO.

Tabela A.21: Tipos de interações, tempo e direção da comunicação

| Interação        | Síncrona                                                                                                                                                                                                                                                                                                                                                                            |                                                                                                                                                                                                                                                                                                                                                             | Assíncrona                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                                                      |
|------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  | Unidirecional                                                                                                                                                                                                                                                                                                                                                                       | Bidirecional                                                                                                                                                                                                                                                                                                                                                | Unidirecional                                                                                                                                                                                                                                                                                                                                 | Bidirecional                                                                                                                                                                                                                                                         |
| Um p/ um         | <p>Ex: Telemarketing de máquina;</p> <ul style="list-style-type: none"> <li>• Privilégios diferentes de interação;</li> <li>• Receptor restrito à audiência;</li> <li>• Difusão limitada (entre pares);</li> <li>• Alcance limitado;</li> <li>• Compartilhamento restrito;</li> <li>• Intencionalidade interlocutor;</li> </ul>                                                     | <p>Ex: Vídeo-chamada;</p> <ul style="list-style-type: none"> <li>• Privilégios iguais de interação;</li> <li>• Difusão restrita;</li> <li>• Resolução de conflito possível;</li> <li>• Tempos de resposta curto;</li> <li>• Compartilhamento restrito;</li> </ul>                                                                                           | <p>Ex: Recado de voz (ligação celular);</p> <ul style="list-style-type: none"> <li>• Privilégios diferentes de interação;</li> <li>• Receptor restrito à audiência;</li> <li>• Difusão limitada (entre pares);</li> <li>• Tempo de interação indefinido;</li> <li>• Velocidade de difusão;</li> <li>• Compartilhamento facilitado;</li> </ul> | <p>Ex: E-mail;</p> <ul style="list-style-type: none"> <li>• Privilégios iguais de interação;</li> <li>• Resolução de conflito possível;</li> <li>• Argumentação direta;</li> <li>• Tempo de interação indefinido;</li> <li>• compartilhamento facilitado;</li> </ul> |
| Um p/ muitos     | <p>Ex: Transmissão de TV ao vivo;</p> <ul style="list-style-type: none"> <li>• Privilégios diferentes de interação;</li> <li>• Interlocutor como curador da informação;</li> <li>• Maior controle da interação;</li> <li>• Receptor restrito à audiência</li> <li>• Difusão abrangente;</li> <li>• Resolução de conflito restrita;</li> <li>• Compartilhamento restrito;</li> </ul> | <p>Ex: Live;</p> <ul style="list-style-type: none"> <li>• Privilégios semelhantes de interação;</li> <li>• Interlocutor autoridade;</li> <li>• Maior tempo de interação;</li> <li>• Controle sobre a interação de outros;</li> <li>• Resolução de conflitos possível;</li> <li>• Tempos de resposta curto;</li> <li>• Compartilhamento restrito;</li> </ul> | <p>Ex: Pop-up;</p> <ul style="list-style-type: none"> <li>• Privilégios diferentes de interação;</li> <li>• Receptor restrito à audiência;</li> <li>• Difusão abrangente;</li> <li>• Maior alcance e velocidade;</li> <li>• Compartilhamento restrito;</li> </ul>                                                                             | <p>Ex: E-mail (mala-direta);</p> <ul style="list-style-type: none"> <li>• Privilégios iguais de interação;</li> <li>• Resolução de conflitos possível;</li> <li>• Tempos de resposta curto;</li> <li>• Compartilhamento facilitado;</li> </ul>                       |
| Muitos p/ muitos | <p>Ex: Plataforma de trabalho colaborativo;</p> <ul style="list-style-type: none"> <li>• Privilégios iguais de interação;</li> <li>• Resolução de conflitos possível;</li> <li>• Maior quantidade de conflitos;</li> <li>• Compartilhamento restrito;</li> </ul>                                                                                                                    | <p>Ex: Videoconferência;</p> <ul style="list-style-type: none"> <li>• Privilégios iguais de interação;</li> <li>• Tempos de resposta curto;</li> <li>• Resolução de conflitos possível;</li> <li>• Compartilhamento restrito;</li> </ul>                                                                                                                    | <p>Ex: Feed (Facebook);</p> <ul style="list-style-type: none"> <li>• Privilégios semelhantes de interação;</li> <li>• Moderadores;</li> <li>• Argumentação indireta;</li> <li>• Resolução de conflitos possível;</li> <li>• Tempo de interação indefinido;</li> <li>• Compartilhamento facilitado;</li> </ul>                                 | <p>Ex: Grupos virtuais (WhatsApp);</p> <ul style="list-style-type: none"> <li>• Privilégios iguais de interação;</li> <li>• Resolução de conflitos possível;</li> <li>• Tempo de interação indefinido;</li> <li>• Compartilhamento facilitado;</li> </ul>            |

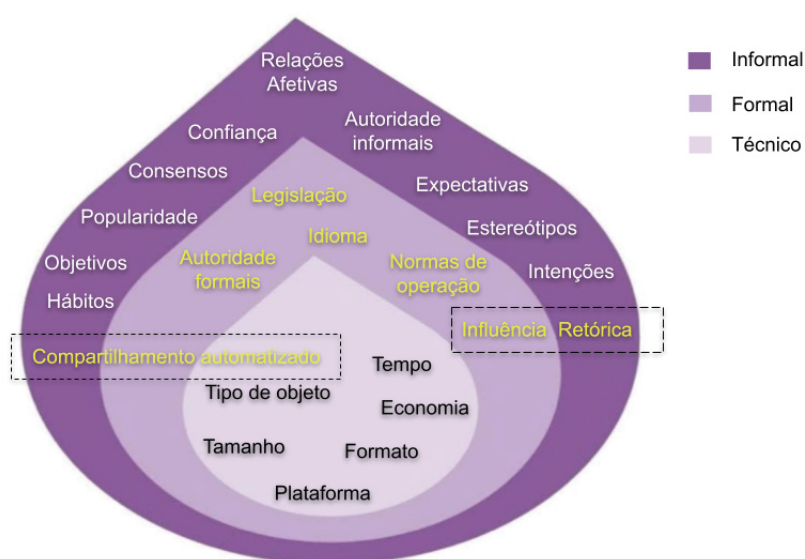


Figura A.12: Aspectos do consumo na Cebola Semiótica.

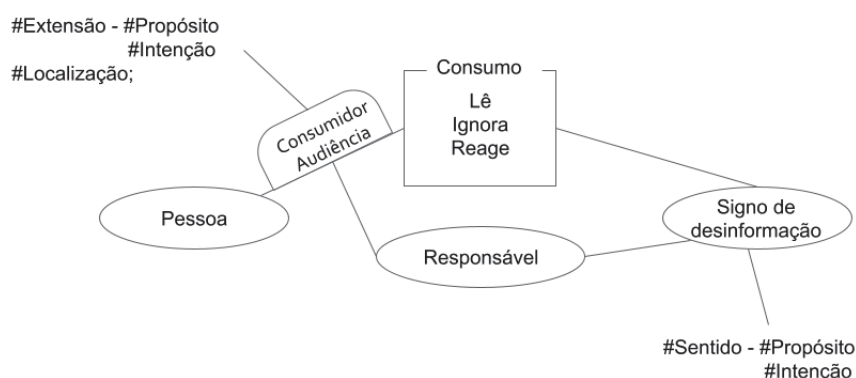


Figura A.13: Aspectos do estágio de criação na CO.

O consumo de desinformação depende de diversos aspectos sociotécnicos e do tipo de desinformação (acidental ou deliberada). No caso de uma ocorrência de desinformação acidental, um objeto digital foi interpretado de modo a não satisfazer a intenção de quem criou a mensagem. Nesse sentido, os mesmos fatores relacionados às limitações que a tecnologia de comunicação impõe na capacidade de expressão (estágio de Criação) podem ser observados para análise de como a audiência observa a comunicação e interpreta uma mensagem. É preciso entender qual foi o enquadramento usado pelas partes interessadas para observar o objeto desinformativo e quais aspectos sociotécnicos estão relacionados ao consumo, como os Princípios de Influência. Assim,

é preciso identificar as partes interessadas que interagem com o objeto e caracterizar os aspectos sociotécnicos do contexto da comunicação que são relevantes para entender o enquadramento.

A audiência faz parte de uma configuração social, com um campo de informação que impacta sua interação. A interpretação do conteúdo de um objeto desinformativo tem relação com a construção social e o contexto do indivíduo que interage com ele. Assim, aspectos de influência como a proximidade afetiva (Cheung e Lee, 2009; Rieh, 2002), autoridades para o grupo (Doty, 2015), crenças comuns (Cholvy, 2014), e ideologias (Jiang et al., 2019) são aspectos que influenciam a interpretação. No caso de desinformação deliberada, partes interessadas mal intencionadas exploram intencionalmente estes e outros aspectos de influência. Outros exemplos são ambiguidades e falácias lógicas que influenciam a percepção e interpretação de mensagens pelas partes interessadas (Cialdini, 1993).

As propriedades do meio de comunicação moldam o objeto desinformativo. O tamanho da mensagem (Kumar et al., 2016b; Zannettou et al., 2019), a quantidade de caracteres (Kumar et al., 2016b; Zhou e Zhang, 2004), se contém links (Aigner et al., 2017), se permite imagens (Taylor-Smith et al., 2018), a diagramação do conteúdo (Leite et al., 2019; Aigner et al., 2017), as cores (Taylor-Smith et al., 2018), são aspectos da plataforma digital de comunicação que implicam em características para o formato da desinformação. Estes são exemplos de aspectos que podem facilitar o consumo do objeto desinformativo, como uma mensagem pequena, ou um título destacado (Ahmad et al., 2019).

Na desinformação deliberada, o formato do objeto desinformativo pode explorar a similaridade com outros objetos digitais de uso comum da audiência. Por exemplo, o formato de uma notícia falsa (*fake news*), quando o objeto desinformativo imita uma notícia, diminui percepção de risco dos usuários e facilita o consumo (Shen et al., 2019; Kasra et al., 2018; Aigner et al., 2017). O objeto desinformativo pode ser construído para conduzir a atenção do leitor e guiar o foco, com elementos de ênfase e destaque (Shen et al., 2019), como os objetos do tipo Iscas-clique que chamam a atenção com legendas sensacionalistas em imagens (Chen et al., 2015c). Uma parte interessada em manipular a audiência, pode explorar aspectos da forma do objeto desinformativo para obter uma maior taxa de sucesso no consumo (Shrestha e Spezzano, 2019; Traylor et al., 2017).

Ainda, aspectos físicos, como o dispositivo (TV, celular, etc.) e a plataforma digital (WhatsApp, Blog, etc.) que o usuário recebe a desinformação tem impacto na confiabilidade percebida e percepção de risco dos usuários (Hölig et al., 2020), podendo o levar a consumir um conteúdo sem checar as informações (Kovic e Laissue, 2016). Além disso, os recursos de interação do sistema operacional do dispositivo, como cópias de tela e o compartilhamento de links facilitado, sobrepujam as restrições dos aplicativos e plataformas de comunicação quando não oferecem tais possibilidades, podendo impactar tanto na interação do consumo quanto na difusão da desinformação.

Os fatores humanos, como a memória, têm limitações que os tornam vulnerabilidade para exploração, como a suscetibilidade à introdução de viés na reconstrução de eventos e experiências passadas de uma parte interessada, através de fotos e descrições adulteradas (Larson et al., 2019), ou explorar aspectos da cognição como vieses cognitivos conhecidos (Li e Santos Jr, 2012). Os aspectos sociais, como a influência do difusor e hábitos de consumo de informações das comunidades (Zannettou et al., 2019; Torres et al., 2018b), também influenciam a percepção de risco e o consumo de desinformação (Flintham et al., 2018). Nesse sentido, esses aspectos representam vulnerabilidades que podem ser exploradas por partes interessadas que criam a desinformação digital.

#### A.6.0.4 O estágio de Reiteração

O estágio de ressurgimento ou reiteração contempla as atividades e transformações que ocorrem no objeto desinformativo quanto a sua resiliência a mitigação e seu reaparecimento. Ao ser divulgado, o meio e a audiência se apropriam do objeto desinformativo e o transformam, agregando seus próprios propósitos (Bandeli e Agarwal, 2018; Starbird et al., 2019). Essas transformações dificultam o rastreo e mitigação, e ampliam os impactos da desinformação, como o alcance, velocidade de difusão e na neutralização do objeto desinformativo.

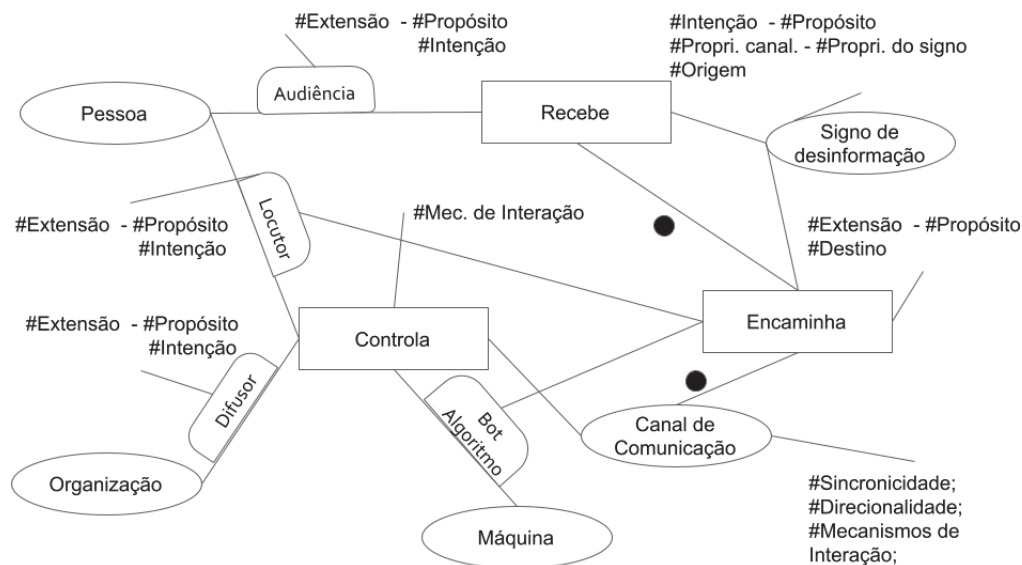


Figura A.14: Aspectos do estágio de Reiteração na DO.

As transformações podem ser tanto orgânicas, quando a audiência transforma a mensagem, quanto sintéticas, quando bots recriam ou redirecionam o conteúdo (Starbird et al., 2019; Yu e Han, 2018). A audiência pode adaptar a mensagem para ser transmitida em uma mídia diferente, migrando para uma plataforma digital com propriedades diferentes para a comunicação. Transformações também ocorrem artificialmente com o objetivo de aumentar a eficácia da mensagem. Nesse caso, agentes intencionalmente ajustam a mensagem para evitar detecção (Bandeli e Agarwal, 2018), ou promover mensagens já desmascaradas em formato novo (Finn et al., 2015; Tang et al., 2017).

Após a divulgação do conteúdo original, o objeto desinformativo pode migrar, tanto entre diferentes plataformas digitais de comunicação (Janicka et al., 2019; Del Vicario et al., 2017), quanto entre grupos na mesma plataforma (Melo et al., 2019), passando por diversas tecnologias de comunicação. Quando um objeto migra para uma nova organização, o contexto do seu ciclo de vida se altera, considerando novos aspectos para a plataforma tecnológica e novos aspectos para o contexto. Por exemplo, um conteúdo originalmente criado para ser divulgado em um formato com propriedades bem definidas (como o pdf), pode se transformar em links, texto e imagens quando migra de uma plataforma de comunicação para outra. Assim, o objeto desinformativo se torna heterogêneo quando se difunde, difícil de rastrear e conter (Zhou e Zhang, 2007), dificultando o trabalho de mitigação e possibilitando o reaproveitamento/reaparecimento em outros formatos, contextos e períodos de tempo.



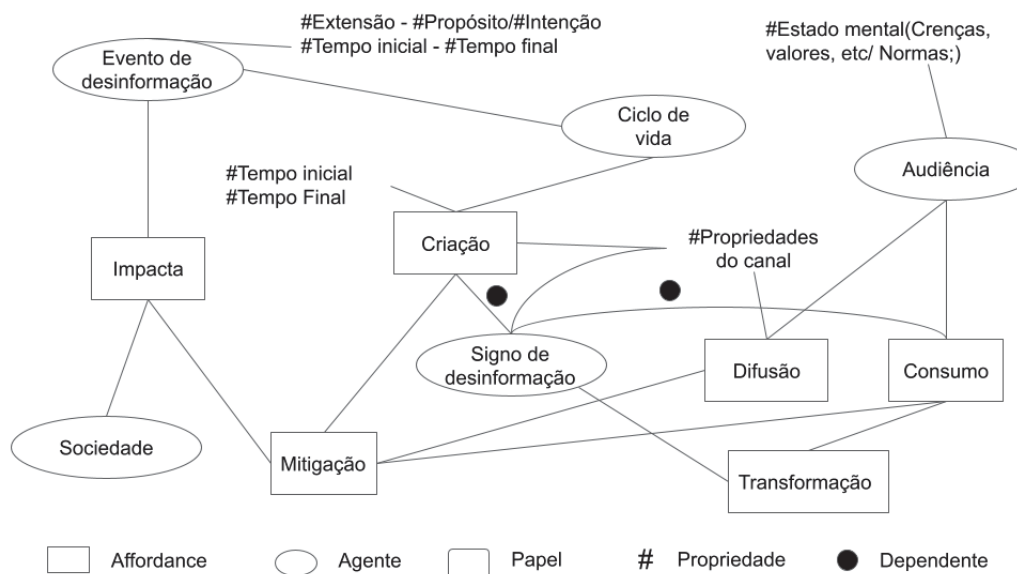


Figura A.15: Aspectos de um caso de desinformação (evento) na DO.

## A.7 APÊNDICE 7 - RELATÓRIOS DAS AVALIAÇÕES

Esta seção apresenta os relatórios detalhando as análises de eventos desinformativos feitos utilizando a abordagem.

### A.7.1 Sobre os estudo em pequena escala

Essa seção detalha os estudos em pequena escala, especialmente os dois estudos que utilizaram grupo focal. Os outros estudos são detalhados em video-tutoriais disponíveis na plataforma Youtube (Uma Fake News <sup>29</sup> e uma Teoria da Conspiração <sup>30</sup>).

#### A.7.1.1 Análise de um esquema de phishing (Fraude e-mail)

Em 2011, a RSA Security era a divisão de segurança da corporação EMC, criando produtos de rede, criptografia e segurança computacional (Westervelt, 2011). Em março de 2011 a empresa sofreu um ataque de *spear phishing* (Zetter, 2011). Este tipo de ataque usa a desinformação para persuadir indivíduos selecionados à executar uma ação aparentemente inofensiva, mas que causará prejuízos (Caputo et al., 2013). Nesse caso, um grupo de hackers selecionou quatro empregados sem importância na hierarquia da empresa como alvos. A mensagem falsa foi construída usando o e-mail como mídia digital, contendo o assunto "Plano de Recrutamento 2011" e um arquivo do Excel com um malware oculto. O e-mail foi detectado como spam pela infraestrutura de segurança da empresa e direcionado para o lixo eletrônico. Um dos quatro empregados que receberam o e-mail abriu o arquivo do Excel, infectando o seu terminal. Com o acesso, os hackers usaram o terminal para escalar os privilégios na infraestrutura e roubar dados dos clientes (Mikko, 2021). A Figura A.16 mostra mensagem.

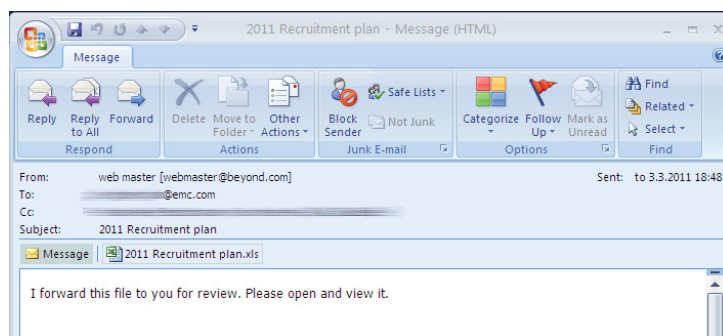


Figura A.16: Representação do objeto desinformativo(phishing). Fonte: F-Secure (Mikko, 2021).

#### A.7.1.2 Clarificação do domínio do evento - exploração do contexto

A versão utilizada da Dismiss contemplava uma fase para caracterização do campo de informações de uma organização <sup>31</sup>. Isso era feito utilizando a Cebola Organizacional (CO) para mapear as características Informais, Formais e Técnicas do contexto do caso de desinformação, os dados da CO eram passados de entrada para a técnicas de Análise Semântica, onde era encontradas as relações entre agentes, *affordances* e propriedades, que então serviam de entrada para a Modelagem Ontológica. O campo de informações foi caracterizado com base em documentos oficiais que detalhavam do ataque (RSA, 2021), notícias em portais informativos

<sup>29</sup><https://www.youtube.com/watch?v=hWyjbR9-KIE>

<sup>30</sup><https://www.youtube.com/watch?v=NzWYrbrLjfc>

<sup>31</sup>Mais informações sobre o campo de informações, *affordances* e agentes na Seção A.1.1.3

confiáveis (Westervelt, 2011; Zetter, 2011) e relatórios de empresas de segurança (Mikko, 2021). Identificamos 13 partes interessadas para o contexto do evento desinformativo. As partes interessadas foram divididas em dois grupos principais, os interessados em manipular (atacantes), e a organização. Figura A.17 mostra o mapa de partes interessadas organizadas no Diagrama de Identificação de Partes Interessada (DPI).

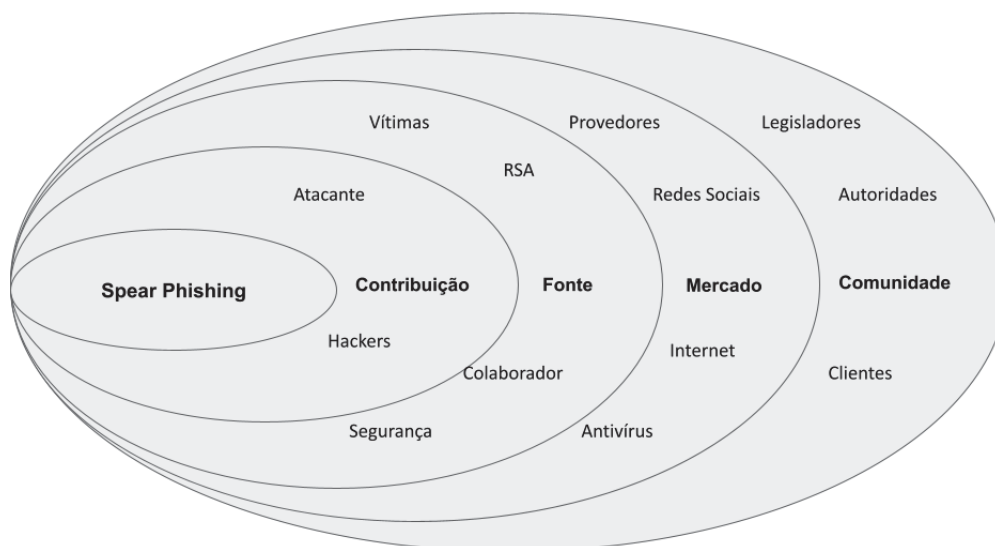


Figura A.17: Partes interessadas (RSA Phishing)

O objetivo das partes interessadas "Hackers"(grupo contribuição) era o ganho monetário, através da venda de dados privados da parte interessada "Clientes"(grupo comunidade), que contratam a parte interessada "RSA"(grupo fonte). A empresa RSA se dedica a atividades de segurança digital, possuindo a obrigação de proteger os dados dos seus clientes. Para isso, ela emprega mecanismos de prevenção contra o acesso e uso indevido desses dados, como *firewalls*, antivírus, normativas de segurança, e criptografia (Rivner, 2011). Nesse sentido, o acesso aos dados é uma atividade com um custo elevado para aqueles que não tem permissão. Por esse motivo, a parte interessada "Hackers"investiu tempo e esforço para enganar os mecanismos de segurança da RSA, encontrando na desinformação uma forma menos custosa de conseguir uma porta de entrada na rede da empresa. Os usuários de sistemas, representam o elo mais frágil na cadeia de segurança cibernética (Caputo et al., 2013).

Os hackers planejaram o ataque, coletando informações sobre o contexto da empresa e seus colaboradores, com o objetivo de encontrar aspectos passíveis de exploração (Rivner, 2011). A parte interessada "Redes sociais"(grupo mercado), serviu como uma fonte para coleta de dados das partes interessadas alvo, revelando informações sobre o perfil, como seus cargos, seus endereços de correspondência, e escolaridade (Rivner, 2011). A Figura A.18 apresenta os aspectos do campo de informação identificados na fase de Clarificação do Domínio do evento no Artefato da Cebola Organizacional.

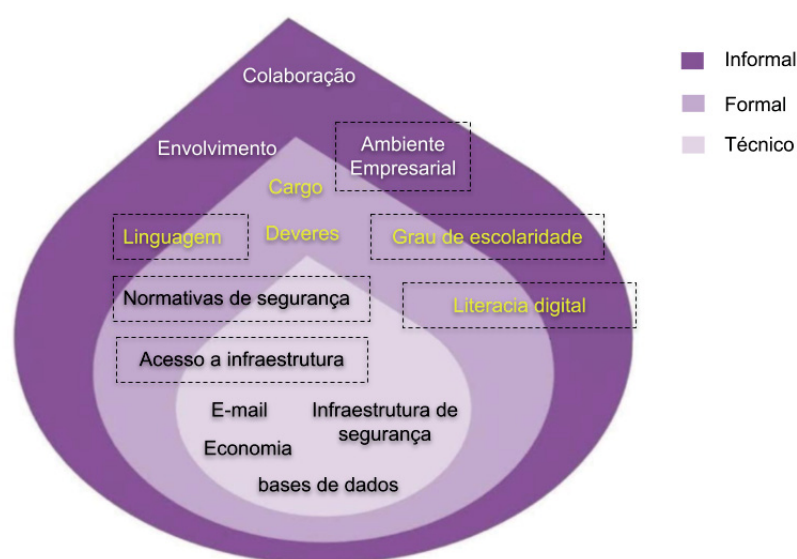


Figura A.18: Aspectos do campo de informações relevantes para o evento desinformativo.

A Figura A.19 apresenta os aspectos do campo de informação identificados na fase de Clarificação do Domínio do evento no Formulários de Caracterização do Campo de Informações. Esse formulário foi pensado para direcionar a análise do caso considerando as perspectivas de cada parte interessada.

|          | Hackers                                |                                       |                                                       | RSA                                     |                                 |                                               | Empregados                                      |                                 |                                                                 |
|----------|----------------------------------------|---------------------------------------|-------------------------------------------------------|-----------------------------------------|---------------------------------|-----------------------------------------------|-------------------------------------------------|---------------------------------|-----------------------------------------------------------------|
| Informal | Grupo; estelionato; não seguem normas; | Recebe correspondência diariamente    | Uso deliberado de linguagem informal                  | Atividade Economia                      | comunicação interna organização | Colaboração entre empregados                  | Redes sociais                                   | Comunicação informal            | linguagem, correspondência interna informal, Hábitos de consumo |
| Formal   | Preparação Plano; Não seguem leis;     | Conta de e-mail remete a cargo/função | Uso deliberado de nomes significativos ambiente empre | Amb. Empresarial Protocolos Hierarquias | normativas de segurança         | Amb. Empresarial Normativas de segurança;     | Cargo, privilégio de acesso; Literacia digital; | Cargo Atividades                | Cargo; nomes significativos atividades; privilégio de acesso;   |
| Técnico  | E-mail; vírus;                         | Origem externa                        |                                                       | Base de dados                           | Infraestrutura Segurança        | Mecanismos de segurança, Antivírus; Anti-spam | Acesso a rede local RSA                         | conta serviço e-mail rede local | Conta e-mail;                                                   |

Figura A.19: Aspectos do campo de informações relevantes para o evento desinformativo.

O dados do Formulários de Caracterização do Campo de Informações passavam por uma fase de Análise Semântica, onde agentes, *affordances* e propriedades foram identificadas. A Figura A.20 apresenta as Unidades Semânticas identificadas e categorizadas.

| Unidades semânticas    | Agente-affordance-papel | Descrição                                                      |
|------------------------|-------------------------|----------------------------------------------------------------|
| Sociedade              | Agente                  |                                                                |
| Organização RSA        | Agente                  | Grupo de agentes que compartilham normas e consensos           |
| Pessoa                 | Agente                  | Parte da sociedade                                             |
| Empregado              | Papel                   | Parte da organização                                           |
| Cliente                | Papel                   | Parte da sociedade                                             |
| Hacker                 | Papel                   | Parte da sociedade                                             |
| Comunicação            | Agente                  | Atividade de uma ou mais pessoas                               |
| Signo de desinformação | Agente                  | Parte da comunicação; Todo - tipos de signos de desinformação; |
| E-mail                 | Papel                   | Tipo de serviço de comunicação;                                |
| Envia                  | Affordance              | Atividade de uma ou mais pessoas                               |
| Recebe                 | Affordance              | Atividade de uma ou mais pessoas                               |
| Abre                   | Affordance              | Atividade dependente de arquivo                                |
| Arquivo                | Agente                  |                                                                |
| Excel                  | Papel                   | Tipo de arquivo                                                |
| Vírus                  | Papel                   | Tipo de arquivo                                                |

Figura A.20: Exemplo de Unidade Semânticas - Caso RSA.

As Unidades Semânticas informaram a construção do diagrama Ontológico do Caso RSA. A Figura A.21 apresenta as Unidades Semânticas identificadas e categorizadas.

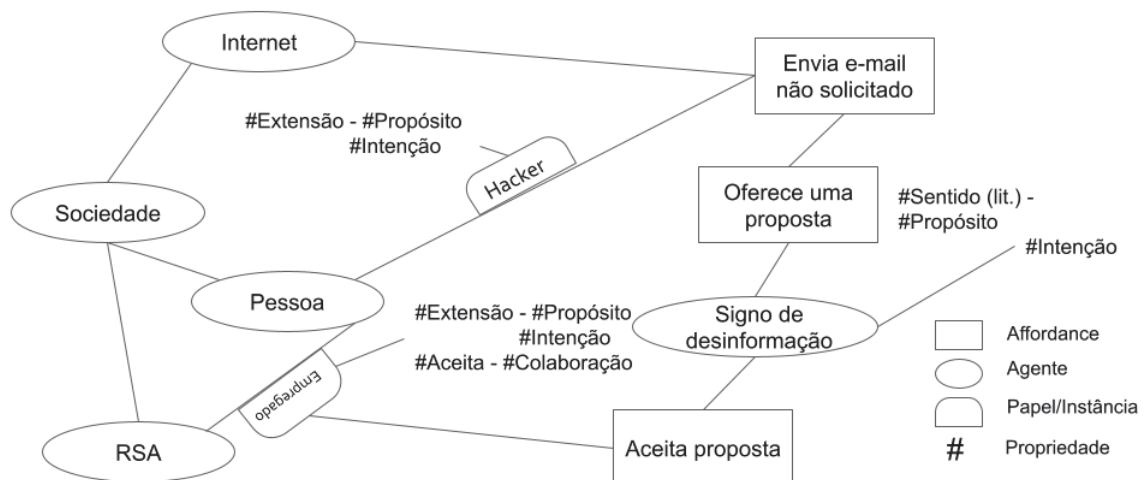


Figura A.21: Exemplo de Diagrama Ontológico de atividade Spear Phishing. - Caso RSA.

Nessa primeira versão da Dismiss, a Etapa 1 foi proposta para identificar as partes interessadas e construir uma visão do contexto do caso de desinformação, para a partir desses elementos básicos realizar a análise sociotécnica do caso. No entanto, devido à complexidade de conceitos e procedimentos relacionados a CO, a Análise Semântica e a Modelagem Ontológica (juntamente com os outros artefatos da SO usados em outras Etapas da Dismiss), o autor decidiu, com base nos resultados das avaliações dos Grupos Focais sobre a Dismiss, retirar estes artefatos e técnicas para simplificar a aplicação da Dismiss.

#### A.7.1.3 Análise Semiótica do objeto desinformativo

O objeto desinformativo não era sofisticado, com uma estrutura simples e um texto impessoal, que apelava para a curiosidade do colaborador, levando-o a abrir um arquivo infectado.



Os mecanismos técnicos de segurança detectaram o e-mail como possível spam e encaminharam a mensagem para caixa de spam. Um dos quatro colaboradores que recebeu o e-mail retirou a mensagem da caixa de spam e a abriu. A Figura A.22 apresenta alguns dos aspectos sociotécnicos identificados no objeto desinformativo.

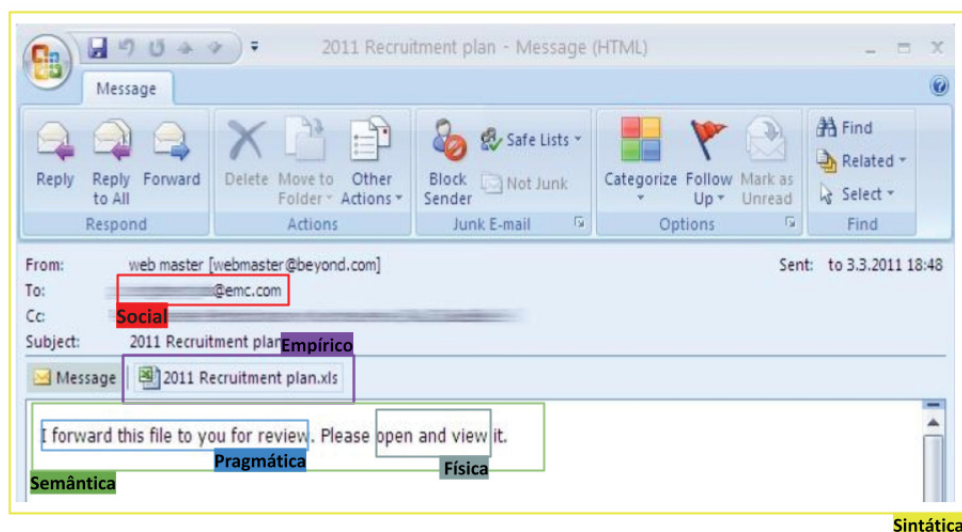


Figura A.22: Propriedades semióticas no objeto desinformativo.

Foram identificados 46 aspectos sociotécnicos, 29 do sistema de informação humana e 17 da plataforma tecnológica. Os dados da análise foram registrados em tabelas do Google Sheets<sup>32</sup>. O Quadro A.22 apresenta um resumo dos aspectos identificados.

<sup>32</sup>[https://docs.google.com/spreadsheets/analise\\_spear](https://docs.google.com/spreadsheets/analise_spear)

Tabela A.22: Resumo dos aspectos sociotécnicos identificados na análise

| Dimensão Semiótica | Aspecto                                                                                                  | Exemplos                                                                                                                                                                                                                                                                                                           |
|--------------------|----------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Social             | Atividades profissionais;<br>Impacto social;<br>Normas de segurança;                                     | A desinformação explora normas do campo de informação da empresa;<br>Causa impactos negativos na imagem da empresa;<br>Prejudica o relacionamento com os 'clientes';                                                                                                                                               |
| Pragmática         | Exploração de expectativas;<br>Exploração de intenções;                                                  | A desinformação tem o objetivo de influenciar os alvos a abrir o arquivo com o <i>malware</i> anexado no e-mail;<br>Empregados tem funções e responsabilidades dentro da organização;<br>Exploração da curiosidade do alvo;                                                                                        |
| Semântica          | Percepções da audiência;<br>Vieses cognitivos;<br>Inexperiência/Ignorância;<br>Linguagem representativa; | O pedido de ajuda para revisar o documento configura o viés cognitivo Quid-pro-quo - a tendência do ser humano em querer ajudar;<br>O aspecto persuasivo da temática, com título e nomes representativos de tarefas da organização;<br>Usuário é convencido a abrir e-mail marcado como spam;                      |
| Sintática          | Linguagem;<br>Diagramação;<br>Perfis de usuários;                                                        | Informações dos perfis dos colaboradores em redes sociais, para identificar vulnerabilidades passíveis de exploração, como seu cargo, idade, escolaridade;<br>E-mail na caixa de spam;<br>E-mail explora a similaridade com tarefas diárias;                                                                       |
| Empírica           | Tempo de leitura;<br>Interação da Comunicação;<br>Codificação;                                           | Quantidade de mensagens com o mesmo remetente (4 e-mails);<br>Arquivo contém um <i>malware</i> que explora uma falha <i>Zero-day</i> ;<br>Arquivo codificado do Excel com um tamanho adequado para transmissão e armazenagem;<br>Desinformação requer a abertura de um outro arquivo, fora da aplicação de e-mail; |
| Física             | Localização;<br>Infraestrutura de rede;<br>Normas de segurança;                                          | Atacantes precisam que o arquivo seja aberto dentro do domínio da empresa;<br>A Internet favorece o anonimato dos criadores;                                                                                                                                                                                       |

#### A.7.1.4 Leitura e análise dos resultados

A criação do objeto desinformativo foi uma atividade planejada que envolveu um grupo de pessoas e atividades de *preparação*. A desinformação foi criada para explorar características do contexto empresarial, como as atividades das funções desempenhadas pelos colaboradores da empresa RSA. O rótulo "Pragmática" apresentado na Figura A.22, representa um aspecto da **intenção falsa** expressa pelos criadores, pedindo para o destinatário revisar um arquivo. Os alvos foram selecionados evitando perfis de alto nível da organização. O arquivo contém um *malware* criado para ser usado no evento desinformativo, e que explora uma falha *Zero-day* - não conhecida ou documentada.

Em relação a difusão, os atacantes prepararam o ataque criando uma conta em um serviço de e-mails legítimo (Rivner, 2011). Estes serviços são populares e de fácil acesso,

comumente usados para manter a anonimato ao permitirem registros de qualquer lugar do mundo, sem requisitar muitas informações do contratante. Outro aspecto identificado está relacionado ao princípio da **Reciprocidade**, onde os usuários recebem uma mensagem não solicitada, que apresenta informações significativas sobre a organização e requisita uma ação. Ademais, a infraestrutura de acesso à Internet favorece o anonimato dos criadores, onde qualquer pessoa em qualquer lugar do planeta pode enviar um e-mail. O e-mail de destino de domínio da empresa foi uma decisão deliberada para forçar o acesso a mensagem por um terminal conectado à rede da empresa.

Ainda, alguns elementos sobre a **Comprometimento**, como termos significativos<sup>33</sup> para o título e o arquivo, remetem a coerência entre o pedido, o arquivo, e a função da parte interessada. Ainda, o apelo para o **Comprometimento** das partes interessadas, explora a responsabilidade em executar suas funções. A mensagem atrai o usuário a clicar no arquivo do Excel, que tem um tamanho adequado para transmissão e armazenagem, e não levanta suspeitas.

Em relação a mitigação, a legislação e fiscalização são insuficientes para inibir ou rastrear a origem da desinformação de modo efetivo. Os atacantes podem estar em qualquer lugar do planeta, e as leis e órgãos responsáveis tem dificuldade de lidar com ações criminosas realizadas em outros países. As vítimas tinham treinamento contra engenharia social que não foi suficiente, elementos como **Autoridade** do mecanismo de spam e seus avisos, não foram efetivos para alertar sobre o risco. A origem desconhecida do e-mail, algo incomum para uma mensagem com conteúdo interno da empresa, não foi suficiente para fazer com que o usuário descartasse o e-mail. Além disso, a quantidade de destinatários com a mesma mensagem e remetente (4 e-mails), não levantou suspeitas ou disparou a detecção automática (Rivner, 2011). Ainda, o *malware* explorava uma falha não documentada, e era impossível dos sistemas de detecção identificarem o código malicioso presente no arquivo do Excel. O arquivo estava vazio e fechava automaticamente ao ser aberto, sem indicações de erros ou qualquer mudança aparente, impactando na percepção da vítima sobre anormalidades<sup>34</sup>. Como consequências, o desfecho do evento acarretou na perda de confiança da parte interessada "Clientes" para com a parte interessada "RSA", além de prejudicar a imagem pública da empresa.

#### A.7.1.5 Considerações sobre a análise

A análise revelou aspectos sociotécnicos que impactam na criação, difusão, consumo e mitigação, que embasaram discussões sobre algumas mecânicas do ciclo de vida da desinformação. Por exemplo, aspectos sociotécnicos inter-relacionados que impactaram na criação, como a popularidade e situação econômica da empresa (motivação dos atacantes), a disponibilidade de informações relevantes sobre a organização e as partes interessadas em canais públicos e de livre acesso, o acesso privilegiado das partes interessadas à infraestrutura da empresa, e o anonimato proporcionado pela natureza livre da Internet, são características que facilitaram a elaboração do ataque que usem a desinformação como arma. A difusão foi facilitada pelas vulnerabilidades dos mecanismos técnicos e tecnológicos de segurança da empresa, como o *malware* que explora uma falha desconhecida, e os avisos de spam pouco significativos. O consumo foi impactado por aspectos organizacionais, como normativas e treinamentos de segurança insuficientes. A desinformação foi mitigada tardiamente, após o sucesso do ataque, impactado no relacionamento da empresa com seus clientes e na sua imagem pública. Dessa forma, a diversidade de aspectos sociotécnicos que influenciam de algum modo o ciclo de vida

<sup>33</sup>São termos adequados para o contexto empresarial, para atividades que as partes interessadas executam, e para temporalidade

<sup>34</sup>Vídeo da interação com o arquivo. Fonte: F-secure

da desinformação deliberada, representam a complexidade multidisciplinar do fenômeno da desinformação no contexto analisado.

Com base nos resultados, algumas orientações podem contribuir com as estratégias de mitigação. Os mecanismos tecnológicos de mitigação de ataques presentes na infraestrutura (antivírus, AntiSpam), informam aos usuários o estado do sistema e o risco apresentado por uma mensagem para a segurança das operações da empresa. Como no caso analisado, mesmo com normativas de segurança e treinamento contra engenharia social, o usuário ignorou o aviso e abriu o arquivo na mensagem suspeita, o que indica um desvio do comportamento esperado por esta parte interessada nessa situação. Os motivos dessa atitude não são claros pelos resultados da análise, mas há inferências de como aspectos relacionados aos Princípios da Influência podem impactar nas decisões das partes interessadas. Por exemplo, aspectos relacionados ao princípio de **Afinidade**, como o uso de um **meio de comunicação e arquivos comuns** as atividades da empresa, o **assunto da mensagem é representativo** das funções das partes interessadas, a **linguagem significativa e familiar** do contexto, são aspectos que influenciam na redução da percepção de risco (Cialdini, 1993).

Podemos considerar que o aviso de spam não foi significativo o suficiente para a consciência situacional <sup>35</sup> do usuário. Um aspecto importante sobre os avisos de sistemas computacionais, é que eles precisam ser significativos para a parte interessada que interage com o sistema, em especial quando lidando com desinformação (Fernandez e Alani, 2018). Nesse sentido, o serviço AntiSpam da RSA poderia utilizar uma escala de classificação de risco com base em diferentes aspectos identificados na análise. Por exemplo, os aspectos como "e-mails de fontes desconhecidas, tipos comuns de arquivos infectados, uso de linguagem informal" são três aspectos inter-relacionados que podem servir para classificar níveis de ameaças e avisar os usuários de acordo com esses níveis de modo mais representativo. O uso de escalas para os níveis de risco (Gao et al., 2018) transmite mais informações do que somente colocar um e-mail na caixa de spam.

Para evitar que vulnerabilidades humanas sejam exploradas pela desinformação, uma segunda camada de prevenção - além das ferramentas tecnológicas - reforça a capacidade de mitigar a desinformação. O treinamento para aprimorar a capacidade das partes interessadas em lidar com a desinformação pode ser informado por aspectos sociotécnicos identificados na análise. Conhecer os Fundamentos e Princípios da Influência, e como são usados, sobretudo suas adaptações e transformações no contexto da desinformação deliberada, pode aprimorar a capacidade de detecção de farsas (Cialdini, 1993). Isso pode ser feito com a conscientização através de exemplos de cenários de eventos desinformativos (Nicol, 2005). Ainda, a conscientização dos usuários em relação a suas responsabilidades são aspectos do SI Formal que podem auxiliar na mitigação - como estatutos e normativas sobre conduta e consequências, palestras e treinamentos para atualizar os conhecimentos sobre as ferramentas tecnológicas usadas, e manuais e guias para interação com os SI das empresas.

Além disso, aspectos do SI Técnico devem ser considerados no treinamento das partes interessadas, como a educação relevante sobre aspectos dos sistemas que os usuários operam que são usados para explorá-los. Por exemplo, os mesmos aspectos que podem ser usados para aprimorar o sistema de avisos de spam, como "e-mails de fontes desconhecidas, tipos comuns de arquivos infectados", podem ser ensinados para as partes interessadas. Estes são aspectos do SI Técnico que podem auxiliar na percepção de risco das partes interessadas, e levá-los a detectar a

---

<sup>35</sup>Este termo é usado no campo de Gerenciamento de Risco em Sistemas de Apoio a Decisão, representando o nível de consciência que um usuário ou equipe tem sobre uma situação. Nesse campo, falhas de consciência situacional podem acarretar decisões prejudiciais que acarretam riscos à vida, ao patrimônio ou ao meio ambiente (Kokar e Endsley, 2012).

desinformação. Ainda, atualizações dos treinamentos precisam ser constantes para capacitar os usuários a mitigar ataques e atacantes que estão sempre explorando novas formas de manipulação e influência.

### A.7.2 Análise de uma farsa no WhatsApp

Em março de 2020 circulou pelo WhatsApp no Brasil uma mensagem que relacionava o combate da COVID-19 com o nível de Ph do corpo humano (Uol, 2020b; G1, 2021b). A mensagem apresentava um tratamento onde o indivíduo deveria comer certos alimentos para atingir um nível de Ph que supostamente aniquila a COVID-19. A mensagem trazia em sua estrutura referência a autoridades autenticadoras, informações sobre alimentos com Ph errado, além da orientação de que é possível alterar o Ph do corpo (Johnson, 2020). A Figura A.23 apresenta a primeira aparição dessa desinformação no Brasil em março de 2020. Ainda, essa desinformação foi corrigida no final de março por agências brasileiras de checagem de (Lupa e G1) (Uol, 2020b; G1, 2021b), apesar das correções, o objeto desinformativo continuou circulando.

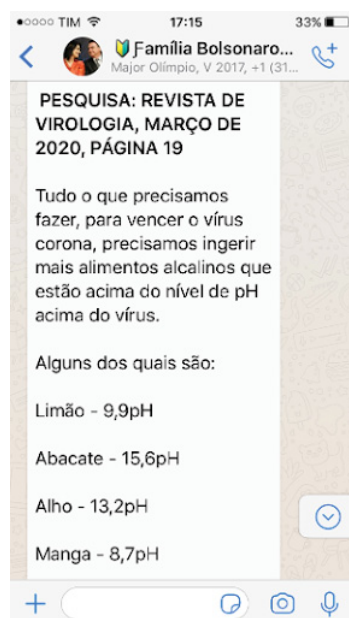


Figura A.23: Mensagem desinformativa. Fonte: Uol (Uol, 2020b).

#### A.7.2.1 Clarificação do domínio do evento

Devido a natureza privada de dados como, rede de origem, data de criação, e o usuário que criou, é complicado identificar onde e quando este objeto desinformativo foi criado. Nesse sentido, o evento foi considerado a partir do momento que o objeto surgiu no Brasil. Assim, o contexto desta análise engloba o objeto desinformativo na rede social WhatsApp, detectado em março/2020, e seus derivados que apareceram posteriormente.

O campo de informações foi caracterizado com base em relatórios de checadores de fatos (Uol, 2020b; G1, 2021b; Johnson, 2020) notícias em portais informativos confiáveis (Santos-Pinto et al., 2021; dos Anjos, 2021; CE, 2020; Times, 2020; Joel, 2021; Povo, 2020; Almirante e Marés, 2020), relatórios sobre hábitos de consumo de informações por populações realizados no ano do evento desinformativo (Hölig et al., 2020), e artigos científicos relacionados ao contexto do evento publicados naquele ano (Vicario et al., 2019). A Figura A.24 apresenta os aspectos do campo de informação identificados na fase de preparação.



Figura A.24: Aspectos do campo de informações relevantes para o evento desinformativo.

Foram identificadas 28 partes interessadas, divididas em três grupos principais: aquelas com ideologia alinhada a desinformação (simpatizantes), aqueles que combatem a desinformação (resistência) e os controladores dos meios de difusão. A Figura A.25 apresenta as partes interessadas mais relevantes mapeadas no DPI.

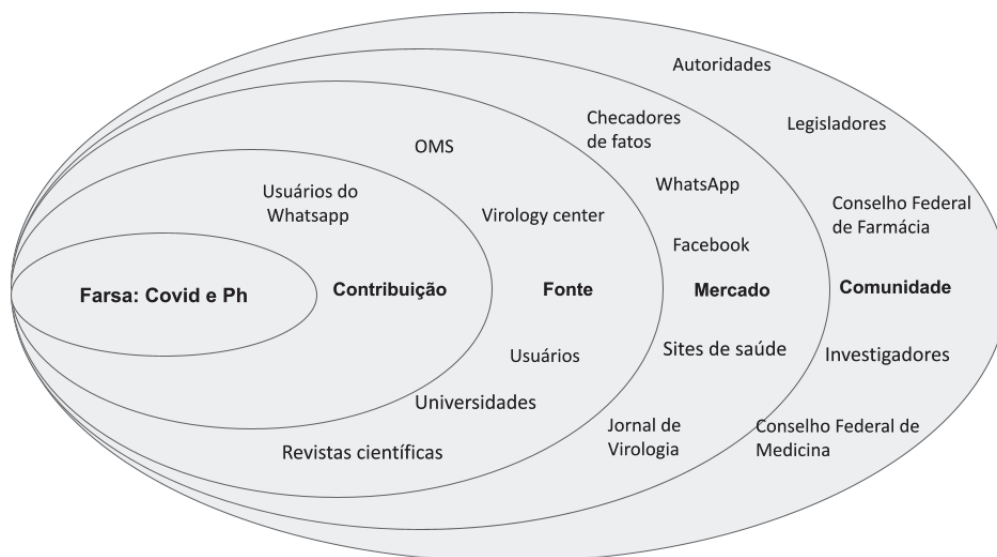


Figura A.25: Partes interessadas (Covid e pH)

### A.7.3 Análise Semiótica do Ciclo de Vida

O objeto desinformativo não era sofisticado, com uma estrutura simples e um texto genérico, que apelava para curiosidade do colaborador para que abrisse um arquivo infectado. Mesmo com os mecanismos técnicos de segurança detectando o e-mail como possível spam e encaminhando a mensagem para caixa de spam, um colaborador abriu a mensagem. A seguir, uma descrição detalhada dos aspectos que influenciaram no desfecho do caso, discutidos para cada estágio do ciclo da desinformação.



O criador do objeto pode ter sido qualquer "Usuário do WhatsApp", e não há garantias de que este objeto seja original do Brasil. O objeto desinformativo é puramente textual, e tem uma estrutura simples, com cabeçalho, uma breve introdução e uma lista de itens. O tema explora o contexto extraordinário da pandemia, que influencia o estado emocional (medo e incerteza) e o pensamento lógico (intenção de se informar, cuidar da saúde, busca por tratamentos) da audiência. Há indícios do objetivo dos criadores ser ganho ideológico, como a propaganda da eficácia de tratamentos alternativos, na tentativa de influenciar a percepção da audiência ("Outros usuários do WhatsApp") do risco da pandemia. Há a indicação de "Autoridades científicas" como fonte da desinformação, e o conteúdo deturpa os resultados do artigo original referenciado (Joel, 2021). O objeto foi difundido no aplicativo WhatsApp, que possui mecanismos que facilita a difusão para outros grupos e usuários, além carecer de recursos de mitigação de desinformação (Puska et al., 2020).

A Figura A.26 apresenta aspectos sociotécnicos identificados no objeto.

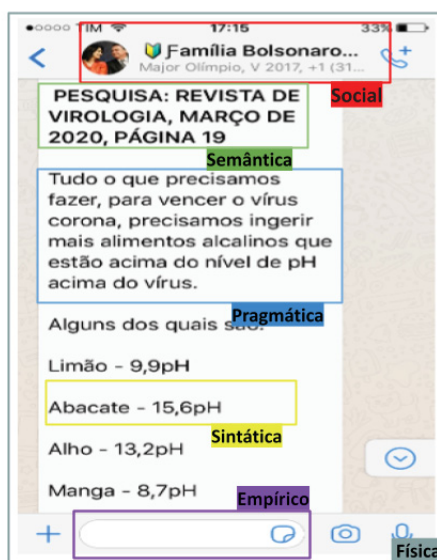


Figura A.26: Análise de Farsa no WhatsApp.

Foram identificados 127 aspectos sociotécnicos, 94 do sistema de informação humana e 33 da plataforma tecnológica. Os resultados podem ser visualizados na planilhas do Google Sheets<sup>36</sup>. O Quadro A.23 apresenta um resumo dos aspectos sociotécnicos e exemplos.

<sup>36</sup>[https://docs.google.com/spreadsheets/d/analise\\_p/h](https://docs.google.com/spreadsheets/d/analise_p/h)

Tabela A.23: Resumo dos aspectos sociotécnicos identificados na análise

| Dimensão Semiótica | Aspecto                                                                                                                                                                        | Exemplos                                                                                                                                                                                                                                                                                                                                                        |
|--------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Social             | Popularidade do tema;<br>Impacto social;<br>Crise de saúde pública;<br>Polarização política;<br>Hábitos de consumo de notícias;<br>Bolhas sociais;<br>Legislação insuficiente; | Explora o contexto extraordinário de uma pandemia;<br>A novidade da doença;<br>Usa autoridades científicas para influenciar consensos;<br>Promove uma percepção desinformada da realidade da pandemia que é custosa para a saúde pública;<br>WhatsApp é um aplicativo popular no contexto, muito usado para consumo de informações (notícias, novidades, etc.); |
| Pragmática         | Exploração de expectativas, intenções, e necessidades;<br>Explora percepção de autoridade e credibilidade da fonte e da mensagem;<br>Propósito do uso da tecnologia;           | Indivíduos tendem a se informar sobre a doença, seus tratamentos e como se proteger;<br>Autoridade certificadora e a alteração das informações do artigo original/referência;<br>Credibilidade de informações recebidas pelo Smartphone/WhatsApp;                                                                                                               |
| Semântica          | Percepções da audiência;<br>Vieses cognitivos;<br>Inexperiência/Ignorância;                                                                                                    | Ignorância sobre a pandemia;<br>Ignorância sobre biologia e pH;<br>Informações significativas e representativas;<br>Medo e incerteza.<br>Diminui a percepção de risco da pandemia.                                                                                                                                                                              |
| Sintática          | Linguagem;<br>Diagramação;<br>Perfis de usuários;                                                                                                                              | A forma e o tipo do objeto desinformativo;<br>O cabeçalho e um corpo que chama a atenção para a listagem de consumíveis;<br>A estrutura do texto reverso, construído a partir de uma argumentação conclusiva;<br>Termos técnicos não são de amplo entendimento (escala do pH = 0<>14; pH do abacate 15,6);                                                      |
| Empírica           | Criptografia;<br>Tempo de leitura;<br>Interação com a comunicação;<br>Restrições de acesso;                                                                                    | Canal de comunicação bidirecional possibilita discussões;<br>WhatsApp carece de mecanismos para auxiliar na mitigação de desinformação;<br>Planos de acesso restringem capacidade de usuários verificarem mensagens recebidas;                                                                                                                                  |
| Física             | Mídias sociais;<br>Penetração de Internet;<br>Limitações dos dispositivos;                                                                                                     | Variedade de canais para difusão;<br>Tamanho do objeto desinformativo;<br>Quantidade de dispositivos móveis com acesso Internet;                                                                                                                                                                                                                                |

Aspectos como, a popularidade do tema, a pandemia, a legislação de regulação da desinformação digital precária, favorecem o surgimento (Watney, 2018; Kshetri e Voas, 2017). O uso de autoridade certificadora significativa, o uso da linguagem pessoal, e a intenção de cuidado expressado na mensagem são características persuasivas ligadas ao princípio da **Afinidade**, e amplamente usadas na construção de discursos apelativos (Li e Santos Jr, 2012). O objeto desinformativo é puramente textual, sem links para a fonte, isso somado a planos de acesso a Internet restritivos, dificulta a verificação da informação no objeto desinformativo. A quantidade

de dispositivos móveis conectados a Internet no Brasil, se traduz em mais usuários do WhatsApp, o que o torna um vetor popular para espalhar desinformação.

#### A.7.3.1 *Leitura e Análise dos Resultados*

A criação do objeto desinformativo foi uma atividade intencional. Há indícios de que a desinformação foi criada para explorar os princípios da **Escassez** (ignorância sobre o funcionamento do vírus; apresentar uma solução sobre o tratamento do vírus; o uso de termos de urgência como "tudo que precisamos fazer");, **Prova Social** (quando incerto, o indivíduo tende a seguir o comportamento do grupo;), e da **Unidade** (Grupo tem alinhamento ideológico de crenças; uso de palavras como "nós" e "precisamos"). O rótulo "Pragmática" apresentado na Figura A.22, representa um aspecto da **intenção falsa** do criador, com elementos que exploram a **Afinidade** e **Unidade**. O texto é associado ao viés ideológico do grupo naquele tempo<sup>37</sup>, apelando para os Princípios da Influência de **Comprometimento**, **Coerência** e **Reciprocidade** (solicitar uma ação) (Cialdini, 1993).

Em relação a difusão, o "WhatsApp" agrupa partes interessadas com interesses parecidos. A organização das pessoas em grupos facilita o recebimento de desinformação com o viés do grupo (Vicario et al., 2019), o que condiz com a realidade social brasileira, apontado pela polarização política do grupo em que o objeto desinformativo foi identificado. A intenção dos usuários do WhatsApp de se comunicar e os mecanismos do WhatsApp que facilitam o compartilhamento com poucas ações do usuário (Puska et al., 2020), ampliam a capacidade de difusão do objeto desinformativo. As configurações de qualidade de serviço e segurança da transmissão, como as restrições a um tipo de tráfego na rede e a criptografia, dificultam a detecção e facilitam a exposição seletiva de desinformação para algumas partes interessadas (**Reciprocidade**). O tamanho do objeto desinformativo é uma característica relevante que facilita a difusão, com um conteúdo sem imagens ou vídeos que consomem recursos do serviço de acesso.

Em relação ao consumo, a desinformação deliberada chama a **atenção** da audiência por ser um tema popular, e explora o desejo e expectativa das partes interessadas de se informar e se proteger. A **Autoridade** certificadora e aspectos relacionados a **Afinidade** como o viés ideológico de confirmação - tendência de aceitar perspectivas que confirmam as do indivíduo -, induzem a aceitação da desinformação (Li e Santos Jr, 2012; Flintham et al., 2018; Bailin e Siegel, 2002). A linguagem emocional, e dados técnicos são aspectos que influenciam o julgamento de credibilidade (Li e Santos Jr, 2012; Bailin e Siegel, 2002). Por exemplo, o pH é uma informação técnica que requer treinamento para ser interpretado corretamente, e termos como "vencer" associam um viés positivo para a desinformação. O texto reverso explora o princípio da **Reciprocidade**, construído a partir de uma argumentação conclusiva, junto com uma semântica que explora o princípio de **Unidade**, como a segurança e o pertencimento ao grupo, como a afirmação "tudo o que *precisamos* fazer para nos proteger", são capazes de induzir uma conclusão pretendida pelo criador do objeto (Teixeira e e Silva, 2007).

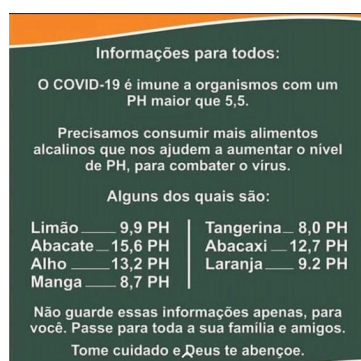
O objeto desinformativo explora aspectos da **Afinidade**, como o título em negrito, o tamanho do texto, a listagem, que facilitam e o direcionam o fluxo de leitura. Além disso, o meio de comunicação bidirecional do WhatsApp, possibilita conversações e discussões, levando a resolução de conflitos (de Jongab, 2009), com um potencial de influência dos princípios de **Prova Social** e **Comprometimento**. Os hábitos de multitarefa são características de uso do WhatsApp no Brasil (Newman et al., 2019), o que pode impactar na capacidade de atenção do usuário e levá-lo a consumir desinformação sem verificar (Shen et al., 2019).

<sup>37</sup>O posicionamento público de autoridades do estado ligadas ao propósito do grupo (Vicario et al., 2019)

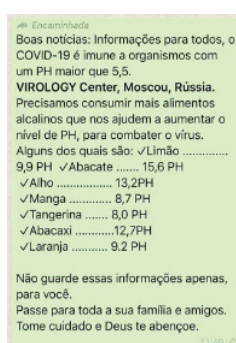
Em relação a mitigação, existem indícios de que os brasileiros possuem hábitos pouco saudáveis de consumo de informações, como usar o WhatsApp como fonte confiável informações, se informar em portais informativos não tradicionais e consumir informações sem verificar (Hölig et al., 2020). A plataforma de comunicação não tem um posicionamento ativo quanto ao bloqueio e correção de desinformação. Faltam mecanismos que auxiliem os usuários de maneira significativa. Os "Checadores de Fatos" corrigiram a desinformação em março de 2020 (Uol, 2020b), porém este é um meio pouco acessado pelos Usuários do WhatsApp (Hölig et al., 2020), ainda mais aqueles com planos restritivos de acesso à Internet móvel (Puska et al., 2020), como ocorre no Brasil. Estes são exemplos de aspectos sociotécnicos inter-relacionados que impactam nas estratégias de mitigação de checadores de fatos e propagação de correções.

### A.7.3.2 Reiteração

Entre abril e maio de 2020, versões diferentes do objeto desinformativo viralizaram via WhatsApp, no Brasil, na Índia, na África do Sul e em Portugal (CE, 2020; Times, 2020; Joel, 2021). A primeira aparição em órgãos de checagem de fatos foi em meados de Março de 2020 na África do Sul (Joel, 2021) e no Brasil (Uol, 2020b), nos quais o objeto desinformativo referenciava um artigo publicado em 1991 em uma revista de virologia (Joel, 2021). Em Abril de 2020, o objeto desinformativo foi detectado em um novo formato no Facebook, sendo identificada pela agência de checagem de fatos africana Africacheck.org (Joel, 2021). Depois disso, as correções foram difundidas pela aliança internacional de verificadores de fatos do Coronavírus, o CoronavirusFacts (Pointer, 2021; Johnson, 2020). Ainda, em julho de 2020, mesmo após a desinformação ter sido identificada e as correções divulgadas por checadores de fatos, foi registrado o reaparecimento no Brasil, como parte de mensagens maiores, em meio a dicas de prevenção e tratamento alternativo, atribuídas ao Ministério da Saúde (MS) (Almirante e Marés, 2020). A Figura A.27 mostra as variações da desinformação.



(a) Portugal.



(b) Brasil.

[4/1, 10:04 PM] Karen Allen: This is to inform us all that the pH for corona virus varies from 5.5 to 8.5.

All we need to do, to beat corona virus, we need to take more of an alkaline foods that are above the above pH level of the virus.

Some of which are:  
**Lemon - 9.9pH**  
**Lime - 8.2pH**  
**Avocado - 15.6pH**  
**Garlic - 13.2pH**  
**Mango - 8.7pH**  
**Tangerine - 8.5pH**  
**Pineapple - 12.7pH**  
**Dandelion - 22.7pH**  
**Orange - 9.2pH**

(c) Índia

Figura A.27: [

Variantes do objeto desinformativo. Fontes: IndianTimes (Times, 2020), Lupa (Afonso, 2020) e Polígrafo (Sapo, 2021).

A desinformação apresentava variações no formato e estrutura da mensagem. O objeto desinformativo foi transformado em uma imagem, ou traduzido para outros idiomas. Em algumas versões, a mensagem alterava (G1, 2021b) ou não referenciava nenhuma fonte para os dados (Johnson, 2020). As variações são exemplos de novos ciclos originados por outros objetos desinformativos. Isso implica que o rastreamento da origem da mensagem não é uma tarefa trivial, e sua mitigação não foi efetiva.

O tema de tratamento alternativo, explorado em reiterações no evento desinformativo, indica a construção de uma narrativa. Isso não quer dizer que o evento foi planejado, porém é um indício de que o tema foi explorado continuamente, possivelmente pelo impacto de aspectos do contexto, como a popularidade do tema, a ignorância sobre o vírus, os vieses de confirmação e ancoragem em grupos polarizados, e as funções técnicas da plataforma tecnológica de comunicação que dificultam a verificação. Assim, a desinformação foi usada para fomentar discussões em torno do tema, indicando a tentativa de manipular as conclusões e ações das partes interessadas, ou causar confusão e sobrecarregar as partes interessadas de informações e problemas desnecessários.

Uma investigação mais elaborada, considerando um contexto maior, surge como necessidade para caracterização dos impactos e explorações relacionados ao tema. Para isso, uma investigação mais abrangente é necessária, com mais informações sobre o surgimento dos objetos desinformativos, como as informações privadas de difusão, origem, destino e tempo. A migração do objeto desinformativo entre organizações internacionais, como as comunidades brasileira e da África do Sul, mostra a importância de iniciativas e atividades de mitigação integradas, representando um desafio nada trivial para mitigação. Além disso, o trabalho colaborativo entre organizações privadas e públicas, amplia a capacidade de gerenciamento da desinformação, e representa outro desafio de integração entre organizações para mitigação do ressurgimento. A Figura A.28 apresenta uma representação do estágio de Reiteração no evento analisado.

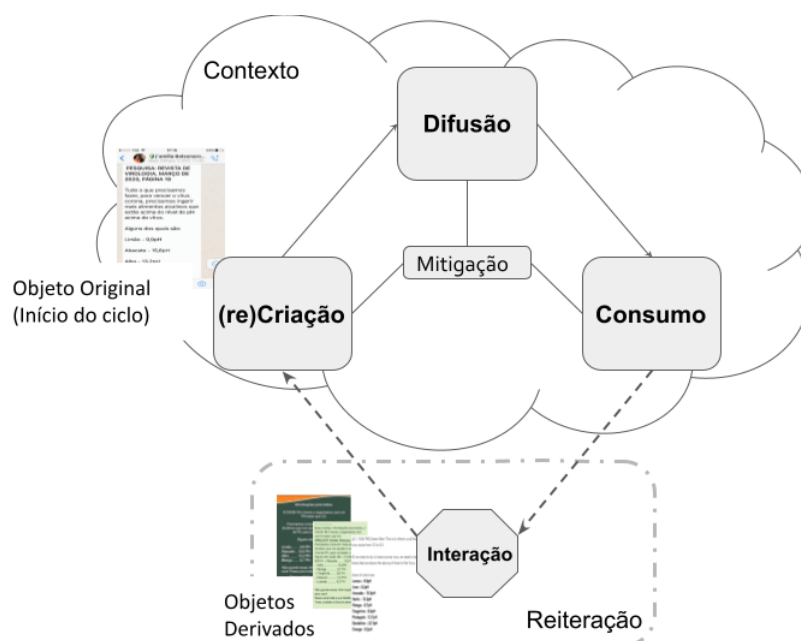


Figura A.28: Ciclo de vida da desinformação - Reiteração.

As características culturais das organizações implicam em interações contextualizadas carregadas de subjetividade. Os consensos explorados pelo conteúdo do objeto desinformativo, consistem de estereótipos, visões de mundo, e normas sociais que são compartilhados pelo grupo em que a mensagem foi enviada. A temporalidade significa a existência de consensos particulares em uma organização, vinculados a um período e uma autoridade. Nesse sentido, a reutilização dos objetos desinformativos em outro momento, explora vulnerabilidades relacionadas a mitigação ineficaz, causando confusão e desconforto enquanto o conflito entre consenso do grupo e correções durarem. Assim, parece existir uma relação entre a exploração das expectativas da audiência e a construção de uma narrativa contínua em torno do tema de tratamento alternativo,



que foi explorado intermitentemente em comunidades de usuários vulneráveis, facilitando as tentativas de exploração de várias pessoas ao mesmo tempo.

O próximo passo será analisar a construção da narrativa em torno da temática explorada pelo objeto desinformativo do segundo caso. Podemos enquadrar o contexto do evento desinformativo em torno da narrativa com viés ideológico relacionado ao tratamento alternativo, e investigar as relações entre diferentes objetos desinformativos em diferentes contextos em uma mesmo evento.

A mitigação de uma narrativa deve se preparar para adaptação do contexto ao longo do tempo. Isso pode ser feito com bancos de eventos desinformativos, que reúnam os objetos desinformativos identificados, os aspectos sociotécnicos analisados e as estratégias de mitigação empregadas. Estas informações poderão ser usadas para comparação com futuras reiterações da desinformação. Ainda, são estratégias necessárias em conjunto com a divulgação de informações corretivas, esclarecer as motivações, mostrar comparativos com outros grupos, e apresentar as consequências sociais da desinformação para conscientização efetiva (Fernandez e Alani, 2018).

#### *A.7.3.3 Considerações sobre a análise*

A análise revelou aspectos sociotécnicos inter-relacionados nos os estágios do ciclo de vida da desinformação deliberada. Por exemplo na criação, aspectos como a situação do contexto (Crise, Pandemia), a popularidade do tema ("saúde"), e o anonimato proporcionado pela Internet, são características que influenciam o surgimento de desinformação sobre o tema. Na difusão, a natureza do meio de comunicação assíncrono de duas vias, a popularidade do aplicativo WhatsApp e suas características que facilitam o compartilhamento de mensagens para usuários em quantidade, se somam aos aspectos informais da temática e do momento que partes interessadas experienciam, ampliando a capacidade de difusão do objeto desinformativo. O consumo foi impactado pela incerteza e o medo relacionados a pandemia, os hábitos dos brasileiros ao consumir informações sem verificação, e aspectos formais e técnicos relacionados a literacia digital e a educação das partes interessadas. Estes aspectos são exemplos de como o SI humanas e as plataformas tecnológicas influenciam os estágios de criação, difusão e consumo de modo abrangente.

No evento desinformativo, o objeto tem um potencial de difusão que engloba a população total de smartphones com acesso ao WhatsApp no Brasil - cerca de 120 milhões de usuários (Hölig et al., 2020), de uma população de aproximadamente 220 milhões de pessoas. Em um contexto com essas dimensões, as estratégias de mitigação precisam ser organizadas e abrangentes (Starbird et al., 2019), que corrijam desinformações nos locais onde o usuário consome suas informações (Fernandez e Alani, 2018), e que sejam representativas e significativas para eles (Ecker et al., 2014). No período analisado, o contexto do evento desinformativo carecia de estratégias de mitigação organizadas, com atividades desenvolvidas por poucos órgãos do setor privado - somente o trabalho jornalístico não é suficiente para mitigar a desinformação em situações de crise (Starbird et al., 2018) -, e quase nenhum movimento do setor público. A incerteza e o medo são aspectos relacionados aos Princípios de Influência, como **Aprovação Social** e **Escassez**, que são facilmente explorados em situações de crise (Cialdini, 1993), quando não se tem uma estratégia ampla de mitigação com campanhas preventivas (Basol et al., 2021).

A informação precisa chegar aos usuários no local que estão habituados a consumir informações. A impossibilidade de integração do WhatsApp com *plugins* capazes de promover a verificação no ambiente de comunicação do usuário, e a violações da neutralidade da rede das operadoras, impõe ainda mais complexidade para mitigação, como propagação de correções. Nesse sentido, campanhas informativas são ferramentas úteis para mitigar os efeitos nocivos da desinformação em populações abrangentes (Basol et al., 2021), em especial quando se pode prever



que a desinformação sobre um tema irá surgir, como em situações de crise (Zhu et al., 2018). Estas campanhas precisam de uma legislação que regule os direitos e deveres de plataformas de comunicação, como as campanhas de vacinação fomentadas pelo setor público de 1990 2015, onde se reservavam horários para transmissão na Televisão e no Rádio (Teixeira et al., 2020). Atualmente, existem outros canais de comunicação que também poderiam ser regulados da mesma forma, como os *streams* (Youtube, Netflix, etc.), que são locais bastante acessados pelos usuários no Brasil (Hölig et al., 2020), e que chamam a atenção (Trattner e Kappe, 2013).

As campanhas informativas podem auxiliar na conscientização dos grupos, podendo neutralizar o objeto desinformativo sem a necessidade de apelar para a censurar (Karduni et al., 2018; Raynal e Gaspard, 2010). Entretanto, apenas propagar informações tende a encontrar resistência dos indivíduos (Pal et al., 2018). Para efetividade deste tipo de campanha, é necessário maior qualidade da informação para tornar a mensagem significativa para a audiência. Nessa direção, incorporar explicações (Vraga e Bode, 2018; Chan et al., 2017), apresentar motivações (Cobb et al., 2013), e exemplos de consequências (Fernandez e Alani, 2018), são estratégias promissoras.

As redes sociais precisam encontrar formas significativas para as diferentes partes interessadas, de alertar o nível de risco que uma desinformação representa, considerando o aspecto subjetivo na interpretação de um conteúdo. Por exemplo, estratégias como uso Rótulos em mensagens não é efetivo, pois a racionalização dos usuários depende do contexto e como eles percebem a mensagem, possibilitando a negligência dos avisos (Gao et al., 2018; Pal et al., 2018), e impactos do viés de ancoragem que tem o potencial de influenciar a propagação da desinformação (Ecker et al., 2014). A tecnologia pode ser explorada pelo uso de sistemas autônomos (bots) de comunicação que propagam informações corretivas, como fizeram algumas iniciativas (Melo et al., 2019). Nesse caso, os obstáculos consistem da dificuldade de aceitar correções e usar ferramentas digitais, que podem ser mitigados por ações de educação e conscientização dos usuários.

#### A.7.3.4 *Análise da conspiração do tratamento preventivo*

## A.8 APÊNDICE 8 - PROTOCOLO DAS OFICINAS I E II

Esta seção apresenta o desenho do estudo para avaliação da abordagem sociotécnica para apoiar o estudo do fenômeno da desinformação. Serão avaliadas a significância, utilidade e efetividade da proposta com o público alvo.

### A.8.1 O estudo - ( 120 min)

Os projetistas de estratégias e mecanismos de mitigação da desinformação frequentemente desenvolvem ou empregam modelos com base em características do contexto do problema que enfrentam, para projetar suas soluções. Eles se apoiam em ferramentas epistêmicas para analisar o contexto e extrair os aspectos relevantes para a solução do problema. Como a literatura reporta, as soluções sendo propostas precisam lidar com a multidisciplinaridade que caracteriza o fenômeno da desinformação digital (Fernandez e Alani, 2018). Entretanto, as caracterizações do fenômeno da desinformação tendem a desconsiderar aspectos do contexto social e organizacional integrados aos aspectos técnicos e humanos (Starbird et al., 2019; Fernandez e Alani, 2018), e tem foco nos estágios da comunicação de desinformação, segmentando o fenômeno sociotécnico. Estas caracterizações são pouco representativas da realidade que enfrentam, levando a limitações na análise e definição de problemas, restringindo a capacidade das soluções geradas em mitigar a desinformação deliberada. Nesse sentido, propomos uma abordagem para análise de ocorrências de desinformação na comunicação para apoiar o estudo do fenômeno de modo sistêmico e sociotécnico por projetistas de soluções computacionais. Este estudo pretende avaliar a utilidade e significância da abordagem com o público-alvo.

**Objetivo:** Avaliar a significância da abordagem e a utilidade percebida com relação a atividade de análise e caracterização do contexto de ocorrências de desinformação na comunicação, do ponto de vista de projetistas de soluções computacionais interessados no fenômeno de desinformação, no sentido de apoiar o estudo de aspectos que impactam a criação, difusão e consumo da desinformação que devem ser considerados ao projetar uma estratégia de mitigação.

**Questões de Pesquisa:** Qual a percepção dos participantes sobre a significância da abordagem? Qual a percepção dos participantes sobre a utilidade da abordagem?

**Métricas:** Utilidade Percebida; Facilidade de Uso Percebida; a Auto-eficácia Percebida, e Relevância para o Trabalho percebidas pelos participantes quanto ao uso da abordagem (Fathema et al., 2015).

#### A.8.1.1 Participantes

O estudo será realizado com pesquisadores em Interação Humano-Computador (IHC) do laboratório de IHC da UFPR.

#### A.8.1.2 Atividades a executar

Os participantes utilizarão a abordagem individualmente para análise de um evento desinformativo, onde deverão caracterizar as partes interessada, identificar aspectos sociotécnicos que impactam na criação, difusão e consumo de desinformação no contexto analisado e, sintetizar e registrar os resultados da avaliação nos formulários para este fim.

Além do preenchimento dos materiais da abordagem, os participantes responderão a um questionário desenvolvido para atingir os objetivos do estudo. Cada participante aplica a abordagem para analisar o mesmo evento desinformativo, recebem as mesmas instruções, e respondem as mesmas perguntas, individualmente. Serão analisados dois objetos desinformativos

em dois eventos diferentes previamente selecionados: Uma desinformação acidental e uma fake news;

Modo:

- Apresentação e assinatura do Termo de Consentimento Livre e Esclarecido (ApêndiceA.10).
- Apresentação sobre a perspectiva sociotécnica da desinformação e o ciclo de vida;
  - Os participantes terão acesso antecipado ao encontro de um documento para contextualização. O documento consiste de um artigo de 2 páginas sobre a perspectiva sociotécnica do fenômeno da desinformação, um entendimento sobre a desinformação acidental e deliberada no formato digital, o ciclo de vida da desinformação (5 min de leitura). No dia da sessão, os participantes assistirão a um apresentação sobre conteúdo do artigo. Será apresentado um exemplo de ocorrência de desinformação, passando pelos estágios do ciclo - material de consulta ( 5 min).
- Apresentação da abordagem e das atividades de avaliação aos participantes;
  - Apresentar a abordagem, como conduzir a avaliação e o material de apoio para aplicação da abordagem - Materiais: Diagrama de partes interessadas e Escada Semiótica; Quadros com o material de apoio para identificação de aspectos sociotécnicos (tipos de desinformação, princípios de influência, tipos de interação, e falácias lógicas) - ( 15 min).
    - \* Apresentar os casos de desinformação a serem analisados.
    - \* Os participantes irão interagir com a abordagem por um período de teste, seguindo um exemplo de análise para aprimorar a eficiência e eficácia na aplicação ( 10 min).
    - \* Avaliação de evento de desinformação ( 50min). Os participantes aplicam o método e registram os dados nos artefatos. Os participantes serão guiados na aplicação da abordagem por um passo a passo, e o pesquisador responsável marcará o tempo de execução de cada etapa, (15 20 min).
- Condução da avaliação e coleta de dados;
  - Coleta de dados para observação do uso, entrevista com os participantes e por um questionário pós-teste (questões objetivas - Google Forms) ( 10 min).
    - \* Formulário de avaliação embasado em Technology Acceptance Model (TAM) e Escala Likert;
      - Questões relacionadas a caracterização dos participantes: idade, sexo, escolaridade, experiência com computação, cargo em que trabalha e tempo de experiência no mercado;
      - Questões relacionadas à capacidade da abordagem (TAM3);
- Análise dos dados coletados;
- Relato dos resultados.

#### A.8.2 Questionários

Questões de Utilidade e Facilidade de Uso Percebidas usadas para coleta de dados (coleta com Likert):

- A abordagem me apoiou a identificar aspectos sociotécnicos relevantes do evento desinformativo analisado;

- A abordagem me apoiou a caracterizar o contexto do evento desinformativo analisado;
- A abordagem me fez pensar em problemas que eu dificilmente teria identificado sem usá-la;
- Considero a abordagem útil para identificar aspectos sociotécnicos no evento desinformativo analisado;
- Considero a abordagem útil para analisar eventos desinformativos de modo abrangente;
- Considero que a abordagem é clara e fácil de entender;
- A abordagem ajudou a me sensibilizar para a importância de se pensar em aspectos sociotécnicos na comunicação de desinformação;
- Considero fácil utilizar a abordagem para identificar problemas de desinformação na comunicação digital;
- Os materiais usados na etapa de análise semiótica da comunicação (Etapa II da abordagem) apoiaram a atividade de identificação de problemas;
- Considero os materiais de apoio (princípios de influência, falácias, lógicas, tipos de interação) úteis para aplicação de abordagem;

Questões abertas:

- Quais os pontos positivos da abordagem para avaliação de redes sociais online sob a perspectiva de valores? No que o método contribuiu, ajudou, etc.
- Quais os pontos negativos do método? Dificuldades, como eles poderiam ser melhorados, etc.

Questões sobre Auto-Eficácia Percebida: (coleta com Likert):

- Eu me sinto confiante em aplicar a abordagem.
- Eu tenho as habilidades necessárias para usar a abordagem.
- Eu poderia completar o trabalho usando a abordagem. . .
  - . . . se não houvesse ninguém por perto para me dizer o que fazer.
  - . . . se eu tivesse apenas os recursos de ajuda embutidos para assistência.
  - . . . se alguém me mostrar como fazê-lo primeiro.

Questões sobre Relevância para o Trabalho (coleta com Likert):

- No meu trabalho, o uso da abordagem é importante;
- No meu trabalho, o uso da abordagem é relevante;
- O uso da abordagem é pertinente às minhas várias tarefas relacionadas ao trabalho;

## A.9 APÊNDICE 9 - PROTOCOLO DA OFICINA III

### Plano de uma oficina de análise de desinformações para avaliação da Dismiss

#### A.9.1 O estudo - ( 120 min)

O fenômeno da desinformação digital é um problema complexo que envolve tanto a tecnologia quanto a sociedade. Nesse sentido, as soluções sendo propostas para mitigar os problemas da desinformação digital precisam lidar com a complexidade dessa natureza sociotécnica característica do fenômeno (Fernandez e Alani, 2018). Entretanto, no melhor do nosso conhecimento, as ferramentas e artefatos conceituais propostos para apoiar a análise abrangente do fenômeno tendem a desconsiderar aspectos do contexto social integrados aos aspectos técnicos (Starbird et al., 2019; Fernandez e Alani, 2018). Sem uma visão sociotécnica abrangente, que entenda como a tecnologia e a sociedade interagem para produzir e espalhar desinformação, as soluções ficam limitadas, restringidas pelo entendimento insuficiente de problemas relacionados ao fenômeno. Considerando essas limitações desenvolvemos a Dismiss, uma abordagem para apoiar o diagnóstico de casos de desinformação de modo sistêmico e sociotécnico por pesquisadores e projetistas de soluções computacionais interessados no fenômeno. Assim, propomos esta oficina para avaliar a utilidade percebida e a facilidade de uso percebida da Dismiss com o público-alvo.

#### A.9.2 Introdução

O objetivo desta oficina é avaliar a Aceitação de Tecnologia (TAM) da Dismiss - uma abordagem para a análise sociotécnica da desinformação digital. A Dismiss foi desenvolvida em dois níveis de complexidade construtivos. O primeiro nível, chamado de Estratégia de Análise Sociotécnica de Objetos Desinformativos engloba as técnicas e artefatos básicos para conduzir uma análise sociotécnica de um objeto desformativo em um único canal de comunicação. O segundo nível, chamado de Estratégia de Análise Sociotécnica do Ciclo de Vida da Desinformação, expande a análise de objetos para abordar diferentes estágios do seu ciclo de vida possibilitando um estudo abrangente que engloba as transformações do objeto e o entendimento de aspectos inter-relacionados nos estágios do ciclo de vida de um modo sistêmico.

Esta oficina é proposta com um foco para avaliar as melhorias na Estratégia de Análise Sociotécnica de Objetos e na Análise do Ciclo de Vida. A oficina será realizada em três fases. Na primeira fase, os participantes receberão uma visão geral da Dismiss e suas principais características. Na segunda fase, os participantes serão solicitados a auxiliar o facilitador (autor) na aplicação da Dismiss para analisar um caso real de desinformação. Na terceira fase, os participantes serão convidados a preencher um questionário para avaliar a utilidade percebida e a percepção de facilidade de uso da Dismiss.

**Grupo de participantes:** Pesquisadores e analistas que trabalham ou têm interesse no campo da análise de desinformação digital.

#### A.9.3 Fase 1: Introdução à Dismiss (20 minutos)

**Boas-vindas e introdução:** O facilitador (autor) dará as boas-vindas aos participantes e apresentará o objetivo da oficina. O facilitador irá se apresentar e explicar brevemente a sua experiência no campo da desinformação e apresentará o TCLE.

**Visão geral da Dismiss:** O facilitador fornecerá uma visão geral da Dismiss, incluindo seus principais recursos, objetivos e benefícios.



**Sessão de perguntas e respostas:** Os participantes terão a oportunidade de fazer perguntas e esclarecer quaisquer dúvidas que possam ter sobre a Dismiss. O facilitador encorajará os participantes a fazer perguntas e fornecer respostas claras e concisas.

#### A.9.4 Fase 2: O estudo de caso (60 minutos)

**Introdução ao estudo de caso:** O facilitador deve apresentar um caso real de desinformação para os participantes analisarem. O facilitador guiará os participantes para ilustrar como a Dismiss pode ser aplicada a situações do mundo real. Os participantes apoiarão a aplicação da Dismiss, auxiliando na coleta de dados e preenchimento dos artefatos da abordagem de modo individual.

##### **Materiais:**

- A Dismiss e suas técnicas e artefatos propostos para análise;
  - Inclui o Ciclo de Vida da Desinformação digital, Análise de Enquadramento, Diagrama de Partes Interessadas, Escada Semiótica e os Formulários para anotação dos dados.
- Um conjunto de tarefas projetadas para testar a utilidade percebida da Dismiss.

##### **Procedimento:**

- Os participantes receberão um conjunto de tarefas a serem executadas usando a Dismiss. O facilitador guiará a utilização da ferramenta, observará os participantes enquanto concluem as tarefas e anotará suas interações com a Dismiss. Tarefas:
  - Explorar o contexto do caso para identificar os elementos básicos para análise (objetos desinformativos, canais de comunicação e partes interessadas) e preencher os artefatos da Dismiss;
  - Análise sociotécnica do caso para identificar aspectos sociotécnicos que caracterizam o caso e preencher os artefatos da Dismiss;
  - Análise dos dados coletados para identificar e entender problemas no caso e preencher os artefatos da Dismiss;

#### A.9.5 Fase 3: Avaliação de Aceitação de Tecnologia (40 min)

**Introdução à avaliação:** o facilitador fornecerá uma breve visão geral do Modelo de Aceitação de Tecnologia (TAM), incluindo seus principais componentes.

**Distribuição do Questionário:** os participantes receberam um questionário que foi elaborado para avaliar o TAM da Dismiss.

**Pesquisa:** os participantes serão solicitados a responder ao questionário que mede suas atitudes em relação à Dismiss baseada no Modelo de Aceitação de Tecnologia (TAM). As perguntas da pesquisa se concentrarão nos seguintes pontos:

- Utilidade percebida: em uma escala de 1 a 5, o quanto você concorda com as seguintes afirmações sobre a Dismiss?
  - A Dismiss me ajudou a entender a natureza complexa do caso de desinformação.
  - A Dismiss foi útil para identificar a problemas no caso de desinformação.

- A Dismiss me ajudou a identificar aspectos sociotécnicos para serem considerados no desenvolvimento de soluções que ficariam ocultos sem o uso da abordagem.
- Comentários adicionais: Forneça qualquer feedback ou comentário adicional que você possa ter sobre a utilidade da Dismiss.
- Percepção de facilidade de uso: em uma escala de 1 a 5, o quanto você concorda com as seguintes afirmações sobre a Dismiss?
  - A Dismiss é fácil de usar.
  - A Dismiss é fácil de entender.
  - Considero fácil de integrar a Dismiss em minhas atividades de trabalho.
  - Comentários adicionais: Forneça qualquer feedback ou comentário adicional que você possa ter sobre a facilidade de uso da Dismiss.
- Intenção de uso: em uma escala de 1 a 5, o quanto você concorda com as seguintes afirmações sobre sua intenção de usar a Dismiss?
  - Acredito que a Dismiss será útil em meu trabalho.
  - Acredito que usar a Dismiss aumentará a qualidade de soluções geradas.
  - Comentários adicionais: Forneça qualquer feedback ou comentário adicional que você possa ter sobre a sua intenção de uso da Dismiss.
- Comentários adicionais: Forneça qualquer feedback ou comentário adicional que você possa ter sobre a Dismiss.

**Preenchimento do Questionário:** Os participantes terão 20 minutos para completar o questionário. O facilitador irá instruir sobre como preencher o questionário e responder a quaisquer perguntas que os participantes possam ter.

**Grupo Focal:** após o preenchimento do questionário, o facilitador conduzirá uma discussão em grupo (20 minutos) sobre as experiências dos participantes quanto ao uso da Dismiss. A discussão se concentrará na identificação dos pontos fortes e fracos da Dismiss, bem como nas áreas de melhoria.

**Análise de dados:** as respostas da pesquisa feita durante a avaliação, os artefatos preenchidos produzidos na atividade, e as anotações relacionadas a observação dos participantes durante a atividade serão analisadas para identificar temas relacionados às questões do TAM.

**Conclusão:** o facilitador irá resumir as principais conclusões da oficina e fornecer algumas recomendações sobre como a Dismiss pode ser melhorada. O facilitador irá agradecer aos participantes por sua participação e incentivá-los a fornecer *feedback* futuros.

## A.10 APÊNDICE 10 - TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

**Termo de Consentimento Livre e Esclarecido (TCLE)**

Nós, Alisson Andrey Puska e Prof. Dr. Roberto Pereira, da Universidade Federal do Paraná, estamos convidando o senhor(a) a participar de um estudo intitulado “Estudo de Caso para o entendimento da significância e utilidade da abordagem para apoio do estudo de desinformações digitais”. Esta pesquisa é importante para a compreensão da significância e efetividade da abordagem em apoiar os participantes na caracterização dos estágios do ciclo de vida da desinformação deliberada, das partes interessadas, do campo de informações, e na identificação de aspectos sociotécnicos relevantes para projetar soluções.

Aspectos importantes desta pesquisa:

- Sua participação na pesquisa não é obrigatória.
- Caso você participe da pesquisa, será necessário realizar atividades sobre o design de soluções computacionais, que envolvem a análise de situações-problema envolvendo desinformação deliberada no formato digital, e o uso de ferramentas de apoio. Para participar deste estudo, solicito a sua especial colaboração em permitir que os dados resultantes da sua participação sejam estudados. Como forma de complementar e auxiliar o melhor entendimento da realidade estudada, este estudo poderá ser registrado por meio de anotações, imagens como printscreen e gravação de áudio/vídeo. Nenhum tipo de informação pessoal que o identifique será divulgada.
- Para a participação na pesquisa, você precisará de um computador conectado à Internet, e e-mail para eventuais transferências de arquivos. Isso será necessário para interação com outros participantes e recursos necessários ao estudo, assim como para preenchimento de questionário, o que levará aproximadamente 90 minutos.
- Houve planejamento para evitar constrangimentos e/ou desconfortos por parte dos participantes, mas caso isso venha a ocorrer, você pode interromper a sua participação a qualquer momento, sem qualquer tipo de prejuízo.
- Serão tomadas as providências durante a coleta de dados de forma a garantir sua privacidade e seu anonimato. Os dados coletados durante o estudo destinam-se estritamente às atividades de pesquisa relacionadas à investigação de aplicações no processo de design, não sendo utilizados em qualquer forma de avaliação profissional ou pessoal.
- Os benefícios esperados com essa pesquisa envolvem a identificação de limitações no entendimento do fenômeno da desinformação digital com relação a complexidade de aspectos sociotécnicos no contexto onde a desinformação ocorre e sua relevância para o desenho de soluções de mitigação.
- O pesquisador Alisson Andrey Puska, responsável por este estudo, poderá ser localizado pelo e-mail [alisson.puska@gmail.com](mailto:alisson.puska@gmail.com) e no telefone (46) 99136-0880 para esclarecer eventuais dúvidas.
- A sua participação neste estudo é voluntária e se o(a) senhor(a) não quiser mais fazer parte da pesquisa poderá desistir a qualquer momento e solicitar que desconsiderem este Termo de Consentimento Livre e Esclarecido assinado.
- As informações relacionadas ao estudo poderão ser conhecidas por pessoas autorizadas, como o Prof. Dr. Roberto Pereira, orientador desta pesquisa. No

entanto, se qualquer informação for divulgada em relatório ou publicação, isto será feito sob forma codificada, para que a sua identidade seja preservada e mantida sua confidencialidade.

- O material obtido - questionários, áudios e vídeos – será utilizado unicamente para essa pesquisa e será destruído/descartado após transcrição em texto anonimizado ao término do estudo, dentro de 1 semana. As imagens capturadas serão armazenadas, de modo anonimizado, com o propósito de publicação em artigos científicos.
- Não haverá qualquer tipo de remuneração pela sua participação.

Eu, \_\_\_\_\_ li esse Termo de Consentimento e compreendi a natureza e objetivo do estudo do qual concordei em participar. A explicação que recebi menciona os riscos e benefícios. Eu entendi que sou livre para interromper minha participação a qualquer momento sem justificar minha decisão e sem qualquer prejuízo para mim. Eu concordo voluntariamente em participar deste estudo.

## A.11 APÊNDICE 11 - PLANO DE AVALIAÇÃO DA DISMISS

Esta seção apresenta o plano de uma oficina realizada com o objetivo de avaliar a versão simplificada da Dismiss.

## Plano de uma oficina de análise de desinformações para avaliação da Dismiss

Alisson A. Puska e Roberto Pereira<sup>1</sup>

<sup>1</sup>Programa de Pós-Graduação em Informática  
Departamento de Informática – Universidade Federal do Paraná (UFPR)  
Curitiba – PR – Brazil

[\[alissonpuska,rpereira.inf\]@gmail.com](mailto:[alissonpuska,rpereira.inf]@gmail.com)

**Resumo.** O fenômeno da desinformação digital é um problema complexo que envolve tanto a tecnologia quanto a sociedade. Nesse sentido, as soluções sendo propostas para mitigar os problemas da desinformação digital precisam lidar com a complexidade dessa natureza sociotécnica característica do fenômeno (Fernandez e Alani, 2018). Entretanto, no melhor do nosso conhecimento, as ferramentas e artefatos conceituais propostos para apoiar a análise abrangente do fenômeno tendem a desconsiderar aspectos do contexto social integrados aos aspectos técnicos (Starbird et al., 2019; Fernandez e Alani, 2018). Sem uma visão sociotécnica abrangente, que entenda como a tecnologia e a sociedade interagem para produzir e espalhar desinformação, as soluções ficam limitadas, restringidas pelo entendimento insuficiente de problemas relacionados ao fenômeno. Considerando essas limitações desenvolvemos a Dismiss, uma abordagem para apoiar o diagnóstico de casos de desinformação de modo sistêmico e sociotécnico por pesquisadores e projetistas de soluções computacionais interessados no fenômeno. Assim, propomos esta oficina para avaliar a utilidade percebida e a facilidade de uso percebida da Dismiss com o público-alvo.

### 1. Introdução

O objetivo desta oficina é avaliar a Aceitação de Tecnologia (TAM) da Dismiss - uma abordagem para a análise sociotécnica da desinformação digital. A Dismiss foi desenvolvida em três níveis de complexidade construtivos. O nível *Light-Weight* engloba as técnicas e artefatos básicos para conduzir uma análise sociotécnica de modo simples, porém abrangente. O nível *Custom* que adiciona materiais de apoio para aprofundar a análise e identificar aspectos sociotécnicos relevantes com mais rigor. Por fim, o nível *Expert* que integra técnicas de exploração e identificação de padrões para permitir uma análise abrangente e aprofundada, capaz de revelar as dimensões de cada caso.

Esta oficina é proposta para avaliar a versão *Light-Weight*, tendo em vista que as demais versões foram avaliadas em momentos anteriores. A oficina será realizada em três fases. Na primeira fase, os participantes receberão uma visão geral da Dismiss e suas principais características. Na segunda fase, os participantes serão solicitados a auxiliar o facilitador (autor) na aplicação da Dismiss para analisar um caso real de desinformação. Na terceira fase, os participantes serão convidados a preencher um questionário para avaliar a utilidade percebida e a percepção de facilidade de uso da Dismiss.



**Grupo de participantes:**

Pesquisadores e analistas que trabalham ou têm interesse no campo da análise de desinformação digital.

**Fase 1: Introdução à Dismiss (20 minutos)**

**Boas-vindas e introdução:** O facilitador (autor) dará as boas-vindas aos participantes e apresentará o objetivo da oficina. O facilitador irá se apresentar e explicar brevemente a sua experiência no campo da desinformação e apresentará o TLCE.

**Visão geral da Dismiss:** O facilitador fornecerá uma visão geral da Dismiss, incluindo seus principais recursos, objetivos e benefícios.

**Sessão de perguntas e respostas:** Os participantes terão a oportunidade de fazer perguntas e esclarecer quaisquer dúvidas que possam ter sobre a Dismiss. O facilitador encorajará os participantes a fazer perguntas e fornecer respostas claras e concisas.

**Fase 2: O estudo de caso (60 minutos)**

**Introdução ao estudo de caso:** O facilitador deve apresentar um caso real de desinformação para os participantes analisarem. O facilitador guiará os participantes para ilustrar como a Dismiss pode ser aplicada a situações do mundo real. Os participantes apoiarão a aplicação da Dismiss, auxiliando na coleta de dados e preenchimento dos artefatos da abordagem.

**Materiais:**

1. A Dismiss e suas técnicas e artefatos propostos para análise;
  - a. Inclui o Ciclo de Vida da Desinformação digital, Análise de Enquadramento, Diagrama de Partes Interessadas, Escada Semiótica e os Formulários para anotação dos dados.
2. Um conjunto de tarefas projetadas para testar a utilidade percebida da Dismiss.

**Procedimento:**

1. Os participantes receberão um conjunto de tarefas a serem executadas usando a Dismiss. O facilitador guiará a utilização da ferramenta, observará os participantes enquanto concluem as tarefas e anotará suas interações com a Dismiss. Tarefas:
  - a. Explorar o contexto do caso para identificar os elementos básicos para análise (objetos desinformativos, canais de comunicação e partes interessadas) e preencher os artefatos da Dismiss;
  - b. Análise sociotécnica do caso para identificar aspectos sociotécnicos que caracterizam o caso e preencher os artefatos da Dismiss;
  - c. Análise dos dados coletados para identificar e entender problemas no caso e preencher os artefatos da Dismiss;

### Fase 3: Avaliação de Aceitação de Tecnologia (40 min)

**Introdução à avaliação:** o facilitador fornecerá uma breve visão geral do Modelo de Aceitação de Tecnologia (TAM), incluindo seus principais componentes.

**Distribuição do Questionário:** os participantes receberam um questionário que foi elaborado para avaliar o TAM da Dismiss.

**Pesquisa:** os participantes serão solicitados a responder ao questionário que mede suas atitudes em relação à Dismiss baseada no Modelo de Aceitação de Tecnologia (TAM). As perguntas da pesquisa se concentrarão nos seguintes pontos:

1. Utilidade percebida: em uma escala de 1 a 5, o quanto você concorda com as seguintes afirmações sobre a Dismiss?
  - a. A Dismiss me ajudou a entender a natureza complexa do caso de desinformação.
  - b. A Dismiss foi útil para identificar problemas no caso de desinformação.
  - c. A Dismiss me ajudou a identificar aspectos sociotécnicos para serem considerados no desenvolvimento de soluções que ficariam ocultos sem o uso da abordagem.
  - d. Comentários adicionais:
    - i. Forneça qualquer feedback ou comentário adicional que você possa ter sobre a utilidade da Dismiss.
2. Percepção de facilidade de uso: em uma escala de 1 a 5, o quanto você concorda com as seguintes afirmações sobre a Dismiss?
  - a. A Dismiss é fácil de usar.
  - b. A Dismiss é fácil de entender.
  - c. Considero fácil de integrar a Dismiss em minhas atividades de trabalho.
  - d. Comentários adicionais:
    - i. Forneça qualquer feedback ou comentário adicional que você possa ter sobre a facilidade de uso da Dismiss.
3. Intenção de uso: em uma escala de 1 a 5, o quanto você concorda com as seguintes afirmações sobre sua intenção de usar a Dismiss?
  - a. Acredito que a Dismiss será útil em meu trabalho.
  - b. Acredito que usar a Dismiss aumentará a qualidade de soluções geradas.
  - c. Comentários adicionais:
    - i. Forneça qualquer feedback ou comentário adicional que você possa ter sobre a sua intenção de uso da Dismiss.
4. Comentários adicionais:
  - a. Forneça qualquer feedback ou comentário adicional que você possa ter sobre a Dismiss.

**Preenchimento do Questionário:** Os participantes terão 20 minutos para completar o questionário. O facilitador irá instruir sobre como preencher o questionário e responder a quaisquer perguntas que os participantes possam ter.

**Grupo Focal:** Após o preenchimento do questionário, o facilitador conduzirá uma discussão em grupo (20 minutos) sobre as experiências dos participantes quanto ao uso da Dismiss. A discussão se concentrará na identificação dos pontos fortes e fracos da Dismiss, bem como nas áreas de melhoria.

**Análise de dados:** As respostas da pesquisa feita durante a avaliação, os artefatos preenchidos produzidos na atividade, e as anotações relacionadas a observação dos participantes durante a atividade serão analisadas para identificar temas relacionados às questões do TAM.

**Conclusão:** O facilitador irá resumir as principais conclusões do workshop e fornecer algumas recomendações sobre como a Dismiss pode ser melhorada. O facilitador irá agradecer aos participantes por sua participação e incentivá-los a fornecer feedback futuros.

## A.12 APÊNDICE 13 - MANUAL DA DISMISS

Esta seção apresenta o Manual utilizado pelos participantes das oficinas de avaliação da Dismiss.

## DISMISS: ABORDAGEM PARA ANÁLISE SOCIOTÉCNICA e SISTÊMICA DE DESINFORMAÇÕES

Passo a passo para aplicação da DISMISS

Alisson Andrey Puska, 26-Novembro, 2022. v.0.6

O fenômeno da desinformação digital apresenta uma natureza sociotécnica complexa que representa um desafio para mitigação de seus efeitos na sociedade [Fernandez, 2018] e que pode levar ao desenvolvimento de soluções prejudiciais como a censura [Puska, 2021].

A DISMISS é uma abordagem proposta para apoiar a análise sociotécnica de desinformação digital, instrumentalizando o diagnóstico de casos de desinformação e auxiliando na caracterização do contexto, no entendimento de problemas e na identificação de requisitos para a ideação e o desenvolvimento de soluções.

A DISMISS pode auxiliar a:

- Identificar o tipo de desinformação (acidental ou deliberada);
- Identificar as partes interessadas em um caso de desinformação;
- Entender como as partes interessadas são afetadas;
- Caracterizar o Ciclo de Vida da Desinformação Digital;
- Identificar problemas em um caso de desinformação;
- Identificar indícios de uma campanha de desinformação deliberada;
- Entre outros;

Este material apresenta os artefatos e técnicas da DISMISS para realizar a análise sociotécnica de casos de desinformação e exemplifica como utilizá-los.

## Analise Sociotécnica de Desinformação Digital

A desinformação digital existe a partir da comunicação de pessoas através de Tecnologias de Informação e Comunicação (TIC), englobando uma diversidade de aspectos sociais e técnicos que influenciam como um caso de desinformação se desenrola e na extensão de suas consequências.

Para lidar com a complexidade de elementos sociotécnicos do fenômeno, a DISMISS estrutura a análise de um caso de desinformação e utiliza artefatos adaptados da Semiótica Organizacional (SO) para auxiliar no entendimento da natureza sociotécnica da desinformação digital.

A DISMISS é dividida em 4 etapas que abordam diferentes aspectos da análise (Figura 3).

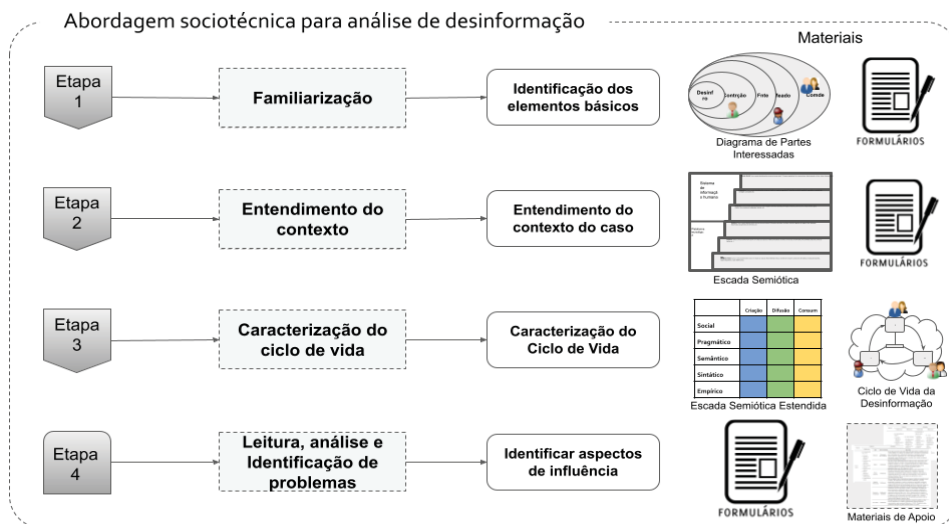


Figura 3 - Etapas da DISMISS.

- Na Etapa 1, o analista explora o contexto do caso de desinformação para identificar os elementos básicos que fundamentam a análise. Nessa etapa o analista identificar os meios de comunicação, as partes interessadas e os objetos digitais de desinformação.
- A Etapa 2 engloba a análise sociotécnica do caso de desinformação, visando caracterizar o Ciclo de Vida da Desinformação para compreender como os elementos e aspectos sociotécnicos podem estar relacionados e ampliar o entendimento do contexto. Nessa etapa o analista revela aspectos sociotécnicos inter-relacionados na criação, difusão e consumo da desinformação no caso analisado.
- A Etapa 3 visa identificar problemas no caso analisado através da identificação de aspectos de influência na criação, difusão e consumo de desinformação. O analista revê os aspectos no ciclo considerando os Princípios de Influência e os Aspectos da tecnologia.



Para explicar o uso da DISMISS, esse tutorial apresenta um exemplo real (Figura 5) de análise de um caso de desinformação.

### Caso: Farsa no Zap

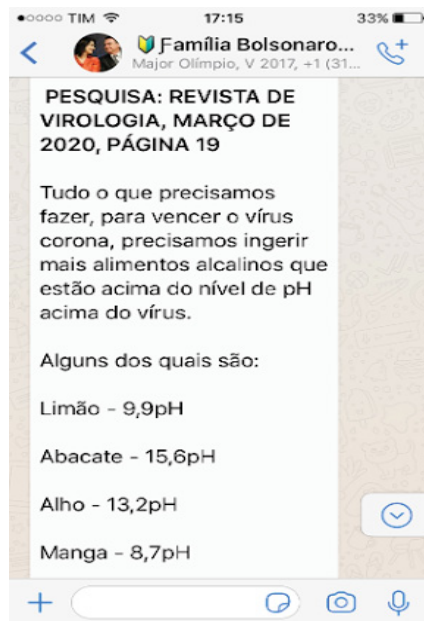


Figura 5 - Objeto desinformativo do caso em análise.

Em março de 2020 circulou pelo WhatsApp no Brasil uma mensagem que relacionava o combate do SARS-CoV-2 com o nível de Ph do corpo humano<sup>1 2</sup>. A mensagem indicava um tratamento onde o indivíduo deveria comer certos alimentos para atingir um nível de Ph que supostamente eliminaria a SARS-CoV-2 de seu organismo. A mensagem trazia em sua estrutura referência a autoridades científicas, informações sobre alimentos com pH errado, além da orientação de que é possível alterar o Ph do corpo<sup>3</sup>. Ainda, essa desinformação foi corrigida no final de março por agências brasileiras de checagem de fatos (Lupa e G1)<sup>1 2</sup>. Apesar das correções, o objeto desinformativo continuou circulando.

### Fontes

1-<https://piaui.folha.uol.com.br/lupa/2020/04/16/verificamos-ph-alimentos-covid>

2-<https://g1.globo.com/fato-ou-fake/noticia/2020/03/30/e-fake-que-a-ingestao-de-alimentos-alcalinos-combate-o-novo-coronavirus.ghtml>

3-<https://healthfeedback.org/claimreview/eating-alkaline-food-cannot-prevent-or-cure-covid-19>

### Etapa 1 - Familiarização

Na primeira etapa, o analista explora o contexto do caso de desinformação para se familiarizar. Isso significa entender os elementos básicos do caso, empatizar com as partes interessadas, e identificar objetos desinformativos.

O analista irá identificar as partes interessadas, os canais de comunicação, o objeto desinformativo e suas transformações (se houverem), além de definir um escopo para análise.

#### Exploração do contexto

Como primeiro passo, o analista deve coletar informações básicas sobre os elementos que constituem um caso de desinformação para análise.

Os elementos básicos de um caso de desinformação digital são: partes interessadas, canais de comunicação e objeto desinformativos. As partes interessadas são os participantes do caso. Os canais de comunicação são os meios por onde a desinformação se propaga. Os objetos desinformativos são a desinformação no formato digital.

Para guiar o analista nessa etapa, a DISMISS utiliza o [Formulário de Registro de Caso Desinformativo \(FRCD\)](#). Os dados da coleta informam o analista para definir o escopo da análise e dar início a caracterização do caso.

| Formulário de Registro de Evento |                        |  |  |
|----------------------------------|------------------------|--|--|
| Sobre o evento                   |                        |  |  |
| Título do evento em análise:     | Temática do evento:    |  |  |
| Quantidade de objetos:           | Miniatura do objeto:   |  |  |
| Data de início do evento:        | Data de fim do evento: |  |  |
| Descrição do evento:             |                        |  |  |

Figura 5 - Formulário de Registro de Casos - (Etapa 1).

**Dica:** Todos os campos deste e de outros formulários da DISMISS possuem descrições e exemplos para auxiliar o preenchimento, bastando posicionar o mouse sobre o campo para exibir a ajuda.

### Coleta de informações

Primeiramente, o analista deve abordar o caso a partir do objeto desinformativo. Para coletar as informações, o analista deve observar o meio de comunicação por onde a desinformação se propaga, buscando as informações que respondam as seguinte perguntas-guia:

- Qual o conteúdo da desinformação?
  - A temática é um dos atributos usados para estudar casos com mais de um objeto desinformativo ou rastrear as transformações de um objeto em um caso ao longo do tempo. É importante identificar a temática pois ela está intimamente ligada a aspectos sociais do caso de desinformação, como a popularidade e relevância do assunto atraem a atenção e se difundem mais rápido e com maior alcance [Feng 2013]. O FRED apresenta uma lista de temas para facilitar a identificação.
- Qual o tipo do objeto desinformativo?
  - Existem vários tipos de objetos desinformativos e cada tipo possui características particulares. Cada tipo de objeto está ligado a um propósito de uso e possui aspectos sociotécnicos que induzem um certo tipo de interação. Por exemplo, iscas-clique são um tipo de objeto desinformativo que busca monetizar acessos, atraindo cliques de usuários. Para isso, esse tipo de objeto apela para imagens, títulos e subtítulos sensacionalistas e com temas populares, além de serem frequentemente difundidos na forma de anúncios (propaganda, notícias, assuntos populares, etc.). Para apoiar o analista a identificar o tipo de desinformação a DISMISS oferece o [Quadro de Tipo de Desinformação](#) com descrições e exemplos.
- Onde o objeto aparece?
  - Esta informação está relacionada com a definição do escopo da análise. Um caso de desinformação pode ser fruto de uma operação orquestrada, com vários tipos de objetos desinformativos em torno de uma mesma temática que aparecem em vários canais de comunicação [Starbird 2020], ou então pode ser uma ocorrência oportunista/acidental, localizada em um único canal de comunicação [Qin 2018]. Além disso, um objeto desinformativo pode se propagar para diferentes canais de comunicação durante o seu ciclo de vida. Isso é importante para compreender o estágio de difusão do ciclo de vida e rastrear um objeto desinformativo. Dessa forma, é preciso identificar os elementos que constituem os canais de difusão, como os **tipos de dispositivo**, as **tecnologias de comunicação**, **aplicativos**, as **plataformas digitais**, etc.
- Qual o escopo temporal do caso?
  - O início do caso. É importante situar a análise no tempo para evitar viés na identificação de aspectos sociotécnicos. Quando se considera as partes interessadas, o aspecto "tempo" é fundamental para construir um entendimento abrangente dos problemas relacionados à uma determinada desinformação digital e identificar aspectos relevantes em cada caso de desinformação analisado. Isso é importante pois aspectos sociotécnicos como crenças, valores, autoridades, consensos, e mesmo fatores tecnológicos como os mecanismos de interação/mitigação em uma determinada TIC tem

efeitos válidos em um tempo específico, para partes interessadas específicas e em um espaço específico.

#### **Como coletar informações**

A coleta de informações deve ser situada, ou seja, a partir do canal no qual o objeto desinformativo surge, considerando o escopo temporal do caso e o contexto da comunicação. Nesse sentido, a DISMISS sugere duas abordagens para coleta de informações: em profundidade ou em largura.

No primeiro caso, o analista identifica um objeto e inicia a exploração do contexto onde ele ocorreu, buscando por elementos que respondam às perguntas-guia. A DISMISS sugere iniciar a coleta a partir do objeto, identificando o meio de comunicação imediatamente ligado ao objeto, as partes interessadas, as informações relacionadas à temática e o contexto da comunicação neste meio.

De outro modo, o analista pode buscar por relatos do aparecimento de objetos com um mesmo assunto em diferentes meios de comunicação em um mesmo período de tempo.

Por exemplo, matérias em portais informativos, checadores de fatos e postagens em redes sociais podem revelar como as partes interessadas lidam ou lidaram com a desinformação, e como partes interessadas que não estão diretamente envolvidas podem ser impactadas.

**Dica:** Busque as imagens dos objetos desinformativos no Google ou por partes do conteúdo para encontrar fontes de dados para a coleta. Sempre utilize fontes confiáveis, de reputação reconhecida, auditáveis e que tenham compromisso com a qualidade do conteúdo que fornecem.

Além disso, estas fontes podem conter detalhes sobre o caso, como quando o objeto surgiu, por quais plataformas ele se difundiu, as estratégias de mitigação, e outras partes interessadas.

#### **Resultados esperados**

No final desta etapa o analista terá identificado e registrado os meios de comunicação, os objetos desinformativos, a temática e o escopo temporal do caso. A Figura 6 apresenta o FRCD preenchido para o caso analisado.

| Formulário de Registro de Evento           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|--------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sobre o evento                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Título do caso em análise:<br>Farsa no Zap | Temática do evento:<br>Saúde                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Canais de comunicação:<br>- WhatsApp;                                                                                                                                                                                                                                                                                                                                                                                                             |
| Quantidade de objetos:<br>1                | Descrição do evento:<br>Em março de 2020 circulou pelo WhatsApp no Brasil uma mensagem que relacionava o combate do SARS-CoV-2 com o nível de Ph do corpo humano <sup>1 2</sup> . A mensagem apresentava um tratamento onde o indivíduo deveria comer certos alimentos para atingir um nível de Ph que supostamente aniquila a SARS-CoV-2. A mensagem trazia em sua estrutura referência a autoridades científicas, informações sobre alimentos com pH errado, além da orientação de que é possível alterar o Ph do corpo <sup>3</sup> . Ainda, essa desinformação foi corrigida no final de março por agências brasileiras de checagem de (Lupa e G1) <sup>4</sup> e <sup>2</sup> , apesar das correções, o objeto desinformativo continuou circulando. |                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Data de início do caso:<br>Março de 2021   | Data de fim do evento:<br>Julho 2021                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Fonte de informações:<br>1- <a href="https://piaui.folha.uol.com.br/lupa/2020/04/16/verificamos">https://piaui.folha.uol.com.br/lupa/2020/04/16/verificamos</a><br>2- <a href="https://g1.globo.com/fato-ou-fake/noticia/2020/03/30/e-fa">https://g1.globo.com/fato-ou-fake/noticia/2020/03/30/e-fa</a><br>3- <a href="https://healthfeedback.org/claimreview/eating-alkaline-foc">https://healthfeedback.org/claimreview/eating-alkaline-foc</a> |

Figura 6 - FRCD preenchido .

### Identificando as partes interessadas

Para identificar as partes interessadas, o analista utiliza o [Formulário com o Diagrama de Partes Interessadas](#). O formulário é baseado no artefato adaptado da SO - Diagrama de Identificação de Partes Interessadas.

A Figura 7 apresenta o [Formulário com o Diagrama de Partes Interessadas](#) preenchido para o caso do exemplo analisado.

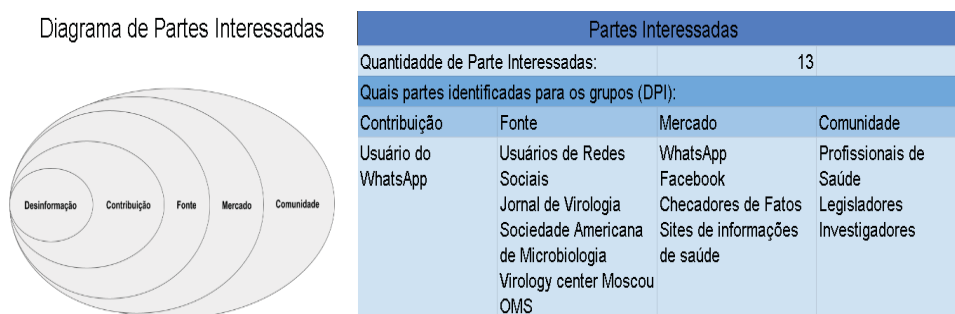


Figura 7 - Formulário de identificação de partes interessadas da DISMISS.

O analista deve observar cada camada do DPI e buscar identificar ao menos uma parte interessada em cada grupo enquanto coleta informações.

- **Contribuição:** Essa categoria inclui as partes interessadas que atuam na concepção e divulgação do objeto desinformativo, como quem cria, financia, planeja, prepara e disponibiliza.
- **Fonte:** Nessa categoria estão aqueles que interagem com a desinformação fornecendo dados para a sua concepção, disseminação e consumo, como a audiência pretendida ou a audiência casual (observador que não do público-alvo).
- **Mercado:** Esta categoria associa as partes interessadas parceiras e concorrentes da desinformação no caso como redes sociais, provedores de Internet, e checadores de fatos.

- **Comunidade:** A categoria comunidade engloba as partes interessadas que são impactadas indiretamente pela desinformação como governos ou legisladores.

As seguintes questões podem auxiliar na identificação de partes interessadas:

- Quais partes interessadas participam do caso?
  - Quem é o interlocutor, quem cria, quem difunde, quem consome, quem controla os meios de difusão, quem sofre as consequências, etc.
- Quais atividades/papéis são exercidos pelas partes interessadas?
  - Identificar autoridades, celebridades, posições de destaque, fiscais, legisladores, referências, pessoas citadas, etc.
- Quais os possíveis relacionamentos entre as partes interessadas?
  - Agrupamentos de participantes, laços familiares, amigos, relações profissionais, etc.

### Resultados Esperados

Ao final desta etapa o analista terá delimitado o escopo da análise e identificado os elementos essenciais para compreensão da natureza sociotécnica do caso. Ele terá uma relação das principais partes interessadas, o objeto desinformativo, onde o objeto aparece, qual a temática e definido o escopo temporal da análise.

Os dados coletados estarão registrados no [Formulários de Registro de Casos de Desinformação](#) e no [Formulário de Identificação de Partes Interessadas](#).

## Etapa 2 - Análise Semiótica do Caso

Nesta etapa, o analista identifica aspectos sociotécnicos relevantes no caso em análise. Estes aspectos devem ser identificados considerando cada um dos elementos da etapa anterior: canais de comunicação, partes interessadas e objeto desinformativo.

Ao final desta etapa o analista terá um conjunto de aspectos sociotécnicos em diferentes níveis de abstração do caso de desinformação, caracterizando o contexto do caso. Estes dados seguem para a Etapa 3 para caracterização dos estágios do ciclo de vida da desinformação.

### Identificando aspectos sociotécnicos

Nesta etapa o analista utiliza a Escada Semiótica (Figura 8), questões-guia e exemplos para identificar aspectos em diferentes níveis de abstração. Os aspectos identificados são registrados no Formulário de Identificação de Aspectos Sociotécnicos (FIAS) - Figura 8.

Para apoiar a análise dos diferentes níveis de abstração, o FIAS emprega a Escada Semiótica Estendida (ESE) - um artefato adaptado da SO usado para análise semiótica de Sistemas de Informação.





Figura 8 - Escada Semiótica.

A ESE agrupa os aspectos sociotécnicos em dois níveis de abstração:

- Aspectos do nível de Sistema de Informação Humano (SIH) - relacionados ao modo como as pessoas aprendem, significam, criam e manipulam informações.
- Aspectos do nível da Plataforma Tecnológica (PT) - relacionados a como o objeto digital é criado, armazenado, difundido e manipulado em uma TIC.

O nível de abstração do SIH é dividido em três subníveis: Social, Pragmático e Semântico.

- **Social:** Neste subnível são registrados os aspectos sociais que impactam e são impactados pela desinformação.
  - Exemplos de aspectos relevantes são os relacionamentos, proximidade afetiva, hierarquia, figuras de autoridade, deveres, direitos, popularidade de uma parte interessada ou tema, entre outros;
- **Pragmática:** Neste subnível são registrados os aspectos que impactam ou são impactados pelo propósito e sentido da desinformação.
  - Exemplos de aspectos relevantes são a conotação da mensagem, se é um pedido, uma afirmação, uma crítica, uma ordem; Declarações e parágrafos finais que buscam concluir ideias; Personificação, deturpação, entre outros;
- **Semântica:** Neste subnível são registrados os aspectos que impactam e são impactados pelo significado e entendimento.
  - Exemplos de aspectos relevantes são a escolha de palavras, dialeto, metáforas, palavras-chave, frases de efeito, jargões, citações, fontes de informação, imagens, expressões, entre outros.

A ES divide o nível de abstração da PT em três subníveis: Sintático, Empírico e Físico.

- **Sintático:** Neste subnível são registrados os aspectos das estruturas e padrões da desinformação.
  - Exemplos de aspectos relevantes são os elementos de ênfase como manchetes, subtítulos, legendas de fotos, posicionamento de figuras,

diagramação, destaque da mensagem, posicionamento da página (primeira página, etc.), entre outros;

- **Empírica:** Neste subnível são registrados os aspectos relacionados às capacidades da plataforma tecnológica, com base na sincronicidade, direcionalidade, e controle.
  - Exemplos de aspectos relevantes são os tipos de interação, sincronicidade da comunicação, direcionalidade do canal, privacidade, etc;
- **Física:** Neste subnível são registrados os aspectos relacionados ao mundo físico;
  - Exemplos de aspectos relevantes são os artefatos tecnológicos de comunicação, defeitos, localização física das partes interessadas, número de dispositivos, custo, entre outros.

A Análise deve ser feita da seguinte forma:

- Adotar a perspectiva das partes interessadas, considerando os elementos do caso de desinformação identificados na Etapa 1 da DISMISS;
- O analista deve utilizar os exemplos na ES para identificar aspectos sociotécnicos em vários níveis;
  - O FAIS representa a escada semiótica e apresenta um conjunto de questões-guia para apoiar a análise sociotécnica;

### Resultados esperados

Como resultado, esta etapa reúne os principais aspectos sociotécnicos identificados e registrados pelo analista no Formulário de Registro de Aspectos Sociotécnicos (FRAS) - Figura 9.

| Formulário de Registro de Aspectos Sociotécnicos |                                                                                                                                                                                                                                                                                                                                                              |
|--------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Social                                           | - Agrupado em comunidades de interesse mútuo e/ou pessoas conhecidas/confiáveis;<br>- Compartilhamento em grupos de amigos e conhecidos/ grupos públicos;<br>- Difícil identificar a fonte inicial da desinformação;- Brasileiro usa whatsapp para discutir notícias e tópicos de interesse;<br>- Pandemia: população precisa se informar sobre tratamentos; |
| Pragmático                                       | -Intencionalidade: enviesamento do conteúdo do artigo é indício de deliberação;- App tem credibilidade;<br>- Pouco interesse em implementar políticas para mitigação;<br>-Desinformação percebida como verídica<br>- Viés de ancoragem podem facilitar o consumo,                                                                                            |
| Semântico                                        | - Linguagem pessoal, positiva e emocional;<br>- Falácia lógica: apelo a autoridade;<br>- Referência a possível autoridade;- Linguagem acessível e termos apelativos;<br>- Tom pessoal da mensagem;<br>- Tema de interesse;                                                                                                                                   |
| Sintático                                        | - Meta dados como a data de criação e compartilhamento, e a mensagem encaminhada;<br>-Origem desconhecida;<br>-Carece de mecanismos de mitigação;                                                                                                                                                                                                            |

|          |                                                                                                                                                                                                                                                                                    |
|----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|          | <ul style="list-style-type: none"> <li>- Planos de acesso restritivos;</li> <li>- Sem links externos, imagens,</li> <li>- Título, listagem;</li> </ul>                                                                                                                             |
| Empírico | <ul style="list-style-type: none"> <li>- Acesso a internet e smartphones com WhatsApp no Brasil;</li> <li>- Meio assíncrono bidirecional</li> <li>Interação muitos para muitos;- Permite discussão e deliberação</li> <li>- Mídia de comunicação muito usada no Brasil;</li> </ul> |
| Físico   | <ul style="list-style-type: none"> <li>- Quantidade brasileiros com Smartphones;</li> <li>- Usuários em diferentes locais do Brasil;</li> <li>- Distância física;</li> </ul>                                                                                                       |

Figura 9 - Exemplo de FRAS preenchido.

### Etapa 3 - Caracterização do Ciclo de Vida

Nesta etapa, o analista observa os aspectos sociotécnicos que contextualizam o caso, buscando instanciar-los no Ciclo de Vida da Desinformação. Estes levantamento deve ser feito considerando cada um dos elementos identificados na etapa anterior: canais de comunicação, partes interessadas e objeto desinformativo.

Ao final desta etapa o analista terá um conjunto de partes interessadas e aspectos sociotécnicos relevantes em cada estágio do ciclo de vida. Estes dados seguem para a Etapa 4 de leitura, análise e identificação de problemas.

Como uma forma de simplificar o estudo sociotécnico de casos de desinformação, a DISMISS considera quatro estágios do Ciclo de Vida um objeto desinformativo:

1. **Criação:** Este estágio foca no surgimento, quando o objeto é construído e divulgado pela primeira vez.
2. **Difusão:** Este estágio foca nos aspectos da propagação do objeto desinformativo, como as capacidades da mídia de comunicação e o comportamento de difusão;
3. **Consumo:** Neste estágio o foco está na interação da audiência com o objeto desinformativo, relacionado a ações como, ler, compartilhar, ignorar e reagir;
4. **Reiteração:** Este estágio foca no reaparecimento de um objeto desinformativo e das transformações que ocorrem durante sua vida.

Estratégias de **Mitigação** podem ser aplicadas em qualquer estágio do ciclo.

O Ciclo (Figura 3) engloba todos os elementos de um caso de desinformação: partes interessadas, objetos desinformativos, e aspectos sociotécnicos (Contexto).

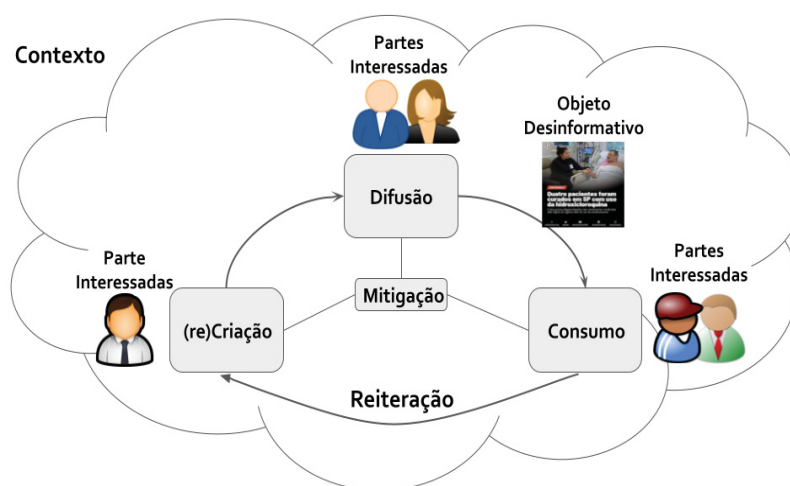


Figura 3 - Representação do ciclo de vida da desinformação.

#### Exemplo de Análise do estágio de Criação

A análise semiótica do estágio de criação considera a perspectiva das partes interessadas envolvidas na concepção e divulgação inicial do objeto desinformativo.

O analista deve adotar a perspectiva da(s) parte(s) interessada(s) envolvidas na criação para observar os diferentes aspectos sociotécnicos com ajuda do ASE. A Figura 11 apresenta um exemplo de análise semiótica do estágio de criação.



Figura 11 - Exemplo de análise do estágio de criação.

#### Quais aspectos sociotécnicos estão relacionados com os Princípios de Influência?

Por exemplo, a relação do conteúdo com a ideologia de grupo, frases que criam o senso de urgência, ou induzem um senso de pertencimento, autoridades como responsáveis pela informação, a autoridade de quem compartilhou a mensagem no grupo, etc;

#### Quais aspectos sociotécnicos do objeto estão relacionados com a intenção do criador?

Por exemplo, a introdução de vies proposital, a escolha de palavras ambíguas, a escolha do público-alvo e de uma temática específica, se é possível identificar uma tendência de posicionamento ideológico, apelo a valores, etc.

#### Quais aspectos sociotécnicos estão relacionados a Plataforma Tecnológica?

Por exemplo, elementos de ênfase que guiam a atenção, como o tamanho da fonte, o uso de negrito, listas e espaços demasiados; erros sintáticos; outro elementos visuais como emojis, gifs e figuras; a formatação permitida pela plataforma, mecanismos que auxiliam ou prejudicam a criação da mensagem, etc;

#### Exemplo de Análise do estágio de Difusão

A análise semiótica da difusão visa identificar aspectos sociotécnicos relacionados a propagação do objeto desinformativo. A análise foca nos recursos da mídia de comunicação e na perspectiva das partes interessadas envolvidas na difusão. A Figura 12 apresenta um exemplo de análise semiótica da difusão.



Figura 12 - Exemplo de análise do estágio de difusão.

#### Quais aspectos sociotécnicos estão relacionados com os Princípios de Influência?

Por exemplo, a motivação de informar outras pessoas, a popularidade do tema, a relação do meio de comunicação com os usuários (credibilidade e confiabilidade), se a fonte da mensagem é mesma parte interessada que criou o conteúdo, etc;

#### Quais aspectos sociotécnicos estão relacionados com a Plataforma Tecnológica?

Por exemplo, restrições para transmissão de dados (como planos de acesso a dados dos prestadores de serviço, dos portais informativos, etc.), aspectos da direcionalidade e da sincronidade que impactam na difusão, como um meio assíncrono que facilita o armazenamento do objeto desinformativo, aspectos de controle de acesso como se a mensagem foi re-encaminhada, etc.

#### Quais aspectos sociotécnicos estão relacionados com a intenção das partes interessadas da difusão?

Por exemplo, as políticas de uso da aplicação e dos serviços prestados, se a legislação da LGP está sendo cumprida, se existem mecanismos ativos de mitigação da desinformação, se existem mecanismos de aviso ou de alerta para o usuário, etc.

#### Exemplo de Análise do estágio de Consumo

A análise semiótica do consumo visa identificar aspectos sociotécnicos relacionados às as partes interessadas que consomem a desinformação. A Figura 13 apresenta um exemplo de análise do estágio de consumo.



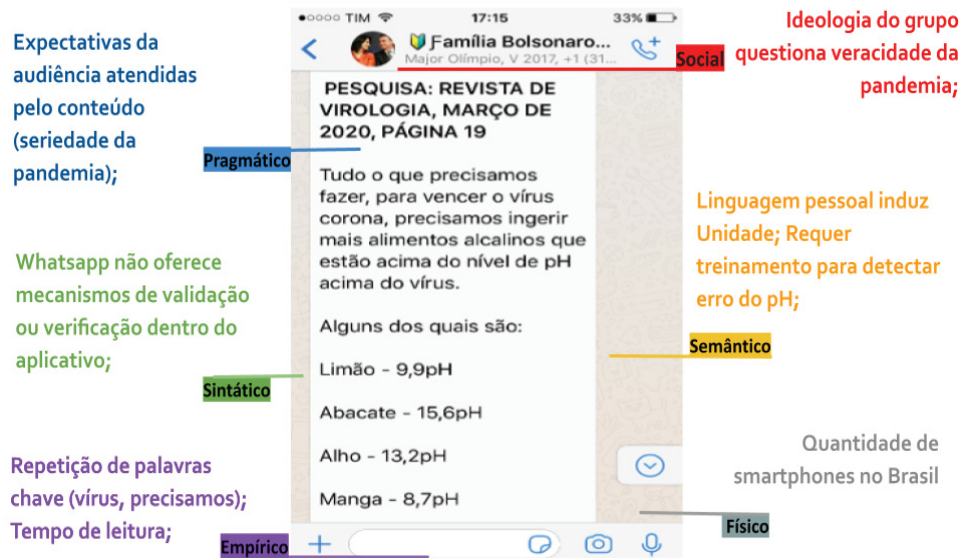


Figura 13 - Exemplo de análise do estágio de consumo.

#### Quais aspectos sociotécnicos estão relacionados aos Princípios de Influência?

Por exemplo, a relação do conteúdo com o ambiente da audiência, palavras que marcam uma ideia, palavras chaves repetidas, se o tema é popular, os relacionamentos afetivos no grupo, elementos que representam autoridades, etc.

#### Quais aspectos sociotécnicos estão relacionados com a Plataforma Tecnológica?

Por exemplo, mecanismos de controle que permitem/facilitam a verificação de informações, emojis/imagens/palavras que induzem sentimento/sensações, etc;

#### Quais aspectos sociotécnicos estão relacionados com o posicionamento das partes interessadas?

Por exemplo, se é possível identificar uma tendência de posicionamento ideológico, se ocorreram discussões, compartilhamentos, alterações, etc.

#### Resultados esperados

Como resultado, esta etapa reúne os principais aspectos sociotécnicos identificados e registrados pelo FIASE. Como resolução do exemplo o Quadro 2 apresenta o ASE preenchido.

|            | Criação                                                                                                                    | Difusão                                                                                                                          | Consumo                                                                                                                                  |
|------------|----------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| Social     | - Agrupamentos em comunidades de interesse mútuo e/ou pessoas conhecidas/confiáveis podem ser alvos de interesse;          | - Compartilhamento em grupos de amigos e conhecidos/ grupos públicos;<br>- Difícil identificar a fonte inicial da desinformação; | - Brasileiro usa whatsapp para discutir notícias e tópicos de interesse;<br>- Pandemia: população precisa se informar sobre tratamentos; |
| Pragmático | Intencionalidade: enviesamento do conteúdo do artigo é indício de deliberação;                                             | - App tem credibilidade;<br>- Pouco interesse em implementar políticas para mitigação;                                           | -Desinformação percebida como verídica<br>- Viés de ancoragem podem facilitar o consumo,                                                 |
| Semântico  | - Linguagem pessoal, positiva e emocional;<br>- Falácia lógica: apelo a autoridade;<br>- Referência a possível autoridade; |                                                                                                                                  | - Linguagem acessível e termos apelativos;<br>- Tom pessoal da mensagem;<br>- Tema de interesse;                                         |
| Sintático  | - Data de criação;<br>- Mensagem encaminhada;<br>- Origem desconhecida;                                                    | Carece de mecanismos de mitigação;                                                                                               | - Planos de acesso restritivos;<br>- Sem links externos, imagens,<br>- Título, listagem;                                                 |
| Empírico   | - Acesso a internet e smartphones com WhatsApp no Brasil;                                                                  | Meio assíncrono bidirecional<br>Interação muitos para muitos;                                                                    | - Permite discussão e deliberação<br>- Mídia de comunicação muito usada no brasil;                                                       |
| Físico     | - Brasileiros com Smartphones;                                                                                             | - Usuários em diferentes locais do Brasil;<br>- Distância física;                                                                |                                                                                                                                          |

Quadro 2 - Exemplo de ASE preenchido.

#### Etapa 4 - Leitura, análise e identificação de problemas

Nesta etapa o analista busca por inter-relações entre aspectos sociotécnicos relacionados às partes interessadas e os estágios do ciclo de vida da desinformação.

O analista pode considerar os Princípios de Influência e as Propriedades do Canal de Comunicação como ponto de partida para identificação de aspectos sociotécnicos.

Para isso, com o ASE preenchido, o analista deve identificar aspectos inter-relacionados, destacando-os no ASE. Os aspectos podem estar inter-relacionados de três formas:

- **Entre dimensões semióticas:** Quais aspectos da plataforma tecnológica restringem a expressão ou interpretação? Existe relação entre aspectos da temática, do tipo de objeto desinformativo, do conteúdo (imagens, cores, etc.) com o ambiente das partes interessadas, como ideologias, posicionamentos, crenças, valores, etc? Como as fontes de informação se relacionam com o conteúdo do objeto? Existem indícios de deliberação?
- **Entre partes interessadas:** Quais interesses, expectativas, e propósito na comunicação? Como impactam a criação, difusão e consumo? Quais partes interessadas representam autoridades? Quais as relações afetivas? Quais os consensos do grupo?
- **Entre estágios do ciclo:** Quais atividades de um estágio tem impacto no outro? Como a criação e o consumo estão relacionados? Existem aspectos que indiquem deliberação entre escolha do conteúdo do objeto com as partes interessadas da audiência? Como a difusão se relaciona com o consumo? Como o objeto foi divulgado, onde ele surgiu?

### Aspectos sociotécnicos de influência e a identificação de problemas

A DISMISS considera que os aspectos sociotécnicos relevantes são aqueles que afetam o julgamento e a percepção de risco das partes interessadas, influenciando sua avaliação de veracidade [Babaei 2019], e impactando como criam e consomem desinformações, e na velocidade e alcance da difusão de objetos desinformativos [Safieddine 2017].

A DISMISS se baseia nos Princípios da Influência [Cialdini 1993] para embasar a identificação de aspectos sociotécnicos e justificar sua relevância com base no contexto e nos elementos de cada caso.

Os Princípios de Influência representam tendências (vieses) no julgamento humano e na forma como as pessoas avaliam situações cotidianas [Cialdini 1993]. Eles são frequentemente empregados na área de vendas (propagandas), campanhas de marketing, relações públicas, e esquemas de fraude, estelionato e manipulação.

Existem [Cialdini 1993] 7 Princípios da Influência:

- **Afinidade:** Pessoas tendem a gostar daquilo que as representa, que se parece com elas. Aspectos como o uso de uma linguagem significativa para uma audiência (gírias e jargões), uma aparência agradável, usar opiniões positivas ou que confirmam crenças pre-existentes, a similaridade na estrutura/diagramação de um conteúdo, o tamanho de uma mensagem, e as cores são exemplos de aspectos que impactam a percepção de risco e a avaliação de veracidade de um objeto digital.

- **Aprovação social:** Pessoas querem se sentir validadas pelo que as outras pessoas estão fazendo ao seu redor, em especial quando em dúvida. Casos ocorrendo no ambiente de um grupo têm impacto no que o grupo percebe como comportamento correto. Aspectos deste princípio são consensos, popularidade de um tema, valores morais, crenças, etc. Este princípio ganha força na presença de incerteza, induzindo a tendência de um indivíduo em agir como o grupo está agindo - o efeito manada.
- **Autoridade:** Pessoas tendem a seguir figuras de autoridade. Identifique aspectos que impactam a percepção de risco como a popularidade de indivíduos em um grupo, como celebridades e influenciadores digitais; A autoridade de pessoas públicas, pessoas que são referência nas atividades que realizam, como médicos, cientistas etc. Busque por imagens, referências nos textos, assinaturas digitais, logotipos, certificações, etc.
- **Comprometimento:** Pessoas querem que seus compromissos sejam consistentes com os seus valores. Comprometimento tem base no que as pessoas estão fazendo, e no que as outras pessoas pensam que ela está fazendo. Assim, aspectos como a coerência, a fluência, domínio de um assunto, a expectativa da audiência de ter suas necessidades atendidas, ou defender as ideologias de um determinado grupo podem gerar afinidade e comprometimento, facilitando o engano.
- **Escassez:** A tendência das pessoas de temerem perder algo é pior que a possibilidade de deixar de ganhar. Algo que se torna (ou pode se tornar em breve) escasso ou indisponível, tem mais valor do que quando era abundante - como as corridas por papel higiênico no início da pandemia [BBC 2020], a inexistência de uma vacina/conhecimento sobre o vírus, ou um determinado grupo que luta por direitos.
- **Reciprocidade:** É a expectativa inerente de quando se recebe algo com valor positivo, um sentimento de dívida surge. Explorado por mensagens ou informações não solicitadas, que geram uma ligação com a necessidade de responder e se engajar, ou atender a pedidos posteriores. É usada para construir um senso de consciência. Alguns elementos explorados são a exposição seletiva, percepção seletiva, e controle do fluxo de informações.
- **Unidade:** A tendência de buscar nos outros, traços de identidade que nos definem, e levam a formação de grupos com indivíduos ideologicamente alinhados. É um princípio que deriva da interação social e da criação de laços de confiança entre indivíduos em grupos. Aspectos tribais da sociedade como gírias, sentimentos de familiaridade, proximidade podem levar a ideia de "nós" contra "eles".

**Importante:** Os Princípios não devem ser tomados como normas que governam comportamentos, mas como dispositivos de interpretação usados pelas partes interessadas para compreensão de uma situação [Sbisa 1980]. Isso quer dizer que eles podem estar ou não em ação no contexto do caso analisado. Na DISMISS eles embasam a identificação de aspectos que podem influenciar a criação, difusão e consumo de desinformações.

Além dos Princípios de Influência, a Plataforma Tecnológica possui características que moldam a forma como um objeto digital é representado, e como ela restringe e facilita a interação da audiência com um objeto desinformativo.

A DISMISS considera aspectos relacionados à **sincronicidade**, **direcionalidade** e os mecanismos de **controle** que constituem a PT.

- A sincronicidade diz respeito ao uso simultâneo do canal de comunicação pelas partes interessadas, podendo ser síncrona ou assíncrona. Em um **formato síncrono** o objeto desinformativo está acessível somente durante a interação da audiência com o meio de comunicação, ou seja, ele não está (ou ficará) armazenado para acesso posterior - como a transmissão de uma partida de futebol ao vivo. No **formato assíncrono** o acesso ao objeto desinformativo pode ocorrer em momentos diferentes, pois o mesmo fica armazenado e disponível para o acesso em um momento futuro. Isso tem implicações para o compartilhamento, a transformação, o rastreo e a responsabilização em um caso de desinformação.
- A direcionalidade está relacionada a forma como as partes interessadas interagem no canal de comunicação, podendo ser unidirecional ou bidirecional. Em um **meio unidirecional**, somente uma das partes envolvida na comunicação pode transmitir ou receber informações no canal de comunicação. Em um **meio bidirecional** ambas as partes podem tanto transmitir quanto receber informações no canal. Isso tem implicações para a forma como o objeto desinformativo é criado e disponibilizado, e para a rastreabilidade e responsabilização em um caso de desinformação.
- Quanto aos mecanismos de controle, todo canal de comunicação possui mecanismos de controle que limitam ou ampliam a capacidade das partes interessadas de se expressar ou interpretar mensagens. Os mecanismos se dividem em **Controle de Acesso** e **Controle de Qualidade**. O Controle de Acesso diz respeito ao que uma parte interessada pode fazer naquele canal e como pode fazer, como transmitir uma mesma mensagem para uma pessoa ou para várias pessoas, permitir uma parte interessada de acessar um determinado site na Internet, ou oferecer mecanismos que facilitam o compartilhamento de objetos digitais, etc. O Controle de Qualidade diz respeito à qualidade de um objeto nesse meio, se relacionando com aspectos da mensagem, como permitir a difusão de vídeos em alta qualidade, somente texto ou o tamanho e formato do objeto desinformativo.

Assim, cada tipo de tecnologia possui aspectos específicos que permitem um certo tipo de interação e tem implicações na interação da audiência com um objeto desinformativo.

O Quadro 3 mostra um exemplo de aspectos sociotécnicos relacionados. Os valores destacados em vermelho tem alguma característica em comum.

|        | Criação                                                                          | Difusão                                                                                                                 | Consumo                                                                              |
|--------|----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| Social | - Agrupado em comunidades de interesse mútuo e/ou pessoas conhecidas/confiáveis; | - Compartilhamento em grupos de amigos e conhecidos/ grupo;;<br>- Difícil identificar a fonte inicial da desinformação; | - Grupos polarizados<br>- Pandemia: população precisa se informar sobre tratamentos; |

|            |                                                                                |                                                                                        |                                                                              |
|------------|--------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------|
| Pragmático | Intencionalidade: enviesamento do conteúdo do artigo é indício de deliberação; | - App tem credibilidade;<br>- Pouco interesse em implementar políticas para mitigação; | - Viés de ancoragem podem facilitar o consumo;                               |
| Semântico  | - Linguagem pessoal, positiva e emocional;<br>- Apelo a autoridade;            |                                                                                        | - Linguagem acessível,<br>- Tom pessoal da mensagem;<br>- Tema de interesse; |
| Sintático  | - Mensagem encaminhada;<br>- Origem desconhecida;                              | Carece de mecanismos de mitigação;                                                     | - Sem links externos, imagens,<br>- Título, listagem;                        |
| Empírico   | - Acesso a internet e smartphones com WhatsApp no Brasil;                      | Meio assíncrono bidirecional<br>Interação muitos para muitos;                          | - Permite discussão<br>- Mídia de comunicação muito usada no Brasil;         |
| Físico     | - Brasileiros com Smartphones;                                                 | - Usuários em diferentes locais do Brasil;                                             |                                                                              |

Quadro 3 - Exemplo de ASE preenchido.

## Resultados Esperados

Ao final desta etapa o analista terá os artefatos da DISMISS preenchidos, representando uma perspectiva ampla de aspectos sociotécnicos em diferentes níveis de abstração que impactam a criação, difusão e consumo de objetos informativos. Isso significa o entendimento de como o objeto desinformativo foi construído, quais aspectos sociotécnicos são influentes para as partes interessadas que interagem com ele, e como ele pode ter sido entendido pelas partes interessadas.

Com estes artefatos preenchidos, o analista pode se perguntar como ele pode minimizar o impacto dos fatores de influência para aprimorar uma TIC, ou definir requisitos para o desenvolvimento de uma estratégia de mitigação adequada ao contexto da análise.

Além disso, a partir dos aspectos levantados, se pode realizar estudos longitudinais, investigar correlações entre os aspectos sociotécnicos em busca de causalidade.

A DISMISS tem potencial de apoiar o desenvolvimento de soluções de mitigação, desde análises jornalísticas até informações de detecção automáticas.

## Agradecimentos

Este passo a passo foi elaborado com ajuda dos membros do laboratório de Interação Humano-Computador da Universidade Federal do Paraná. Minha eterna gratidão pelo apoio e apontamentos que possibilitaram o refinamento da DISMISS.

## Referências

- Bazan, S. (2017). A new way to win the war. *IEEE Internet Computing*, 21(4), 92-97.
- Cialdini, R. B., & Cialdini, R. B. (2007). *Influence: The psychology of persuasion* (Vol. 55, p. 339). New York: Collins.
- Cybenko, G., Giani, A., & Thompson, P. (2002). Cognitive hacking: A battle for the mind. *Computer*, 35(8), 50-56.
- Fernandez, M., & Alani, H. (2018, April). Online misinformation: Challenges and future directions. In *Companion Proceedings of the The Web Conference 2018* (pp. 595-602).



- Floridi, L. (1996). Brave. Net. World: the Internet as a disinformation superhighway?. *The Electronic Library*.
- Fourney, A., Racz, M. Z., Ranade, G., Mobius, M., & Horvitz, E. (2017, November). Geographic and Temporal Trends in Fake News Consumption During the 2016 US Presidential Election. In *CIKM* (Vol. 17, pp. 6-10).
- Freitas Melo, P. D., Vieira, C. C., Garimella, K., Melo, P. O., & Benevenuto, F. (2019, December). Can WhatsApp counter misinformation by limiting message forwarding?. In *International conference on complex networks and their applications* (pp. 372-384). Springer, Cham.
- Guimarães, M. S., Baranauskas, M. C. C., & Martins, E. (2008, June). Communication-Based Modelling and Inspection in Critical Systems. In *ICEIS* (5) (pp. 215-220).
- Hinz, O. (2007, January). False Information in Internet Auction Communities. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)* (pp. 178b-178b). IEEE.
- Kolkman, Martin. "Problem articulation methodology." (1993).
- Liu, K. (2000). *Semiotics in information systems engineering*. Cambridge University Press.
- Puska, A., Piccolo, L., & Pereira, R. (2020). DisMiss False Information: A Value Matter. *Human Computer Interaction and Emerging Technologies: Adjunct Proceedings from*, 75.
- Safieddine, F., Dordevic, M., & Pourghomi, P. (2017, July). Spread of misinformation online: Simulation impact of social media newsgroups. In *2017 Computing Conference* (pp. 899-906). IEEE.
- Stamper, R. (1973). Information in business and administrative systems.
- Tudjman, M., & Mikelic, N. (2003). Information science: Science about information, misinformation and disinformation. *Proceedings of Informing Science+ Information Technology Education*, 3, 1513-1527.
- Zhou, X., & Zafarani, R. (2018). Fake news: A survey of research, detection methods, and opportunities. *arXiv preprint arXiv:1812.00315*, 2.