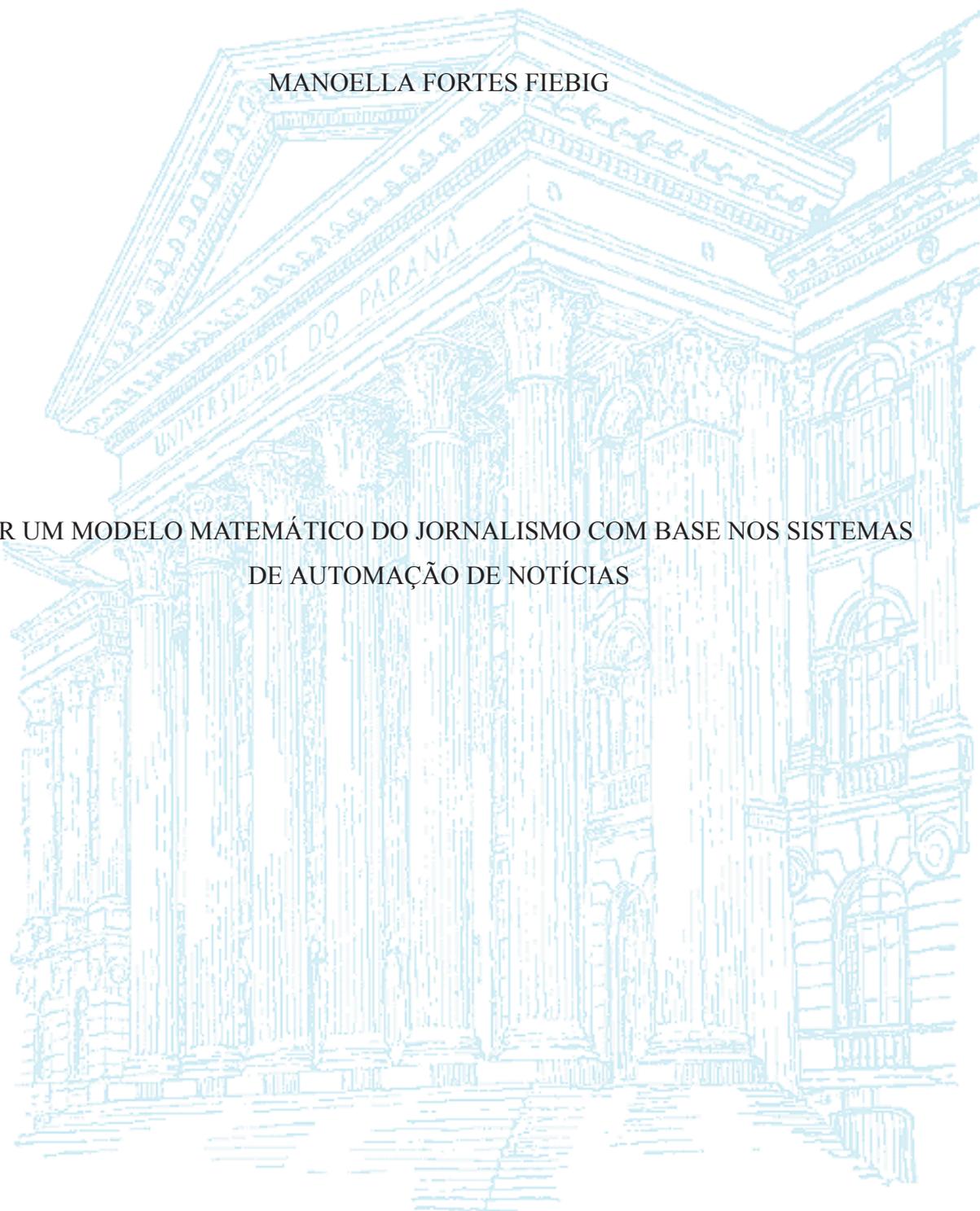


UNIVERSIDADE FEDERAL DO PARANÁ

MANOELLA FORTES FIEBIG

POR UM MODELO MATEMÁTICO DO JORNALISMO COM BASE NOS SISTEMAS
DE AUTOMAÇÃO DE NOTÍCIAS



CURITIBA

2023

MANOELLA FORTES FIEBIG

POR UM MODELO MATEMÁTICO DO JORNALISMO COM BASE NOS
SISTEMAS DE AUTOMAÇÃO DE NOTÍCIAS

Texto apresentado ao Programa de Pós-Graduação em Comunicação, Área de Concentração em Comunicação e Sociedade, Setor de Artes, Comunicação e Design da Universidade Federal do Paraná, como requisito para a obtenção do título de Doutora em Comunicação.

Orientadora: Prof^ª. Dr^ª. Claudia Irene de Quadros

CURITIBA

2023

UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS
BIBLIOTECA DE ARTES COMUNICAÇÃO E DESIGN - CABRAL

- F452 Fiebig, Manoella Fortes
Por um modelo matemático do jornalismo com base nos sistemas de automação de notícias. / Manoella Fortes Fiebig. – 2023.
1 recurso online: PDF
- Orientadora: Profª. Dra. Claudia Irene Quadros.
Tese (doutorado) – Universidade Federal do Paraná, Setor de Artes, Comunicação e Design, Programa de Pós-graduação em Comunicação.
Inclui referências.
1. Comunicação. 2. Jornalismo. 3. Matemática. 4. Teorias do jornalismo. 5. Teoria matemática da comunicação. 6. Lógica. I. Quadros, Claudia Irene. II. Universidade Federal do Paraná. Setor de Artes Comunicação e Design. Programa de Pós-graduação em Comunicação. III. Título.

CDD: 302.2

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação COMUNICAÇÃO da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **MANOELLA FORTES FIEBIG** intitulada: **Por um Modelo Matemático do Jornalismo com Base nos Sistemas de Automação de Notícias**, sob orientação da Profa. Dra. CLAUDIA IRENE DE QUADROS, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutora 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 Junho de 2023.

Assinatura Eletrônica

29/06/2023 17:42:45.0

CLAUDIA IRENE DE QUADROS
Presidente da Banca Examinadora

Assinatura Eletrônica

29/06/2023 17:29:22.0

RODRIGO EDUARDO BOTELHO FRANCISCO
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

29/06/2023 11:36:05.0

MAURÍCIO AUGUSTO PIMENTEL LIESEN NASCIMENTO
Avaliador Externo (UNIVERSIDADE TUIUTI DO PARANÁ)

Assinatura Eletrônica

02/07/2023 11:29:14.0

RAQUEL RITTER LONGHI
Avaliador Externo (UNIVERSIDADE FEDERAL DE SANTA CATARINA)

DEDICATÓRIA

Para João Claudio e Pedro Jair. Enfim, consegui. Conseguimos!

AGRADECIMENTOS

O que pode acontecer em quatro anos? Certamente, muita coisa. O imprevisível acontece, a realização de inúmeros sonhos e a destruição deles. Impossível resumir quatro anos em uma única página de agradecimentos. Seria equivocados, também, buscar citar nominalmente cada um e cada uma que fez parte desses últimos anos de vida. Isso porque o doutorado acontece no mesmo trajeto das belezas e dos desgostos desta vida que é vivida. Adaptamo-nos ao doutorado para que ele caiba na agenda, no pensamento, no cotidiano. Aos poucos, ele vai tornando-se cada vez mais parte de nós, e nós dele. O doutorado é transformador em muitos níveis, não apenas no amadurecimento intelectual; e esse processo é permeado por pessoas, experiências, aprendizados.

Assim, deixo meus agradecimentos, em geral, a cada um que por mim passou durante este período, que deixou vestígios de vida pelo caminho. Embora possa correr o risco de não nomear adequadamente algum desses afetos, arrisco-me, no entanto, a citar alguns poucos nomes. Estes, certamente, merecem o destaque:

Ao PPGCOM da UFPR, que me acolheu lá em 2015, quando iniciei o mestrado em Comunicação, e que me acolheu, em 2019, pela segunda vez, para o doutorado. Gratidão a todos os colegas, professores e afetos que foram construídos dentro do ambiente acadêmico e das paredes geladas do prédio amarelinho da rua Bom Jesus.

À CAPES, por ter concedido uma bolsa de estudos durante os primeiros meses do doutorado, o que viabilizou minha permanência no programa à época.

À minha orientadora, Claudia Quadros, que, desde antes do início de meu doutorado, sempre foi uma figura de afeto, compreensão e, principalmente, de inspiração. Foi um privilégio ser a primeira orientanda de doutorado da Claudia no programa de pós-graduação em Comunicação, da UFPR (PPGCOM/UFPR).

Aos membros da banca de qualificação, Maurício, Raquel e Rodrigo, que contribuíram sobremaneira para o resultado final desta pesquisa. Agradeço a generosidade, a disponibilidade e a inestimável fonte de inspiração que me concederam.

Aos meus familiares que, mesmo em meio aos tempos mais difíceis e desafiadores de nossas vidas, sempre apoiaram e incentivaram a conclusão do doutorado.

À Paula Queiroz e Giovanna Nonemacher, profissionais da Psicologia Cognitivo-Comportamental que acompanharam todo o percurso do doutorado e que, seguramente, foram grandes incentivadoras para que hoje eu pudesse apresentar este estudo a vocês.

E, por último, à Magali, a Shihtzu que esteve comigo durante a escrita da dissertação no mestrado e que se manteve firme ao meu lado durante todo o doutorado.

Escolher escrever é rejeitar o silêncio.

Chimamanda Ngozi Adichie

RESUMO

A proposta desta tese de doutorado é elaborar um modelo matemático para o jornalismo com base nos sistemas de automação de notícias. O jornalismo, tão conhecido pela palavra, pelo texto, pela narrativa, aqui é provocado a operar com números, algoritmos, teoremas e estruturas lógicas. A partir de uma revisão bibliográfica multidisciplinar, a construção deste percurso teórico busca estabelecer uma base sólida para apresentar uma proposta de modelo matemático para o jornalismo, defendendo a tese de que todo sistema de automação de notícias presente na quinta fase do jornalismo (CARREIRA, 2017) é passível de uma derivação lógica que corresponda aos objetivos da atividade jornalística, sendo o principal deles o de informar. Partimos da triangulação teórica envolvendo os campos do jornalismo, ciência da informação e lógica matemática para criar o percurso metodológico deste estudo. Recorremos a autores como Shannon (1949), Russell (1993), Bertocchi (2016), Carreira (2017), Devyatkin et.al. (2019), Ribeiro (2019), Traquina (2020) e Lage (2021) para discutir desde a *techkè*, passando pela história da tecnologia, pela era da informação, pela teoria matemática da comunicação, a teoria dos sistemas e as teorias sobre jornalismo de automação para contextualizar a proposição de um Modelo Matemático do Jornalismo. A pesquisa utiliza as estruturas algorítmicas do jornalismo de automação para propor um modelo matemático baseado no logicismo (Russell, 1993) e busca compreender o estado da arte em relação aos estudos sobre jornalismo e matemática no campo da automação jornalística. Para tanto, utilizamos a linguagem de programação *JavaScript* para o desenvolvimento das estruturas algorítmicas em um ambiente de desenvolvimento integrado (IDE, *integrated development environment*) para compreender, em nível anatômico, o funcionamento dos sistemas de automação de notícias. Assim, assumimos a matemática enquanto uma linguagem capaz de descrever o mundo por meio da simbologia lógica. Como resultado, observamos a relação simbólica entre jornalismo e matemática representada por meio do apontamento de teoremas baseados na estrutura algorítmica apresentada.

Palavras-chave: Comunicação. Jornalismo. Matemática. Teorias do Jornalismo.
Teoria Matemática da Comunicação. Lógica.

ABSTRACT

The purpose of this thesis is to develop a mathematical model for journalism based on news automation systems. Journalism, so well known for the word, the text, and the narrative, here is provoked to operate with numbers, algorithms, theorems, and logical structures. Based on a multidisciplinary bibliographic review, the construction of this theoretical path seeks to establish a solid basis to present a proposal for a mathematical model for journalism, defending the thesis that every news automation system present in the fifth phase of journalism (CARREIRA, 2017) is subject to a logical derivation that corresponds to the objectives of journalism, the main one being to inform, truthfully, the facts of society. We start from the theoretical triangulation involving the fields of journalism, information science, and mathematical logic to create the methodological course of this study. We turn to authors such as Shannon (1949), Bertocchi (2016), Carreira (2017), Devyatkin et.al. (2019), and Ribeiro (2019) to discuss aspects related to the Mathematical Theory of Communication, cybernetics, systems theory, and theories on automation journalism to contextualize the proposition of the Mathematical Theory of Journalism. The research seeks to understand the state of the art about studies on journalism and mathematics in the field of journalistic automation, in addition to creating an algorithmic structure capable of generating news through statistical and mathematical equations. Finally, we seek to prove the possibility of creating automation systems for journalism based on precepts of mathematical logic. For that, we use the JavaScript programming language for the development of algorithmic structures in an integrated development environment (IDE, integrated development environment). As a result, we can observe the relationship between journalism and mathematics through the presentation of theorems based on the algorithmic structure presented.

Keywords: Communication. Journalism. Math. Journalism Theories. Mathematical Theory of Journalism.

SUMÁRIO

1 INTRODUÇÃO	11
2 A TEKHNÈ	22
2.1 BREVE HISTÓRIA SOBRE OS CONCEITOS DE TÉCNICA E TECNOLOGIA	23
2.2 VIEIRA PINTO, O HOMEM MARAVILHADO E A FILOSOFIA DA TÉCNICA COMO RESULTADO	28
2.3 SOBRE A ESSÊNCIA DA TÉCNICA E JUÍZOS DE VALOR	31
3 COMUNICAÇÃO, INFORMAÇÃO E TECNOLOGIA.....	34
3.1 INOVAÇÃO É UMA QUESTÃO DE TEMPO: A HISTÓRIA D'O ANALISADOR DIFERENCIAL	34
3.2 O DIGITAL VENCE O ANALÓGICO E A EVOLUÇÃO D'O ANALISADOR DIFERENCIAL	39
3.3 ENFIM NASCE A CIBERNÉTICA	43
3.4 <i>KUBERNETES</i> DO SÉCULO XX	46
4 THE INFORMATION AGE E AS CONTRIBUIÇÕES DE CLAUDE SHANNON	50
4.1 NO MUNDO DA ENTROPIA, O RUÍDO É REI	61
4.2 SISTEMAS COMPLEXOS: DE SHANNON A BERTOCCHI	68
5 OS JORNALISMOS: DA TRADIÇÃO ORAL AO REPÓRTER-ROBÔ.....	77
5.1 O QUE É, O QUE É?	78
5.2 JORNALISMO É CAPITAL, CAPITAL É PODER	81
5.3 UMA HISTÓRIA CENTENÁRIA: CATEGORIZANDO AS FASES DO JORNALISMO ATÉ A CONTEMPORANEIDADE	85
5.4 O QUINTO JORNALISMO.....	88
6 CAN YOU TELL WHICH ONE OF THESE STORIES WAS WRITTEN BY A ROBOT?.....	94
6.1 PODE UM SISTEMA IMITAR A LINGUAGEM HUMANA?	101
6.2 NLG: A LINGUAGEM É ARTIFICIAL, MAS É NATURAL TAMBÉM	104

6.3 ELES JÁ ESTÃO ENTRE NÓS	110
6.3.1. Automated Insights.....	111
6.3.2. Opportunity Gap.....	112
6.3.3. Arria NLG	113
6.3.4. ChatGPT	115
7 A MATEMÁTICA PARA O JORNALISMO: DA FILOSOFIA MATEMÁTICA AOS ESTUDOS SOBRE O TEMA.....	121
7.1 LOGICISTAS, UNI-VOS	121
8 O ESTADO DA ARTE: ESTUDOS SOBRE JORNALISMO E MATEMÁTICA NO CAMPO DA AUTOMAÇÃO JORNALÍSTICA	125
8.1 BASES DE DADOS: BDTD, CAPES E GOOGLE ACADÊMICO	125
8.2 OS PERIÓDICOS CIENTÍFICOS	138
8.2.1 Revista Científica E-Compós	139
8.2.2 Revista Científica Galáxia.....	139
8.2.3 Revista Intercom.....	140
8.2.4 Revista Matrizes	141
8.2.5. Revista Famecos.....	142
9 CRIANDO ESTRUTURAS ALGORÍTMICAS: AS BASES PARA UM MODELO MATEMÁTICO DO JORNALISMO POR MEIO DE ITERAÇÕES LÓGICAS	144
9.1 ANATOMIA DO SISTEMA: COMPREENDENDO O FUNCIONAMENTO DE UM SISTEMA DE AUTOMAÇÃO JORNALÍSTICA.....	144
9.2 TEOREMAS PARA SISTEMAS DE AUTOMAÇÃO DE NOTÍCIAS	154
9.3 UM MODELO MATEMÁTICO PARA O JORNALISMO.....	160
CONSIDERAÇÕES FINAIS.....	165
REFERÊNCIAS.....	169

1 INTRODUÇÃO

É verdade universalmente reconhecida que jornalistas e a matemática se tornaram rivais combatentes durante a história da profissão. Talvez você, meu interlocutor, tenha compreendido a referência à notável escritora inglesa, Jane Austen, nas primeiras palavras desta tese. Embora pudesse citar trechos de um dos romances mais famosos da literatura inglesa, este texto não se propõe a descrever personagens como Darcy e Lizzie, tampouco ser uma história de ficção. Aqui, nas páginas que se seguem, descrevemos capítulos de uma história real, que se inicia antes mesmo da invenção da escrita, quando a cultura oral era predominante na comunicação humana.

Esta é uma tese que se propõe a compreender um dos inúmeros desafios impostos pela cibercultura e que ressoa em uma comunidade específica de profissionais que, em decorrência dos avanços tecnológicos, têm se reinventado a cada década. Se, no passado, o desafio imposto era o da transição dos equipamentos analógicos para os digitais, atualmente o desafio que se impõe é o da inteligência artificial, do *big data*, do processamento de linguagem natural e do *machine learning* que, aplicados ao exercício profissional do jornalismo, estabelecem uma nova era na labuta jornalística. Nessa nova era, ainda são incipientes as regras, normativas e princípios que regem a profissão, o que torna seu estudo essencial para compreender como se dá essa nova forma de trabalho, na qual coexistem repórteres humanos com algoritmos e sistemas de automação em uma mesma sala de redação. É fato que a inteligência artificial tem avançado no território de inúmeras profissões (na medicina, na advocacia, no marketing, nas instituições financeiras, nas escolas e universidades, na indústria, etc.), e com o jornalismo não seria diferente. Embora muito se fale sobre assuntos relacionados à inteligência artificial, como sistemas de automação, inteligência computacional e processamento de linguagem natural, há uma matéria em específico que permeia todos esses pilares, que parece estar sendo deixada de lado. Esse ponto em particular envolve números, lógica, quociente, equações, estatística e resultados.

Em matéria publicada em agosto de 2021 na *Época Negócios*, a professora e pesquisadora Dora Kaufman apresenta o fenômeno da inteligência artificial no jornalismo enfatizando duas polaridades que permeiam a discussão, tanto acadêmica, quanto profissional, acerca da adoção de sistemas de automação nas redações jornalísticas. No texto **Inteligência Artificial no jornalismo: ameaça ou oportunidade?**, Kaufman afirma que aos poucos, no Brasil, as empresas de mídia estão apropriando-se de sistemas de automação capazes de

produzir, publicar e compartilhar notícias de maneira automatizada, dispensando o trabalho manual feito por um jornalista. No texto, a autora descreve um fenômeno que é reflexo de uma tendência já bastante discutida e implementada em países como Estados Unidos e Inglaterra, mas que chega aos poucos no território brasileiro. Ao apresentar os fatos, a autora concatena à discussão um questionamento central que envolve a automação no jornalismo: ameaça ou oportunidade?

Esse mesmo questionamento é feito por um grupo de pesquisadores que publicou, em conjunto, o relatório *News Automation: The rewards, risks and realities of machine journalism*, organizado por Carl-Gustav Lindén e Hanna Tuulonen (2019). No documento, os autores discutem os efeitos da automação no jornalismo em diversos níveis, evidenciando a necessidade de pensar-se criticamente acerca do fenômeno. Além disso, o relatório também identifica a emergência da qualificação do estudante de jornalismo para enfrentar um mercado profissional mutável, bem como o treinamento de professores e profissionais que já estão no mercado para que acompanhem o avanço da automação tanto em nível acadêmico (preparando docentes para apresentar e repassar, de forma didática, essa nova tendência no jornalismo), quanto em nível profissional (preparando jornalistas maduros para protegerem seus cargos de trabalho enquanto a mudança acontece ao seu redor).

Ambos os artigos foram escritos antes de novembro de 2022. A data marca um momento histórico para aqueles que pesquisam ou se interessam por ferramentas de automação e inteligência artificial e como essas ferramentas interagem com o cotidiano social. Naquele mês, a OpenAI (laboratório norte-americano de pesquisas sobre inteligência artificial) lançou no mercado mundial o ChatGPT, ferramenta de processamento de linguagem natural que busca imitar a linguagem humana e que tem o potencial de criar redações, artigos, notícias, fazer cálculos e até criar códigos de programação, dependendo do comando realizado pelo usuário. Segundo dados da própria OpenAI, seus desenvolvedores treinaram um modelo de inteligência artificial que interage de forma conversacional com o usuário; o diálogo pode ocorrer de diversas formas, como: respondendo a perguntas, admitindo erros, identificando e contestando premissas incorretas, e rejeitando perguntas inapropriadas¹. Em menos de meio ano, a OpenAI já lançou novas versões e, em abril de 2023, momento em que esta tese está sendo escrita, a ferramenta já chegou na sua quarta versão, o GPT-4, que promete ser o sistema² mais avançado

¹ Trecho original: "We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer follow-up questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests" retirado do site www.openai.com.

² Trecho original: "*most advanced system, producing safer and more useful responses*", retirado do site www.openai.com.

de inteligência artificial, produzindo respostas mais úteis e seguras, além de trazer dados mais atuais do que a primeira versão da ferramenta³.

Quando o ChatGPT foi lançado, esta tese já estava em curso. E, mesmo falando sobre os sistemas de automação no capítulo 4, sentimos a necessidade de abordar com maior profundidade essa ferramenta em questão. Por isso, destinamos um espaço, nas páginas que se seguem, para evidenciar potenciais e limitações do ChatGPT e como ele se relaciona com o trabalho jornalístico.

Longe de posicionar-se de maneira inflexível em relação aos avanços da tecnologia na vida em sociedade, esta tese se propõe a discutir criticamente o fenômeno da automação e da inteligência artificial no cotidiano, tomando o cuidado necessário para não cair em um discurso puramente tecnófilo⁴ e excluir qualquer dimensão negativa sobre os impactos da automação no jornalismo, ponderando, também, suas faces obscuras para, por outro lado, não ser acusada de exercer uma tecnofobia⁵ exacerbada. O assunto requer um pensamento panorâmico, plural e cuidadoso, e esta tese não visa a esgotar o assunto nas páginas que se seguem. Ao contrário, nosso objetivo aqui é justamente entrar para a discussão, mostrando que há muitas lentes para enxergarmos o mesmo fenômeno. Aqui, escolhemos uma lente multidisciplinar e criticista, que envolve não apenas o olhar para o jornalismo, mas também suas associações com outras áreas como matemática, estatística e ciência da informação e da computação.

Pode parecer um contrassenso envolver jornalismo e matemática na mesma frase, mas o fato é que a automação no jornalismo existe justamente em decorrência da junção dos preceitos matemáticos às regras jornalísticas (e o nosso propósito é mostrar como isso acontece e refletir sobre esse fenômeno). A frase "sou de humanas, não de exatas" provavelmente seja o grito de guerra mais difundido entre os jornalistas, que, por muito tempo, preferiram residir no conforto das palavras, afastando-se cada vez mais dos números e equações. O jornalismo, tão

³ Em novembro de 2022, quando o ChatGPT foi lançado, as informações contidas no banco de dados que alimenta a ferramenta tinham uma limitação temporal: o ChatGPT só tinha informações até setembro de 2021, não constando dados a partir dessa data. Já na atualização GPT-4, a ferramenta já conta com um banco de dados atualizado, com informações em tempo real. Quando perguntamos à ferramenta se os dados da versão paga são diferentes da versão gratuita, a ferramenta nos respondeu a seguinte frase: "Não, a versão paga do ChatGPT não tem informações mais atualizadas do que a versão gratuita. Ambas as versões são alimentadas pelos mesmos dados e têm o mesmo conhecimento disponível. A diferença entre as versões gratuita e paga está no número de solicitações que você pode fazer ao modelo em um determinado período de tempo e nos recursos adicionais que podem ser disponibilizados. No entanto, em termos de conhecimento e informações, ambas as versões são as mesmas."

⁴ Segundo Rüdiger (2011), tecnófilos são os "advogados de defesa" da tecnologia, também conhecidos como "populistas tecnocráticos".

⁵ Segundo Rüdiger (2011), a tecnofobia é o efeito de aversão aos avanços tecnológicos e os pesquisadores dessa linha formam um grupo de "acusação política e moral do fenômeno" cibernético, também conhecidos como os "conservadores midiáticos".

conhecido pela palavra, pelo texto, pela narrativa, aqui é provocado a operar com números e estruturas lógicas. O desafio a que nos propomos aqui é justamente compreender a relação entre palavra e número, sentença e equação, interpretação e prova real. E, para evidenciar esse relacionamento que pode parecer estranho à primeira impressão, vamos conduzir este estudo por uma linha de raciocínio que busca deixar clara a afeição entre o jornalismo e a matemática.

Na trilha desta pesquisa, dividimos o trabalho em etapas, começando por uma história secular para compreender o que é a técnica e como ela permeia a própria construção da cultura contemporânea. Essa é uma incursão necessária para compreendermos que o fenômeno técnico evolui de forma conjunta à evolução do próprio homem, acompanhando-o durante todas as fases de seu desenvolvimento, desde a cultura oral à digitalização dos dados (ou *datification*, como afirmam alguns autores⁶). Do conceito de técnica, do grego *tekhnè*, adentramos no terreno da filosofia com o intuito de incorporar um pensamento criticista à narrativa com o auxílio de autores como Lévy (1999), Lemos (2002), Vieira Pinto (2005), Rüdiger (2011) e Carreira (2017).

O segundo e o terceiro capítulos da tese possuem uma linha teórica condutora: do princípio do pensamento sobre a técnica e da filosofia da técnica (VIEIRA PINTO, 2005), passamos para a história da cibernética sob um ponto de vista sistêmico, invocando diversos referenciais teóricos na explanação sobre o tema (SHANNON, 1948; WIENER 1968; ISAACSON, 2014). Do centro do pensamento cibernético, é que introduzimos a teoria da informação e sua conceituação, inaugurando uma nova fase no que tange às teorias da comunicação e adicionando novos ingredientes ao caldo das reflexões sobre emissão, mensagem e recepção. Com o auxílio de pesquisadores como Isaacson (2014), Soni & Goodman (2017) e Carreira (2017), trazemos à tona a necessidade de refletir sobre os conceitos cibernéticos sob um ponto de vista histórico e multidimensional.

Com um arcabouço teórico que busca deixar evidente a evolução da tecnologia junto com a sociedade, a partir do capítulo 5 é que apresentamos, com ênfase, o jornalismo e seu desenvolvimento histórico em nossa sociedade. Iniciamos com uma incursão na história do jornalismo (LAGE, 1985; RÜDIGER, 2003; MATTELART, 2012; TRAQUINA, 2012), buscando compreender suas fases e os desafios decorrentes de cada época. Nesse capítulo, utilizamos o neologismo os "jornalisms", nos referindo ao jornalismo plural, justamente para evidenciar sua evolução histórica e suas diferentes características em cada uma de suas fases. Aqui vale recorrer à memória histórica para concluir que as "novas tecnologias" podem não ser

⁶ Ver: DOS SANTOS, 2019.

consideradas tão novas assim, já que, como veremos, muitas das invenções que agregaram grandes desafios à profissão, são inspiradas em tecnologias do passado. Ao discutir as fases do jornalismo (MARCONDES FILHO, 2002), destacamos o conceito de "quinto jornalismo", apresentado por Carreira (2017), que avança nas reflexões a respeito da entrada dos sistemas de automação nas redações jornalísticas, seus impactos, paradigmas e *ethos*.

O jornalismo de automação ganha evidência também nesse capítulo, trazendo as reflexões de autores como Arce (2009), que apresenta, inclusive, alguns caminhos para pensar-se logicamente a estrutura do lead, de forma a ser um ponto de partida para propormos uma equação lógica para o jornalismo, e também autores como Mendonça (2016), Lindén (2018; 2019), Diakopoulos (2019), Milosavljevic & Vobic (2019), Ribeiro (2019), Sirén et. al (2020). Esses pesquisadores somam-se a este estudo com o objetivo de compreender o fenômeno da automação tanto do ponto de vista teórico e algorítmico, como também social, econômico e profissional.

Após apresentarmos o jornalismo de automação de forma teórica, recorrendo a todos os autores acima citados e a outras contribuições teóricas, no capítulo 6 buscamos compreender, com profundidade, o conceito de automação no jornalismo. Nesse momento, invocamos, especialmente, contribuições da área de análise e desenvolvimento de sistemas (RIBEIRO, 2019) para compreender de fato o funcionamento desses sistemas que são característicos da quinta fase do jornalismo (CARREIRA, 2017). Nesse capítulo, buscamos, ainda, expor alguns exemplos de sistemas de automação, apresentando ao leitor softwares diferentes que utilizam *natural language generation* (geração de linguagem natural) para criar narrativas (jornalísticas ou não, como é o caso do ChatGPT). Entre os capítulos 5 e 6, a esta autora busca definir "os jornalismo" trazendo conceitos com o auxílio de referenciais teóricos clássicos. Para isso, realiza um "jogo" de palavras com o uso de símile (o que é, o que é?) para evidenciar o percurso histórico do jornalismo até chegar em sua quinta fase, a saber: jornalismo é técnica, jornalismo é indústria, jornalismo é ciência, jornalismo é automação.

Ainda no capítulo 6, apresentamos o uso do ChatGPT, ferramenta de automação e inteligência artificial que irá, sob comandos, escrever uma notícia em poucos segundos. Em seguida, confrontamos a notícia escrita pela ferramenta com uma notícia escrita por um correspondente internacional. Nesse exercício, será possível visualizarmos algumas inconsistências na ferramenta de inteligência artificial, que embora corresponda à estrutura de uma notícia (LAGE, 2021) de forma satisfatória, também indica uma ruptura no preceito de veracidade indicado pelo autor.

No capítulo seguinte (de número 7), traremos ainda o conceito de lógica, de Bertrand Russell (1993), para ancorar nossas futuras inserções no território da matemática com as proposições de teoremas e do modelo matemático do jornalismo, com base nos sistemas de automação. O capítulo 7 nos trará bases da área da filosofia da matemática para pensar a matemática enquanto linguagem, ou seja, como uma forma de descrever fenômenos. Além disso, com base em Russell (1993), buscamos demonstrar que qualquer fenômeno da natureza - ou, qualquer fenômeno "ordinário", no sentido de "comum", "natural", conforme aponta o autor - é passível, de acordo com os princípios logicistas, de uma dedução lógica.

A partir de uma revisão bibliográfica multidisciplinar, a construção deste percurso teórico busca estabelecer uma base de sustentação para apresentar, nos últimos capítulos (8 e 9), uma proposta de modelo matemático para o jornalismo, defendendo a tese de que todo sistema de automação de notícias presente na quinta fase do jornalismo (CARREIRA, 2017) é passível de uma derivação lógica (Russell, 1993) que corresponda aos objetivos do jornalismo, sendo o principal deles o de informar. Partimos da triangulação teórica envolvendo o jornalismo, a ciência da informação e a lógica matemática para criar o percurso metodológico deste estudo. Recorremos a autores como Arce (2009), Carreira (2017), Devyatkin et.al. (2019) e Ribeiro (2019) para apresentar uma fórmula matemática que corresponda a cada etapa do trabalho jornalístico, passando pela coleta de dados, produção da notícia até sua publicação. Uma vez organizadas dentro de um sistema, essas fórmulas possibilitam a automação do processo jornalístico, demonstrando que é possível concatenar determinados preceitos jornalísticos em estruturas lógicas. Essas estruturas buscam reger o trabalho automático do sistema na busca e coleta de dados, na formação de frases e lexicalização, bem como na checagem e publicação do material produzido.

Com o objetivo de compreender a afeição do jornalismo pelos números e pela lógica, o primeiro passo metodológico foi o de identificar, por meio de uma revisão exploratória de bases de dados científicos, qual o estado da arte a respeito dos termos "jornalismo e matemática" entre os anos 2000 e 2020. Foram consultados repositórios científicos nacionais, como o Catálogo de Teses e Dissertações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) e o Google Acadêmico, onde buscamos por resultados tanto em português, quanto em inglês. Além disso, também realizamos uma busca nos periódicos científicos melhor avaliados na área da Comunicação (E-Compós, Galáxia, Intercom, Matrizes e Famecos) no Brasil.

Essa aproximação foi necessária para evidenciar que há uma grande diferença entre a teoria matemática da comunicação (teoria já largamente estudada, como veremos no segundo

capítulo desta tese) e o modelo matemático para o jornalismo, que estamos propondo neste estudo. Nesta primeira etapa de pesquisa, o estado da arte se faz necessário por três razões:

Investigar e reconhecer os autores, instituições e grupos de pesquisa interessados na temática que envolve jornalismo e matemática;

Compreender o que está sendo pesquisado neste eixo entre jornalismo e matemática para identificar pontos de convergência e assimilações;

Estruturar um plano de palavras-chave e termos de pesquisa que dê conta de integrar não apenas as áreas do jornalismo e da matemática, mas que evidencie a intersecção de outros termos, como "teoria matemática do jornalismo", como veremos a seguir.

Após a aproximação metodológica por meio da análise exploratória do campo em que estamos nos inserindo, partimos para uma pesquisa aplicada com o intuito de verificar a viabilidade da nossa proposição. Essa etapa, dividida em duas fases, foi um ponto essencial para medir as possibilidades e limitações da proposta desta pesquisa. Nesse sentido, vale realizarmos uma incursão cronológica e voltar ao ano de 2019, quando a autora ingressava no Doutorado com a ideia de um projeto de pesquisa voltado às limitações e possibilidades da automação no jornalismo. Com o passar dos anos e o natural contato com novas teorias, o projeto de pesquisa que buscava compreender as limitações e possibilidades da automação no jornalismo e o papel do jornalista (humano) dentro do fenômeno dos sistemas de automação de notícias evoluiu para uma proposta mais ousada, que flerta com a busca pelo desconhecido, pelo inédito.

Observamos, durante o caminho, que os sistemas de automação poderiam ser um objeto de pesquisa extremamente interessante e com diversas consequências na esfera sociotécnica, mas que havia ao seu redor outras possibilidades de pesquisa, de reflexão, de provocações e questionamentos que poderiam desencadear novos olhares sobre o assunto. E foi assim, pela evolução natural da pesquisa e pelo amadurecimento intelectual decorrente do processo de doutoramento, que decidimos por mudar o horizonte deste estudo, compreendendo que ainda há o espaço para refletirmos sobre o jornalismo de automação, mas mudando o foco da pesquisa. O jornalismo e a automação continuam presentes neste novo projeto de tese e figuram como um pilar teórico essencial para avançarmos na discussão sobre um modelo teórico matemático do jornalismo, mas, neste novo cenário, não o consideramos como problema de pesquisa central da reflexão.

Após a análise exploratória sobre o estado da arte acerca das palavras-chave desta pesquisa, passamos para a etapa de desenvolvimento de uma aplicação com foco em automação de notícias em um ambiente de desenvolvimento integrado (IDE, *integrated development*

environment), local onde apresentamos a estrutura lógica capaz de gerar notícias automáticas por meio de comandos (funções dentro do sistema). Essa etapa se faz necessária para compreendermos, em nível anatômico, como funcionam esses sistemas. A última etapa do caminho metodológico apresenta teoremas que correspondem a cada etapa do processo dentro do sistema de automação, que dão conta de ilustrar seu funcionamento. Por fim, elaboramos um modelo matemático do jornalismo baseado nos conceitos de progressão e logicismo, de Russell (1993).

Nas páginas que se seguem, interessa-nos criar uma nova forma de olhar para as teorias do jornalismo e – por que não? – sugerir abertura epistemológica capaz de inaugurar uma nova perspectiva de pesquisa na área. Acima falamos que o projeto deste estudo evoluiu para uma proposta ambiciosa. Foi refletindo sobre estas duas áreas, a matemática e o jornalismo, que identificamos que, no meio desta relação, há espaço para criar novos pontos de convergência. O (novo) problema de pesquisa começava a delinear-se quando, no início de 2022, ao ler uma dissertação sobre sistemas automatizados de geração de notícias⁷, um *insight* veio à tona. Na ocasião, o tensionamento que a autora fazia entre a teoria matemática da comunicação e o jornalismo me causaram certo desconforto. Daqueles desconfortos que te colocam para pensar, sabe? Parecia-me que ali estava inscrita uma nova proposição teórica, mas que ninguém estava olhando para ela com a atenção que merecia. Nascia um problema de pesquisa; "dos bons", pensei. Mas foi somente dias depois, já com a pulga atrás da orelha, que encontrei a coragem para propor uma mudança drástica neste estudo. Afinal, quem mudaria o tema da tese de doutorado àquela altura do campeonato?

Um novo problema nascia junto ao questionamento: é possível criar um algoritmo jornalístico capaz de gerar notícias? E, além disso, é possível criar um modelo matemático para o jornalismo tendo como base o funcionamento de um sistema de automação? Desse questionamento, ainda surgiram outros, como: será que é possível analisar o jornalismo pelas lentes da matemática, afinal de contas?

Essas perguntas começaram a permear os estudos e reflexões acerca desta tese de doutorado. Mas foi pensando na justificativa de pesquisa que tomamos a decisão de seguir adiante, já que, ao observar a história do jornalismo, percebemos que a matemática frequentemente foi uma matéria rejeitada e colocada em segundo plano. Gehlen (2016), após realizar uma pesquisa com profissionais e estudantes de jornalismo, constatou que 73% dos alunos de jornalismo e 60% dos jornalistas convencionais de diversas redações brasileiras

⁷ Notícias automatizadas: a evolução que levou o jornalismo a ser feito por não humanos (CARREIRA, 2017).

afirmaram não gostar de matemática (GEHLEN, 2016, p. 189). Em seu texto sobre o "jornalismo da (im)precisão", o autor discorre sobre os tratamentos jornalísticos concedidos aos números nas práticas profissionais da reportagem e relata que a temática do estudo foi concebida originalmente a partir de inquietações que dizem respeito à relação de jornalistas com os números (GEHLEN, 2016, p. 244).

O autor decidiu dedicar-se ao estudo dessa relação após deparar-se recorrentemente com algumas declarações, tanto de jornalistas quanto de alunos de jornalismo afirmando que não gostam de matemática, mesmo que esta seja matéria presente no cotidiano da profissão, especialmente em editorias como economia e finanças. Para o autor, essa separação trouxe "evidências de uma cultura profissional de distanciamento cognitivo dos jornalistas com o conhecimento matemático-quantitativo" (GEHLEN, 2016, p. 244).

Neste estudo, percebemos também o distanciamento do jornalismo e da matemática, porém, de um ponto de vista técnico. Gehlen (2016) refere-se a alguns problemas especialmente relacionados a como os jornalistas tratam o amplo uso de dados numéricos na *construção da notícia*, ou seja, em seu estudo, o autor avalia essa relação de um ângulo prático. Do ponto de vista técnico que buscamos desenvolver neste estudo, refletimos que, assim como a estrutura de um prédio, que requer medições certeiras e cálculos precisos, uma notícia também pode ser estruturada por meio de um cálculo, de um modelo matemático. Assim como um sistema de gestão de finanças, um sistema de geração de notícias também utiliza progressão aritmética para recuperar dados e informações. Assim como uma receita requer quantidades exatas e procedimentos ordenados, a produção de uma notícia também trabalha com sucessões e procedimentos ordenados, como é o caso do *lead* jornalístico.

A matemática está em todos os locais e com o jornalismo não seria diferente. Foi atribuída a Galileu Galilei a criação da célebre frase "a matemática é o alfabeto no qual Deus escreveu o Universo"⁸. Aqui, a matemática é compreendida no sentido de linguagem, como forma de descrever os fenômenos naturais. É como se houvesse uma linguagem universal, invisível, que sustenta todos os eventos da natureza e que pode ser sintetizada por meio da lógica, da sucessão numérica e dos aspectos relacionais entre esses eventos. Segundo Russell (1993), em **Introdução à Filosofia Matemática**, todo evento ou fenômeno do mundo físico pode ser observado e analisado pelas lentes da matemática, pela abstração e pela lógica, que

⁸ A frase é frequentemente atribuída ao físico e matemático alemão Johannes Kepler. No entanto, não encontramos evidências concretas de que Kepler tenha sido o autor da frase. Consta que essa citação pode ter sido uma variação de uma declaração feita por Galileu Galilei. Portanto, não há uma referência bibliográfica específica para a frase.

levam a deduções sobre os acontecimentos do mundo. Assim, conforme o autor, por meio da lógica é possível criar um novo pressuposto matemático, que não faz parte da "matemática natural"⁹, e sim de outro sentido matemático, que se constitui como "parte" da matemática natural, mas que avança na busca pelo entendimento dos fenômenos do mundo.

O nosso desafio, aqui, é justamente progredir nas abstrações entre o jornalismo e a matemática, na busca por evidenciar essa relação de forma que fique clara não apenas a afeição do jornalismo com a matemática, mas também as possibilidades que nascem desse casamento. Além disso, como um objetivo abstrato, é também foco da pesquisa a abertura de uma nova perspectiva de análise nos estudos em comunicação e jornalismo, trazendo matérias como a matemática, a ciência da informação e outras tantas áreas para discutir fenômenos da contemporaneidade.

Para este estudo, elencamos como objetivo geral propor um modelo matemático do jornalismo com base no funcionamento dos sistemas de automação de notícias, objetos presentes na quinta fase do jornalismo. Como objetivos específicos, listamos três que nos auxiliarão no processo de desbravar as vias para a proposição de uma nova perspectiva dentro do jornalismo e que serão perseguidos ao longo das páginas que se seguem:

Compreender o estado da arte em relação aos estudos sobre jornalismo e matemática, com foco em buscar obras que discutam a relação entre os campos;

Dissecar o funcionamento de um sistema de automação, criando teoremas lógicos relacionados a cada etapa do sistema;

Comprovar a possibilidade de criar um modelo matemático para o jornalismo, com o auxílio de sistemas de automação deste com base em preceitos matemáticos;

Com o objetivo de gerar conhecimento para a reflexão epistemológica, mas principalmente para a aplicação prática e desenvolvimento deste campo de estudos interdisciplinar, o resultado desta tese visa a gerar bases para que o jornalismo e a matemática sejam enfim reconhecidos como disciplinas afins e passíveis de relacionamento, de convergência e colaboração. É por isso que, nos capítulos que se seguem, buscamos criar uma linha de raciocínio que começa na história da técnica e termina na proposição dessa nova perspectiva no jornalismo, que não ousa, em primeira instância, autodeclarar-se uma nova

⁹ Para Russell (1919), a matemática natural pode ser entendida no sentido construtivo, "no sentido da complexidade gradualmente crescente: dos inteiros para os fraccionários, os números reais, os números complexos; da adição e multiplicação para a diferenciação e integração e daí para a matemática superior" (pág. 13).

"teoria matemática do jornalismo", mas que, certamente, inaugura um pensamento acerca dessa possibilidade.

2 A TEKHNÈ

Em 1997, o filósofo Pierre Lévy publicava seu livro com o título **Cibercultura**. Logo na primeira linha da obra, que se tornou um clássico entre pesquisadores do fenômeno da cibercultura, o autor escreve seu propósito com o livro: pensar. Lévy inicia sua obra com um objetivo simples: pensar, refletir, ponderar sobre tudo aquilo que envolve a cibercultura e as potencialidades (positivas e negativas) desse fenômeno técnico nos planos econômico, político, cultural e humano. Não por acaso, Lévy divide seu pensamento em três grandes partes: na primeira, busca oferecer definições sobre termos e conceitos que envolvem a cibercultura (o que é a técnica?, o que é real?, o que é virtual?, o que é digital?, o que é interatividade?, entre outros); na segunda parte, o objetivo é conduzir o leitor a algumas proposições acerca da cibercultura no mundo moderno, identificando traços dos avanços tecnológicos em diversas instâncias, como música, arte, filosofia, economia, educação, democracia; por último, na terceira parte de sua obra, Lévy expõe alguns dos grandes problemas gerados pelo fenômeno técnico na modernidade, tecendo um pensamento crítico¹⁰ acerca dos planos de resistência da tecnologia na sociedade moderna e seus possíveis impactos no futuro.

Mesmo que suas abordagens atualmente sejam consideradas ultrapassadas e até superadas por outros autores¹¹ que também pensam a cibercultura por diversas lentes, é necessário reconhecer, por mérito histórico, a importância dessa e de outras obras¹² do filósofo para os campos da comunicação, cultura e filosofia da técnica. E, assim como fez Lévy nos anos 1990, aqui, nesta tese, também nos propomos a “pensar” aspectos provenientes dessa cibercultura. Pensar sobre os impactos da técnica na sociedade ao longo dos séculos e, em especial, pensar na complexidade do avanço de sistemas de inteligência artificial em nosso tempo e como eles interagem com o jornalismo.

André Lemos, pesquisador brasileiro cujo histórico acadêmico é correlacionado diretamente com o pensamento teórico de Pierre Lévy, também se debruçou a pesquisar sobre esses aspectos. Lemos, no entanto, faz uso de outro verbo para definir seus estudos: “compreender”. Na página 26 do livro que resultou de sua tese de doutorado sob orientação de Lévy, Lemos expõe seu objetivo principal: “compreender os desafios da cibercultura” (LEMOS, 2002, p. 26) e, logo em seguida, defende que, para atingir tal propósito, é obrigatório

¹⁰ Embora Lévy deixe clara sua posição “otimista” e “tecnófila” nas primeiras páginas do livro, busca também evidenciar alguns aspectos da cibercultura de modo criticista no fechamento de sua obra.

¹¹ Francisco Rüdiger argumenta sobre este assunto em **As teorias da comunicação** (2010).

¹² Entre as obras importantes de Pierre Lévy traduzidas para o português estão: *As Tecnologias da Inteligência* (1993); *Inteligência Coletiva* (1994); *O que é o Virtual?* (1996); *Ciberdemocracia* (2002).

"buscar, nas raízes do fenômeno técnico, a compreensão da cultura contemporânea" (idem). Isso quer dizer que, antes de discutirmos sobre questões do nosso tempo, como *machine learning*, *big data*, inteligência artificial e outros termos decorrentes dos avanços tecnológicos no século XXI, é necessário "compreendermos" nossa história. E faremos isso a partir de agora, com uma incursão histórica nas raízes da própria experiência humana, revisitando conceitos que evoluíram simultaneamente à história do homem e que resultam no momento que vivemos: do digital, da automação, do pós-humanismo que nos assombra.

2.1 BREVE HISTÓRIA SOBRE OS CONCEITOS DE TÉCNICA E TECNOLOGIA

"The human world is technological to its core"¹³
Pierre Lévy (1999)

Lemos (2002) argumenta que não há maneira de compreender os paradoxos, as potencialidades e os conflitos gerados pelo avanço tecnológico na modernidade sem incorrer no pensamento histórico e filosófico da técnica. O primeiro passo, no entanto, deve ser o de distinguir conceitos que frequentemente se confundem: a técnica e a tecnologia. Para tanto, o autor recorre ao pensamento grego, valendo-se de sua acepção original e etimológica. A *tekhnè*, do grego, pode ser traduzida por "arte" e pode ser compreendida como um conjunto de atividades práticas, como a "arte" de criar ou elaborar algo, como por exemplo "a elaboração de leis e a habilidade de contar ou medir, passando pela arte do artesão, do médico ou da confecção do pão, até as artes plásticas." (LEMOS, 2002, p. 26).

Em outras palavras, a *tekhnè* compreende o "saber fazer" humano, as atividades práticas desenvolvidas ao longo do tempo pelo homem, aquilo que distingue os seres humanos racionais de outros seres, ou seja, segundo Lemos, a *tekhnè* compreende a distinção do "fazer" humano daquilo que é puramente da natureza (a flor, a folha, a luz do sol, etc.), cujo conceito pode ser atribuído à *phusis* (da mitologia romana considerada como a "natureza").

A *tekhnè* se distingue da natureza porque tudo o que é "técnico" é proveniente do ser humano e não se assemelha ao desenvolvimento de uma planta, por exemplo. Uma referência prática para auxiliar na compreensão das diferenças entre a *tekhnè* e a *phusis* é o desenvolvimento de uma flor, cuja semente brota, por seu próprio curso e segue seu desenvolvimento sozinha de acordo com o tempo da natureza; já o vaso, no qual podemos

¹³ Em sua tradução literal, a frase significa "o mundo humano é tecnológico em essência". Na tradução realizada por Carlos Irineu Costa, e publicada no Brasil pela Editora 34, a frase assume o seguinte sentido: "o mundo humano é, ao mesmo tempo, técnico" (LÉVY, 1999, p. 22).

plantar uma flor, é fruto da *tekhnè* humana, pois, provavelmente, foi criado por alguém, cuja prática e conhecimento resultaram na peça final onde a flor irá desenvolver-se futuramente. Segundo Lemos (2002), "o conceito de *tekhnè* é, assim, fruto de uma primeira filosofia da técnica que visa distinguir o fazer humano do fazer da natureza, este último, autopoietico, guardando em si os mecanismos de sua autorreprodução." (p. 27).

Outro autor que se debruça a estabelecer uma ampla conceituação a respeito dos significados da técnica e da tecnologia, fazendo incursões às origens históricas, é Francisco Rüdiger. Em outras palavras, o autor corrobora as afirmações de Lemos (2002) quando evidencia, em sua obra, que, de acordo com a concepção grega, "a técnica é, com efeito, uma forma de saber de que o homem se serve para produzir o que a natureza não lhe proporciona espontaneamente" (RÜDIGER, 2011, p. 76). Sendo assim, segundo as acepções de Lemos (2002) e Rüdiger (2011), que recorrem aos conhecimentos dos gregos antigos para conceituar a técnica, é possível contextualizá-la como uma forma de conhecimento prático da realidade.

Rüdiger (2011) descreve que, originalmente, "a técnica tinha a ver com a práxis criadora individual, o desenvolvimento da habilidade humana imediata" (2011, p. 76) e que provavelmente a confusão entre a técnica e a tecnologia nasceu da (re)interpretação moderna dessas expressões em termos epistêmicos e logicistas, que levam mais em consideração a forma como "os filósofos se expressam sobre a técnica, situando-se em um ponto de vista próprio, que difere daquele dos cidadãos em geral e muitas vezes em contradição com os vigentes nas várias formas de vida existentes em seu tempo e em seu meio social." (p. 77). Isso quer dizer que, mesmo que façamos incursões históricas para compreender a distinção entre a técnica e a tecnologia, é necessário também levar em consideração o contexto de cada vertente filosófica que se propõe a compreender a origem dessas palavras. Com fôlego, o filósofo brasileiro apresenta um capítulo inteiro em sua obra na busca por explicar, de acordo com diversas fontes, a conceituação da técnica, buscando evitar incursões equivocadas entre a conceituação de técnica e tecnologia.

O autor avança em suas considerações, reconhecendo nos gregos, assim como Lemos (2002), a origem da técnica, perpassando por diversos períodos e pensamentos filosóficos acerca do termo: dos antigos aos modernos, passando pela época de ouro do prometeísmo, a ascensão da visão fáustica até chegar ao pensamento cibernético, inaugurando, enfim, a era da informação. É necessário salientar que, para Rüdiger (2011), a técnica é um conceito fundamental para compreender os avanços tecnológicos na sociedade. Segundo ele, a técnica assume posições de destaque em cada uma das fases de sua evolução e, em suma, acaba por diluir-se na cultura moderna, tornando-se algo intrínseco ao homem e à sociedade. Perceba que,

até o momento, o termo "tecnologia" ainda não aparece de forma tão evidente neste resgate cronológico, mesmo que a história tenha forjado sua existência. E isso acontece por um motivo: porque, mais tarde, entende-se que a tecnologia surge com a combinação da *tekhnè* (que, frisamos, é anterior à tecnologia) à *logos*, também do grego, que significa "linguagem, palavra".

Sendo assim, a tecnologia teria sua origem na combinação entre os termos *tekhnè+logos*, quando, finalmente, nasce a concepção de estudo da técnica. A partir da combinação desses termos, avançava o entendimento de que a técnica não era apenas um conjunto de habilidades práticas que eram passadas de forma predominantemente oral, mas também envolvia a aplicação do pensamento racional num processo que buscava avançar na resolução de problemas de ordem prática. Sendo assim, a tecnologia passava a figurar como uma forma de conhecimento que envolvia tanto as habilidades práticas (empíria) quanto o uso da razão e observação do mundo (teoria). A partir desse entendimento da tecnologia como o estudo da experimentação do mundo, decorrem outras disciplinas do pensamento moderno, incluindo a filosofia da técnica, como veremos a seguir.

Dentre os pensadores e filósofos da técnica mais relevantes do século XX, Gilbert Simondon merece destaque por conta de seus conhecimentos em áreas como a mecânica e a eletrônica. Ele foi o responsável por cunhar a expressão dos "modos de existência dos objetos técnicos" que, segundo Lemos (2002), pode ser considerada como uma tese quase que de caráter biológico da tecnicidade humana. Neste ponto, vale ressaltar que, para os gregos, a *tekhnè* é, de fato, o conjunto de habilidades intrínsecas ao ser humano, ou seja, poderíamos concluir que os gregos já compreendiam, a seu modo, a técnica como algo biológico. Carreira (2017) complementa a afirmação anterior defendendo que Simondon "percebia o objeto técnico como reflexo da realidade humana e atribuía a ele um papel de mediador entre o homem e a natureza" (CARREIRA, 2017, p.24).

Para Simondon, a técnica está para a cultura assim como sua própria natureza e, sendo assim, seria impossível separar uma coisa da outra. Argumentava que a tecnologia poderia ser entendida como uma espécie de extensão das habilidades humanas e jamais algo separado ou distinto de sua própria natureza, já que a tecnologia é um processo que nasce das necessidades humanas. Pensando no século XX, Lemos aponta a "máquina" como a grande "responsável pela sensação contemporânea de que a tecnologia não faz parte da cultura humana (ou é sua inimiga)" (2002, p. 30) e, em sua obra, o autor busca articular uma espécie de "ciência" que naturaliza a tecnicidade humana, naturalizando, assim, também, o produto desta última dentro

da cultura contemporânea: as máquinas. Lemos (2002) descreve brevemente a tese¹⁴ de doutoramento de Simondon da seguinte maneira:

Contra essa visão errônea de separação entre técnica e cultura, Simondon pretende fundar uma verdadeira 'tecnologia' (um logos da *tekhnè*), não como ciência aplicada, mas como uma filosofia dos mecanismos, à qual ele propõe o nome de 'mecanologia'. O modo de existência dos objetos técnicos, que vai caracterizar a tecnologia contemporânea; corresponde a uma lógica interna, um caráter genético do desenvolvimento das técnicas primitivas. (LEMONS, 2002, p. 31).

No trecho acima, fica claro que Simondon buscava explicar a tecnologia como um fenômeno intrínseco ao homem e que os objetos técnicos correspondem às necessidades criadas pelos próprios seres humanos no seu processo de evolução. Carreira (2017) lembra que, mesmo sem uma concepção clara, a tecnologia esteve presente "por toda parte, apesar de não ser assim nomeada" (CARREIRA, 2017, p. 22). A autora resgata um fato histórico que aconteceu em 1802 com o professor alemão Johann Beckmann, que dava aulas de economia e, segundo a autora, reintroduziu o termo tecnologia no ambiente acadêmico e filosófico "porque pensava que a crescente importância das artes úteis demandava que elas fossem ensinadas de forma mais sistêmica.

A partir desse momento, o conceito de tecnologia passou por várias tentativas de sistematização (CARREIRA, 2017, p. 22). O relato que data de 1802 é um marco no pensamento filosófico a respeito do conceito de tecnologia, já que fora empregado com o significado de "estudo da técnica" de forma sistematizada e, desde então, diversos autores, incluindo os brasileiros André Lemos e Francisco Rüdiger, investiram significativo tempo de estudo na construção de conhecimento acerca do tema. Segundo Carreira (2017),

A filosofia demorou tanto tempo para começar a estudar a tecnologia porque vários estudiosos achavam que ela tinha pouco interesse por ser supostamente apenas uma aplicação da ciência. Mas segundo Dusek (2009, p.50), parte deles percebeu que esta concepção poderia ser enganosa, pois tudo dependeria de como a ciência é definida. Se com ela nos referimos a ensaio e erro, experimentos controlados com leis matemáticas, a tecnologia poderia ser vista como ciência aplicada. (p. 22).

Vale ressaltar, no entanto, que mesmo na contemporaneidade, como aponta a autora, período em que a ciência ocupa um lugar essencial na sociedade e é considerada a propulsora para a maior parte das invenções tecnológicas, "podemos dizer que a gênese da tecnologia como ciência aplicada pode ser enganosa, já que muitas das invenções são produtos de ensaio e erro,

¹⁴ Du mode d'existence des objets techniques, 1958.

assim como do acaso, o que é conhecido como serendipidade." (CARREIRA, 2017, p. 23). Exemplo disso foi o caso da penicilina, que foi descoberta pelo médico inglês Alexander Fleming de maneira acidental, após ter deixado uma cultura de bactérias que eram responsáveis por provocar infecções em feridas de soldados da guerra, em placas dentro de seu laboratório e, após voltar de suas férias, observou que as placas onde o material estava depositado haviam sido contaminadas e estavam cobertas por uma espécie de bolor.

Ao analisar novamente as placas que haviam sido deixadas sem nenhuma supervisão durante o período em que esteve fora, Fleming percebeu que, ao redor dos locais onde havia o bolor, a bactéria que ele estava estudando não havia sido detectada, indicando que algo nas placas havia combatido o desenvolvimento daquelas bactérias. O "bolor" tratava-se de um fungo, o *Penicillium Notatum*, capaz de matar aquelas bactérias. Essa descoberta "acidental" viria a mudar a história da medicina ocidental e fora revelada por obra da serendipidade.

As grandes descobertas e invenções do homem são pontos que não devem ser eliminados da discussão a respeito da técnica e da tecnologia, pois, como Simondon, Lemos, Rüdiger e outros teóricos da filosofia da técnica argumentam, o homem é, por essência, um ser técnico. Carreira (2017), por sua vez, relembra que "em todas as épocas, o ser humano buscou, em suas criações, uma forma de expandir seus potenciais, de contornar os obstáculos impostos pela natureza e resolver as divergências que encontrava em seu caminho." (p. 21). De acordo com o filósofo brasileiro Álvaro Vieira Pinto (2005a, p. 74), "a evolução das máquinas nada mais é do que a evolução dos homens que a constroem". Sendo assim, o modo de existência dos objetos técnicos passa, essencialmente, pela discussão das habilidades práticas humanas em busca da resolução de suas próprias necessidades.

Segundo Lemos (2002), "a evolução da espécie humana é fruto desse movimento perpétuo e infundável, sendo a técnica responsável pela criação da segunda natureza - a cultura - num processo de desnaturalização do homem." (p. 31). Para o autor, a filosofia da técnica deixa claro que os objetos técnicos formam uma espécie de ecossistema cultural "onde a naturalização do artifício modifica o meio natural, da mesma forma que o meio natural vai impondo limites à atividade técnica humana." (LEMOS, 2002, p. 31). Sendo assim, para compreender a história da evolução da técnica é necessário também observar os movimentos que o homem empreendeu no seu processo de evolução em sociedade, levando em consideração o fato de que cada nova invenção técnica consiste em modificar, inegavelmente, o sistema de relação entre o homem e a natureza. Esse sistema de inovações que acompanha o homem desde os primórdios de sua existência, no entanto, precisa ser observado de forma dialética e não

isolada, já que cada nova criação é fruto de uma necessidade humana e que, por sua essência, causa algum impacto social, econômico, político ou prático.

O filósofo brasileiro Álvaro Vieira Pinto escreveu quase 1.500 laudas, datilografadas em máquina de escrever e corrigidas a mão, com o objetivo de compreender a essência do conceito de tecnologia. Segundo o autor, os animais são seres que não produzem, enquanto o homem é o único animal capaz de produzir por si próprio, racionalizar e comandar a produção dos meios de vencer suas limitações e dificuldades. Na obra **O Conceito de Tecnologia**, publicada após quase 20 anos após sua morte, o catedrático da Faculdade de Filosofia da Universidade do Brasil (denominada de UFRJ desde 1920), consagra-se postumamente como um filósofo da técnica, com uma formação clássica em filosofia aliada ao saber matemático.

2.2 VIEIRA PINTO, O HOMEM MARAVILHADO E A FILOSOFIA DA TÉCNICA COMO RESULTADO

Em face da era tecnológica, Vieira Pinto resgata o conceito de filosofia como o ato de "maravilhar-se" com algo e, após o êxtase do maravilhamento, passar a refletir sobre o assunto que causou tal admiração. Para compreender melhor esse conceito, o autor recorre às acepções de filósofos da Grécia Antiga, como Platão e Aristóteles, exemplificando a relação entre a filosofia e o ato de maravilhar-se. Vieira Pinto (2005) argumenta que "o homem das civilizações antigas se definia - a ponto de graças a este estado de espírito tornar-se um pensador - pelo ato de maravilhar-se." (p. 30) e que "o começo da filosofia equivale ao começo da capacidade de pensar." (idem). Sendo assim, a filosofia nasce no momento em que o homem, após maravilhar-se com as condições e fenômenos naturais, começa a refletir sobre estes e buscar soluções para questões da natureza. Nessa trilha de pensamento, o filósofo brasileiro recorre a Aristóteles:

por se maravilharem, os homens, tanto agora como no passado, começaram a filosofar, a princípio maravilhando-se com as dificuldades mais imediatas, e depois, avançando pouco a pouco, procuraram resolver problemas maiores, como os que se referem aos fenômenos da Lua, do Sol e das estrelas, e por fim, procuraram descobrir a gênese do Universo. Quem se depara com uma dificuldade e se admira reconhece sua própria ignorância (e por isso o amante de mitos é também de certo modo filósofo, pois o mito é composto de maravilhas). (VIEIRA PINTO, apud *Metafísica*, I, 2, 182 b).

A citação acima entende que o homem começou a "filosofar" observando o mundo e os fenômenos que o cercam. Em primeira instância, o ato de filosofar nasce da observação de coisas simples do cotidiano que causam "admiração". Em seguida, o homem foi progredindo para a observação de situações mais complexas, como a própria intenção de compreender as

leis do Universo. Essas conceituações são importantes, inicialmente, para que Vieira Pinto desenvolva sua concepção de "homem maravilhado" na contemporaneidade. Segundo o filósofo, "o que distingue o maravilhar-se atual do antigo é que agora o homem se maravilha não diante da natureza, mas diante de suas próprias obras" (VIEIRA PINTO, 2005, p. 35). Essa é uma passagem importante para assimilarmos seu pensamento acerca da filosofia da técnica. Para o autor, essa concepção "reedita o velho estado de espanto e maravilha, mas agora em face dos tempos que nos são dados" (Idem).

Sua argumentação é a de que, nos tempos atuais,¹⁵ o homem passa a maravilhar-se com suas próprias obras técnicas, ou seja, com seu próprio "produto", enquanto no passado maravilhava-se e, por conseguinte, filosofava sobre a natureza, ou, segundo a origem grega, a *phusis*. Vieira Pinto (2005) ressalta que "outrora, na pobreza de uma civilização tecnicamente atrasada, o homem só podia, com efeito, maravilhar-se com aquilo que encontrava feito; agora, na época da civilização tecnológica, extasia-se diante do que faz." (p. 36). O autor argumenta que, antes, o homem filosofava acerca de elementos e fenômenos da natureza, pois ainda não existiam avanços tecnológicos significativos. Atualmente, o oposto acontece: com uma infinidade de elementos técnicos criados pelas mãos dos homens, a filosofia volta-se para encontrar a razão da essência da técnica. Ou seja, antigamente a filosofia era

correlata de uma sociedade onde são relativamente poucas e pobres as criações artificiais, onde a intervenção do homem não podia substituir por fabricados seus os objetos oferecidos pela pura causalidade da natureza. À medida, porém, que vão sendo compreendidos os processos naturais e descobertas as forças que os movimentam, com a conseqüente possibilidade de utilização delas pelo homem, para produzir artefatos capazes de satisfazer novas necessidades, e essa fabricação se multiplica constantemente, o mundo deixa de ser simplesmente o ambiente rústico e espontâneo e se converte no ambiente urbano, na casa povoada de produtos de arte e, na época atual de aparelhos que põem as forças naturais a serviço do homem. (VIEIRA PINTO, 2005, p. 37).

Sendo assim, o autor considera que, com isso, a função da filosofia passa daquela que busca entender os fenômenos da natureza e a origem do universo para aquela que coloca o homem e suas técnicas no centro do exercício intelectual. Nesse momento, Vieira Pinto (2005) faz um parêntese em seu texto, trazendo uma reflexão que julgamos interessante frisar: coloca o fato de que o conceito de "natural" se modifica ao longo do tempo. Como exemplo, o autor sugere que na antiguidade era natural viver em meio à natureza; atualmente, para o homem

¹⁵ Vale ressaltar que a obra em questão fora publicada em 2005, mas os escritos de Vieira Pinto datam da década de 1970 já que, segundo Nota do Editor, a última lauda do manuscrito fora escrita no dia 5 de abril de 1973. Sendo assim, quando o autor fala de sua "época" é importante contextualizar temporalmente a obra.

civilizado, é natural viver em meio a artefatos eletrônicos, usufruir de luz elétrica e tantas outras comodidades.

Nesse sentido, vale reproduzir literalmente palavras do autor:

A simples interrupção da corrente elétrica doméstica, causando a escuridão e paralisação dos aparelhos mecânicos ou eletrônicos, parece uma anormalidade, sem que alguém pense tratar-se na verdade do retorno à normalidade antiga, pois a humanidade viveu por incontáveis milênios, em que floresceram grandes civilizações e se realizaram admiráveis obras de cultura, sem conhecer nenhum desses recursos de que hoje dispomos e acreditamos serem tão naturais que desperta admiração se por qualquer motivo nos faltam. (VIEIRA PINTO, 2005, p. 37).

Portanto, assim como Lemos (2002) e Rüdiger (2011), Vieira Pinto (2005) evoca as acepções antigas da *techkè*, recorrendo à filosofia de Aristóteles para compreender o fenômeno tecnológico em si. Para ele, a máquina supõe a racionalidade da técnica de forma material, ou seja, na máquina conseguimos visualizar a forma materializada da técnica. É por meio dos objetos técnicos, como as máquinas, que o homem passa a transformar a sua natureza e considerando válida essa proposição "tanto para a máquina mais complexa quanto para a mais simples das ferramentas, [...] compreenderemos que a noção de *técnica* assume posição central na reflexão sobre a atividade criadora efetuada pelo homem no mundo" (VIEIRA PINTO, 2005, p. 135-136).

Vale destacar, então, que, embora escritas em meados dos anos 1970, as páginas datilografadas pelo filósofo brasileiro, que anos mais tarde se tornaram a obra **O Conceito de Tecnologia**, dividida em dois volumes, servem de estrutura para pensar a técnica na contemporaneidade. Os exemplos que o autor cita, como a chegada do homem à lua, eram, à época, grandes feitos da humanidade. Hoje, no entanto, voltamo-nos para fenômenos como o ChatGPT, uma inteligência artificial com habilidades conversacionais que tem o potencial de responder às mais variadas perguntas, dependendo da forma como o questionamento é realizado pelo usuário que a utiliza.

Ainda, por mais que tenhamos bebido diretamente da fonte do autor, seria impossível esgotar as imensas contribuições de sua obra para os estudos sobre técnica e tecnologia e acerca das relações entre o avanço da tecnologia e sua profunda associação com as estruturas dominantes na sociedade, em especial por tratar-se de um pesquisador brasileiro, latino-americano, nos (já) longínquos anos 1970. Sendo assim, para estender o entendimento sobre os fenômenos técnicos na sociedade, evocamos outros autores que também refletem sobre a filosofia da técnica, ou, como diria Vieira Pinto: recorreremos a outros autores que também se *maravilharam* com a técnica.

2.3 SOBRE A ESSÊNCIA DA TÉCNICA E JUÍZOS DE VALOR

Quando falamos em tecnologia, frequentemente a imagem que se forma no imaginário social é a de objetos altamente tecnológicos, como robôs, super computadores, aeronaves e foguetes espaciais. A filosofia da técnica, por sua vez, inaugura um pensamento que busca compreender, com base na experiência e no questionamento, a essência daquilo que é técnico. Segundo Cupani (2011), "tradicionalmente, filosofar sobre um assunto começa pela indagação acerca do que isso é. A nossa pergunta inicial, portanto, consiste em indagar: que é a tecnologia?" (p.11). Quando colocamos em prática o questionamento a respeito do que é a técnica e do que é a tecnologia estamos, na verdade, realizando um exercício filosófico acerca do tema. Engana-se, no entanto, quem pensar que a filosofia da técnica é uma atividade moderna, pois, como lembra Cupani (2011), a filosofia da técnica acompanha a história (ou pode até mesmo datar de antes) da ciência, que nasce na Grécia Antiga por volta do século VI a.C. Segundo o autor,

a intervenção da ciência na produção de artefatos é vista geralmente como geradora de uma diferença importante entre a técnica tradicional, baseada no conhecimento empírico do mundo, e a tecnologia, resultante da aplicação do saber teórico. Por "ciência" entende-se, nesse contexto, o tipo de pesquisa da Natureza (e da sociedade) que combina a observação e o experimento com a representação e o cálculo matemáticos, tal como se pratica no Ocidente a partir da Modernidade (CUPANI, 2011, p. 14).

Assim como, à sua maneira, Lemos (2002), Vieira Pinto (2005) e Rüdiger (2011) argumentam em suas obras, Carreira (2017) também exemplifica que o pensamento filosófico da técnica, influenciado sobremaneira pelos filósofos da Grécia Antiga, cuja perspectiva coloca a técnica como "uma espécie de fuga, portanto, da condição humana, pois os homens sempre criam meios para enfrentar os limites e transpor os obstáculos que identificam" (p. 25). Isso quer dizer que, quando nos questionamos acerca da essência da técnica, é necessário observar sua "razão de ser" que, em suma é, do que um tipo de ação, motivada exclusivamente pela necessidade do homem de tornar mais eficiente suas condições de trabalho e conquista, além de dominar suas limitações em diversos níveis.

Ancorada em Simondon que, por sua vez, via o homem como um maestro que guiava os objetos técnicos dentro da cultura, Carreira (2017) destaca que essa visão coloca o homem como o centro da técnica, elucidando o papel central do humano nas criações tecnológicas e colocando o objeto técnico como um "reflexo da realidade humana", criado e idealizado

justamente para servir a um propósito antropocêntrico da vontade de dominar a natureza, ou seja, de colocar as necessidades humana no centro do universo e inventar tudo aquilo que a natureza seria incapaz de oferecer de forma orgânica. Esse pensamento de domínio da natureza permanece até o presente, basta voltar o olhar para as tecnologias de inteligência artificial, por exemplo, que buscam de forma implacável (e com eficientes avanços) recriar as conexões do cérebro humano em sistemas inteligentes.

A história da técnica, suas concepções e os questionamentos que dela derivam devem ser observados como uma realidade polifacetada, nas palavras do filósofo brasileiro Alberto Cupani. Para ele, o exercício de compreender a tecnologia exige que a enxerguemos de maneira mais ampla, quase como se fosse uma entidade ubíqua, cuja presença se espalha por todo o território do planeta: "aquilo que definimos como tecnologia se apresenta, pois, como uma realidade polifacetada: não apenas em forma de objetos e conjuntos deles, mas também como sistemas, como processos, como modos de execução, como certa mentalidade" (CUPANI, 2011, p. 11). O autor considera, ainda, que essa "presença múltipla" da tecnologia vem acompanhada, invariavelmente, de alguma valoração, ou seja, toda tecnologia que emerge traz consigo um potencial que é passível de julgamento, seja ele positivo ou negativo, como exemplificado a seguir:

Em certos casos, como no das armas de destruição em massa ou da poluição ambiental resultante da industrialização, consideramo-las como algo condenável, que desejaríamos não existisse. Em outros, como no das vacinas, do transporte confortável ou do cinema, a maioria das pessoas vê nessas realizações algo que veio beneficiar a espécie humana. Mas a propósito da maior parte dos objetos e processos tecnológicos há lugar para a dissensão, permanente ou circunstancial. É melhor dispor de ar condicionado ou repensar a arquitetura e a relação do homem com o meio ambiente? Deslocar-se de carro, embora sendo uma maneira mais cômoda e veloz de encurtar distâncias, não se converte em um hábito que quase elimina o exercício corporal, provocando doenças? O uso da internet, instrumento fantástico de informação e comunicação, não ameaça reduzir o conhecimento a acúmulo de dados e permitir que nossa vida pessoal seja controlada? De modo geral, é melhor ou pior, em algum sentido, a vida numa sociedade tecnológica? (CUPANI, 2011, p. 12).

A filosofia da técnica, enfim, contribui para uma reflexão significativa a respeito não apenas de nosso percurso histórico até a modernidade, mas também como um exercício de juízo sobre tudo aquilo que nos é apresentado. Especialmente no momento em que vivemos, na segunda década do início do século XXI, em que somos impactados à exaustão pelas propagandas e por novos aparatos que nos prometem aumento de performance, ubiquidade, onipresença e total domínio sobre nossas limitações, é necessário desenvolver um pensamento crítico acerca dos fenômenos da tecnologia, sobretudo sobre seus impactos nas mais diversas

esferas da vivência humana, a começar por cultura, saúde, educação, chegando até esferas mais afastadas do "ser humano" mas que o afetam diretamente, como a própria natureza, até a política, a economia e a comunicação. Assim como Cupani (2011), Lévy (1999) também reflete sobre o juízo de valor a respeito da técnica. Para ele, a "técnica não é boa nem má (isso depende dos contextos, dos usos e dos pontos de vista), tampouco neutra (já que é condicionante ou restritiva, já que de um lado abre e de outro fecha o espectro de possibilidades)." (LÉVY, 1999, p. 26).

Desejamos destacar também as contribuições de grupos de pesquisa brasileiros que buscam produzir e compartilhar conhecimentos a respeito da filosofia da técnica. Criada em 1983, a Associação Nacional de Pós-Graduação em Filosofia (Anpof) possui um grupo de trabalho com foco em discussões sobre a Filosofia da Técnica e da Tecnologia. Com encontros anuais, o GT, que conta inclusive com as contribuições de Cupani, busca estabelecer uma rede de pesquisas sobre as temáticas da tecnologia, da técnica e das tecnociências. O grupo, formado por professores, estudantes e pesquisadores de todo o Brasil, tem o objetivo de pensar, de forma multidimensional, o fenômeno, as instituições e os objetos técnicos em suas várias dimensões ontológicas, éticas, epistemológicas e estéticas¹⁶.

Há, ainda, grupos de pesquisas ligados a programas de pós-graduação em Filosofia, com inscrição na CAPES e que também buscam contribuir para os avanços nas pesquisas sobre o tema. O que parece ser uma constante nesses grupos é o pensamento de que a técnica pode ser, como ressalta Vieira Pinto (2005), "um modo de ser exclusivo do homem" (p. 138). Para o autor, é no movimento, ou seja, no "ato humano" que reside a concepção do conceito de técnica. E ao maravilhar-se com seus próprios feitos, o homem passa a filosofar - refletir, questionar, compreender - sobre suas obras, dando origem não apenas ao campo da filosofia da técnica, mas também abrindo vias para pesquisas em outros campos do pensamento.

¹⁶ Dados disponíveis no site da Anpof: <<https://anpof.org/gt/gt-filosofia-da-tecnologia-e-da-tecnica>>

3 COMUNICAÇÃO, INFORMAÇÃO E TECNOLOGIA

"Information is the resolution of uncertainty"¹⁷

Claude Shannon, considerado mundialmente como o "pai" da Teoria da Informação.

Vimos que a técnica é um ponto central quando falamos sobre tecnologia, mas, se aprofundarmos o entendimento sobre o assunto, veremos que a técnica também faz parte de um pilar essencial em abordagens teórico-metodológicas nos campos da comunicação e da informação. A origem da teoria da informação se confunde, inclusive, com a origem da cibernética. E nesta, a técnica está no núcleo das discussões. Nos próximos tópicos, buscamos evidenciar as relações entre esses campos e destacar a importância, tanto da cibernética quanto da teoria da informação, nos campos da comunicação e, em especial, do jornalismo. Neste estudo, queremos deixar evidentes as relações entre os campos e suas associações práticas. Para compreender essas ligações, olhemos para o conceito de informação:

3.1 INOVAÇÃO É UMA QUESTÃO DE TEMPO: A HISTÓRIA D'O ANALISADOR DIFERENCIAL

Com espírito romântico, herdado de seu pai, Lord Byron, e alta capacidade analítica, adquirida pela influência de sua mãe, que a incentivou a estudar matemática desde muito cedo, Ada Byron (Figura 1), que ficou conhecida como Ada Lovelace, será o ponto de partida para entendermos os primórdios da revolução digital e o início da era da informação. Considerada a primeira programadora da história (uma mulher!), enquanto participava de um dos tradicionais saraus noturnos da alta sociedade britânica, Ada teve seu primeiro contato, aos 17 anos, com o arsenal de invenções do matemático Charles Babbage que, segundo Isaacson (2014), era uma celebridade das ciências e da matemática na época.

Entre as invenções apresentadas naquele sarau, uma, em especial, chamou a atenção de Ada: a máquina que Babbage chamava de "analisador diferencial", uma espécie de computador que conseguia resolver equações de polinômios. Na ocasião, Ada teria ficado surpreendida com a máquina do inventor e, segundo Isaacson (2014), teria visto a "*imensa beleza* daquela invenção" (ISAACSON, 2014, p. 20, *grifo do autor*). Ada é conhecida por inaugurar o que se convencionou chamar de "ciência poética" e, segundo Isaacson, "o amor de Ada tanto pela

¹⁷A Mathematical Theory of Communication, 1948.

poesia quanto pela matemática levou-a a ver beleza em uma máquina de computação" (2014, p. 21). Além disso, o gosto pela literatura e a herança poética também se devem a seu pai, o escritor e poeta britânico cujo espírito romântico e boêmio resultou em problemas conjugais com sua esposa, mãe de Ada, que tinha medo de que a filha desenvolvesse o mesmo espírito do pai. Por essa razão, a mãe de Ada a incentivou a estudar matemática, aritmética, geometria euclidiana, álgebra e trigonometria.

Aos 18 anos, maravilhada com o analisador diferencial de Babbage, e tendo contato com outras invenções da época, Ada passou a maravilhar-se também com os estudos de Mary Somerville, uma das poucas mulheres a ter destaque nos estudos em matemática na sociedade britânica da época. Segundo Isaacson (2014), Somerville teve grande influência na formação e “se tornou amiga, professora, inspiradora e mentora de Ada. Encontrava-se com a jovem com regularidade, mandava-lhe livros de matemática, criava problemas para ela resolver e, paciente, explicava as respostas corretas" (p. 26). Além disso, ambas participavam dos saraus, incluindo visitas à propriedade de Charles Babbage, onde discutiam sobre as atualizações e o desenvolvimento das novas invenções do matemático.

FIGURA 1 – ADA LOVELACE, PINTADA POR MARGARETH SARAH CARPENTER, EM 1836



Fonte: Os Inovadores, Walter Isaacson (2014)

É importante destacar o nome de Babbage e Somerville porque ambos contribuíram muito para que Ada pudesse tornar-se a primeira programadora da História. Após seu casamento, Ada se distanciou dos estudos e mudou seu foco para a constituição de sua família. Entretanto, aos poucos, foi sentindo falta da sua rotina de estudos e passou a retomar sua atenção à matemática. Neste momento, ela entra em contato com Babbage e o convida a ser seu tutor. Babbage nega o pedido de Ada e, tempo depois, segundo Isaacson (2014), "apareceu uma oportunidade feita sob medida" (p. 36) para que Ada pudesse juntar-se a Babbage em seu trabalho: após um discurso feito por Babbage sobre sua nova invenção, a Máquina Analítica, uma máquina de "propósito geral", no Congresso de Cientistas Italianos em Turim, o jovem engenheiro militar Luigi Menabrea publicou um artigo em francês sobre a palestra. Após a publicação do manuscrito, Ada foi incentivada a fazer a tradução do artigo para o *Scientific Memoirs*, um periódico científico. Segundo Isaacson,

quando terminou o trabalho, ela informou a Babbage, que ficou agradecido, mas também certa medida surpreso. "Perguntei porque não tinha ela mesma escrito um artigo original sobre um tema que conhecia tão intimamente", disse ele. Ada respondeu que isso não lhe havia ocorrido. Na época, mulheres não costumavam publicar artigos científicos. (ISAACSON, 2014, p. 37).

Foi a partir desse fato que Ada pôde, enfim, publicar suas maiores contribuições para a tecnologia na época. Após aceitar a sugestão de Babbage, Ada começou a escrever algumas anotações para anexar ao trabalho traduzido de Menabrea. Essas anotações, publicadas como "Notas da Tradutora", acabaram "tendo 19.136 palavras, mais do que o dobro do tamanho do artigo original de Menabrea" (ISAACSON, 2014, p. 37). As notas de Ada foram publicadas com a assinatura A.A.L. e ficaram mais famosas do que o artigo original, isso porque Ada trouxe diversas contribuições que podem ser consideradas pioneiras para o pensamento científico da época. Segundo Isaacson,

Em suas "notas", Ada explorou quatro conceitos que teriam ressonância histórica um século mais tarde, quando o computador afinal nasceu. O primeiro era o de uma máquina de propósito geral, que pudesse não apenas desempenhar uma tarefa preestabelecida, mas que pudesse ser programada e reprogramada para desempenhar uma gama ilimitada e mutável de tarefas. Em outras palavras, ela visualizou o computador moderno. Esse conceito estava no cerne de sua "Nota A", que enfatizava a distinção entre a Máquina Diferencial original de Babbage e a nova Máquina Analítica que ele estava propondo. (2014, p. 38).

A "nota A", publicada por Ada, explicava o importante conceito de um "computador com propósito geral", já que os computadores da época eram máquinas cujo principal objetivo

era o de fazer cálculos. No texto, Ada relata que uma máquina analítica, de fato, teria o potencial de ir além de apenas tabular dados ou corresponder a função específica, podendo desenvolver outras funções caso fossem combinados símbolos em geral. Segundo Isaacson (2014), é na primeira "nota" publicada por Ada que ela descreve a essência dos computadores modernos. E reside aí sua primeira grande contribuição para os avanços na área.

Em sua segunda nota, Ada avança nas contribuições e apresenta, ela mesma, a descrição do funcionamento de uma máquina de propósito geral. Além disso, Ada expressa as diversas possibilidades que uma máquina de propósito geral poderia ter:

suas operações, ela imaginava, não precisavam ser limitadas à matemática e aos números. Usando a ideia de De Morgan sobre a extensão da álgebra para uma lógica formal, ela notou que um equipamento como a Máquina Analítica podia armazenar, processar e agir sobre qualquer coisa que pudesse ser expressa em símbolos: palavras, lógica, música e qualquer outra coisa que pudesse usar símbolos para ser transmitida. (ISAACSON, 2014, p. 38).

Como se não bastassem as contribuições que apresentamos até agora, em outra nota publicada, Ada inova mais uma vez e descreve em detalhes o funcionamento de um computador por meio de algoritmos. Foi por essa última contribuição que Ada passou a ser conhecida como a primeira programadora da História, pela capacidade de descrever o algoritmo capaz de computar os números de Bernoulli¹⁸. Ada descreveu uma sequência de operações e depois elaborou um gráfico que demonstrava como cada operação seria codificada pela máquina (Figura 2). Além disso, Isaacson (2014) destaca que, nessa última nota publicada por Ada, ela também fala sobre o conceito de sub-rotinas (uma sequência de instruções que desempenha uma tarefa específica, como computar um cosseno ou calcular juros compostos, e que pode ser inserida em programas maiores, se necessário) e de *loop* recursivo (uma sequência de instruções que se repete). Na época, ambos os conceitos eram praticamente inéditos e pouco discutidos na ciência. Segundo Isaacson (2014), "foi sobretudo por causa desse diagrama, que acompanhava os processos complexos de geração de números de Bernoulli, que Ada recebeu de seus fãs a honra de ser chamada de 'a primeira programadora de computador do mundo.'" (p. 41).

¹⁸ Segundo Isaacson (2014), os números de Bernoulli configuram uma série de números de extrema complexidade. Na teoria matemática moderna, os números de Bernoulli são estudados até hoje e são parte importante na programação de sistemas.

instruções que diziam ao hardware quais operações realizar. Nesse software estavam as fórmulas mágicas que podiam transformar as máquinas de maneira assombrosa. (ISAACSON, 2014, p. 100).

Dessa forma, Ada pode ser considerada como uma pioneira em sua área. E ainda que existam contribuições de outros pesquisadores da época (como o próprio Babbage, que influenciou diretamente as publicações de Ada), é notável a importância que ela representa para os avanços na linguagem de programação e na engenharia da computação. "A verdade é que a contribuição de Ada foi tanto profunda, quanto inspiradora", defende Isaacson (2014, p. 45). Segundo o autor, ela foi capaz de vislumbrar o funcionamento complexo das máquinas que hoje conhecemos como computadores, além de imaginar que as máquinas poderiam ter um potencial de processamento muito além do que apenas realizar cálculos. Ada "ajudou a plantar as sementes de uma era digital que floresceria cem anos mais tarde." (ISAACSON, 2014, p. 45).

3.2 O DIGITAL VENCE O ANALÓGICO E A EVOLUÇÃO D'O ANALISADOR DIFERENCIAL

Como vimos, o Analisador diferencial é uma máquina mecânica inventada por Charles Babbage, no início do século XIX. Entretanto, como denominamos no início deste subcapítulo "a inovação é uma questão de tempo" e, mais do que isso, é uma questão de colaboração. Essa é uma frase dita por Walter Isaacson no início do capítulo "O Computador", no qual o biógrafo refaz o percurso da evolução do Analisador Diferencial até chegar aos computadores modernos. Nesse capítulo, o autor deixa claro que os computadores que conhecemos são resultado da soma de esforços de diversos pesquisadores que foram essenciais para a evolução tecnológica nos séculos XIX e XX, como Vannevar Bush, Alan Turing, Claude Shannon, Howard Aiken, Konrad Zuse e John Vincent Atanasoff.

Destacamos, ainda, que vale a pena ler toda a obra de Isaacson para perceber, inclusive, as disputas de poder que ocorreram durante os mais de 150 anos de evolução dos computadores, o que torna uma tarefa difícil atribuir a apenas um nome a invenção do computador, já que existem diversos atores nesse processo. Aqui, no entanto, não nos interessa evidenciar essas disputas (de poder, de ego, de intelecto), mas entre os nomes citados acima, interessa-nos um em especial: Claude Shannon. E o analisador diferencial, que teve sua origem em Babbage, passando por Ada e tantos outros, também irá protagonizar essa história. Vejamos a seguir.

O Analisador Diferencial, computador mecânico criado por Babbage e desenvolvido na segunda metade do século XIX, com o objetivo de solucionar equações matemáticas e estatísticas, pode auxiliar-nos a compreender a história de outro conceito importante para a

evolução tecnológica: o conceito de *informação*. Após a Primeira Guerra Mundial, com seu exemplar de Analisador Diferencial em pleno funcionamento, o diretor do departamento de Engenharia do Instituto de Tecnologia de Massachusetts (*Massachusetts Institute of Technology*, o MIT), Vannevar Bush, buscava por um assistente que o auxiliasse no desenvolvimento e aplicação de pesquisas para aprimorar os resultados obtidos com o Analisador Diferencial. Em 1936, Claude Elwood Shannon assumia o cargo de assistente no MIT e logo percebeu que o sistema do Analisador Diferencial, que era operado por circuitos elétricos, chamados de relés, poderia ser operacionalizado por meio de um sistema aplicando princípios da lógica simbólica.

Segundo Carreira (2017), à época, o mundo da eletricidade e da lógica pareciam distantes e sua relação parecia pouco provável, até o surgimento das proposições inovadoras de Shannon. Para ele, no processo de funcionamento do computador, o que se pensava que um relé transmitia para outro dentro do circuito era a eletricidade. Entretanto, Shannon descobriu que essa comunicação entre relés não se tratava exatamente de eletricidade, e sim um fato lógico: que o circuito estava aberto ou fechado e que poderia corresponder, dentro de uma equação, a números. Percebam, novamente, a importância dos números para a evolução tecnológica, pois essa constatação será útil nos capítulos seguintes desta tese, quando, finalmente, vamos propor um modelo matemático aplicado ao jornalismo. Mas, antes disso, voltemos à Shannon.

A proposição de Shannon de que o circuito poderia ser explicado de acordo com a transmissão de uma informação de um ponto a outro indicava que a comunicação entre os circuitos do Analisador Diferencial poderia acontecer por meio de códigos binários (0 ou 1), que correspondem a sentenças pré-determinadas como sim ou não, verdadeiro ou falso, ligado ou desligado. Soni & Goodman (2017) destacam que

é claro que a informação existia antes de Shannon, assim como os objetos tinham inércia antes de Newton. Mas antes de Shannon, havia pouco senso de informação como uma ideia, uma quantidade mensurável, um objeto adequado para a ciência exata. Antes de Shannon, a informação era um telegrama, uma fotografia, um parágrafo, uma música. Depois de Shannon, a informação foi totalmente abstraída em bits. (SONI & GOODMAN, 2017, p 71 - tradução nossa¹⁹).

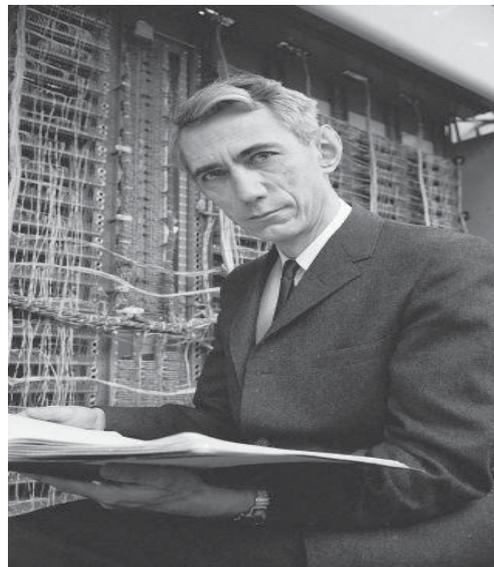
A descoberta de Shannon mudou o pensamento científico acerca da comunicação e da transmissão das informações. Para Soni & Goodman (2017), os estudos empreendidos por

¹⁹ Trecho original: "of course, information existed before Shannon, just as objects had inertia before Newton. But before Shannon, there was precious little sense of information as an idea, a measurable quantity, an object fitted out for hard science. Before Shannon information was a telegram, a photograph, a paragraph, a song. After Shannon, information was entirely abstracted into bits." (SONI & GOODMAN, 2017, p. 71)

Shannon sobre os interruptores elétricos que controlavam o funcionamento do Analisador Diferencial levaram-no a um *insight* que viria a inaugurar o pensamento da era digital. Segundo os autores, os interruptores teriam a capacidade de fazer muito mais do que apenas controlar o fluxo de eletricidade através dos circuitos, "poderiam ser usados para avaliar qualquer afirmação lógica em que pudéssemos pensar, poderiam até parecer 'decidir'. Uma série de escolhas binárias - ligado/desligado, verdadeiro/falso, 1/0 - poderia, em princípio, realizar uma imitação aceitável de um cérebro." (SONI & GOODMAN, 2017, p. 78 - tradução nossa²⁰).

Essa proposição, embora parecesse impensável, incipiente e improvável à primeira vista, seria responsável por revolucionar a forma como a engenharia elétrica enxergava o funcionamento dos circuitos elétricos, inaugurando a era da transmissão da informação por meio de sistemas binários (e que, na segunda metade do século XX, seria tema central para o avanço da cibernética, como veremos no próximo capítulo). Soni & Goodman (2017) retratam a importância dessa descoberta de Shannon quando relatam que suas proposições podem ser vistas integradas "nos circuitos dos nossos telefones, computadores, nossos satélites de TV, nossas sondas espaciais continuam amarradas à terra por finas cordas de 0's e 1's." (p. 99 - tradução nossa²¹).

FIGURA 3 – CLAUDE SHANNON NA CAPA DA TIME MAGAZINE



Fonte: Alfred Eisenstaedt / The LIFE Picture Collection

²⁰ Trecho original: "Shannon's study of the electrical switches directing the guts of that mechanical behemoth led him to an insight at the foundation of our digital age: that switches could do far more than control the flow of electricity through circuits - that they could be used to evaluate any logical statement we could think of, could even appear to "decide". A series of binary choices - on/off, true/false, 1/0 - could, in principle, perform a passable imitation of a brain". (SONI & GOODMAN, 2017, p. 78)

²¹ Trecho original: "that insight is embedded in the circuits of our phones, our computers, our satellite tvs, our space probes still tethered to the earth with thin cords of 0's and 1's". (SONI & GOODMAN, 2017, p. 99).

Segundo Carreira (2017), trabalhando no aprimoramento do Analisador Diferencial, Shannon demonstrou que necessitava apenas de dois números para suas equações diferenciais por integração. Vale ressaltar que ele era um matemático, engenheiro elétrico e também um telégrafo, o que contribuiu para a criação de sua teoria da informação por meio da conceituação de códigos binários. À época, um de seus maiores desafios era o da "possibilidade de alcançar a perfeição do raciocínio a partir da tradução correta do pensamento em códigos" (CARREIRA, 2017, p. 32), ou seja, seu objetivo era criar uma estrutura de códigos capaz de simular o pensamento humano de forma minimamente aceitável. O pesquisador acreditava que o raciocínio humano poderia ser descrito por meio de códigos e estruturas lógicas, como apresentam Soni & Goodman (2017):

Mas aqui estava Shannon trancado em uma sala com uma máquina construída para automatizar o pensamento - construída, em nome da indústria e eficiência, para remover a arte da matemática. E no meio de seu trabalho, percebeu que conhecia outra maneira de automatizar o pensamento, uma que acabaria se mostrando muito mais poderosa do que a máquina analógica. (p. 613 - tradução nossa²²).

Jimmy Soni e Rob Goodman²³ dedicaram-se a criar uma biografia de Claude Shannon. Publicada em 2017, a biografia conta com diversas histórias dos "corredores" do MIT, exemplares da vida pessoal de Shannon e curiosidades sobre o pesquisador que viria a tornar-se o "pai da Ciência da Informação". Como apontam os autores Soni & Goodman (2017), a intenção de Shannon era a de reproduzir o funcionamento do cérebro humano por meio das máquinas. E, uma vez que esses códigos fossem descritos e organizados em um sistema, a máquina seria capaz de desenvolver tarefas antes feitas apenas por humanos. A palavra *bit*, inclusive, deriva dessa descoberta. O neologismo surge como uma abreviação de "*binary digits*²⁴" e passou a identificar uma "unidade de medida", como o minuto, a hora ou o dia, que são consideradas unidades de medida de tempo. No caso dos bits, eles serviam para mensurar uma *unidade de informação*, sendo um conceito probabilístico cujas chances de ser 1 ou 0 são de 50% e isso determinaria a natureza da informação (sim ou não, *on* ou *off*, etc.), de acordo com suas probabilidades. Segundo Carreira (2017), Shannon inseriu esse conceito em sua

²² Trecho original: "But here was Shannon shut in a room with a machine built to automate thought - built, in the name of industry and efficiency, to remove the art from math. And in the midst of his work, he came to understand that he knew another way of automating thought, one that would ultimately prove far more powerful than the analog machine." (SONI & GOODMAN, 2017, p. 613).

²³ Jimmy Soni é jornalista, biógrafo e ex-editor-gerente do The Huffington Post. Rob Goodman é professor assistente de Política e Administração Pública na Toronto Metropolitan University.

²⁴ Dígitos Binários, ou *bits*, conforme abreviação proposta por Shannon.

monografia com o título de "Uma teoria matemática da comunicação²⁵", escrita em 1948 e publicada pela Revista Técnica dos Sistemas Bell.

3.3 ENFIM NASCE A CIBERNÉTICA

O século XX foi palco de inúmeras evoluções tecnológicas, um campo fértil para a inovação, e não foram poucas as consequências desses avanços. Diversos autores já publicaram estudos que discutem, de forma crítica e aprofundada, a complexidade e os efeitos dos avanços tecnológicos no último século; destacamos aqui Lévy (1995; 1999), Castells (2002; 2004), Marcondes (1997) e Rüdiger (2011). Já na segunda metade do século XX, nadando nas ondas da eletricidade e das máquinas elétricas, a sociedade começa a viver um novo ciclo de transformações, como aponta Rüdiger (2011). Para o autor, "após a substituição da força física do homem pela energia da máquina (primeiro a vapor, depois elétrica), assiste-se pouco a pouco à encampação das capacidades intelectuais e cognitivas por autômatos ou máquinas inteligentes" (105, 2011). Lemos (2002) destaca que "o paradigma eletricidade/petróleo, motor elétrico e química de síntese no final do século XIX muda, após o término da Segunda Guerra Mundial, para um novo paradigma: energia nuclear, informática e engenharia genética" (p. 53) na segunda metade do século XX. Segundo Lemos (2002), esse novo paradigma que surge na sociedade no século XX afetaria a vida cotidiana de modo radical, tecendo novos desafios à economia, política, ciência e cultura de forma planetária, figurando como o "pano de fundo para o surgimento da cibercultura" (idem). Vamos abordar, brevemente, a evolução cibernética, pois compreendemos que esta é uma história que tem seu início em Ada Lovelace, passa pelas contribuições de Shannon (que vamos abordar ainda de forma bastante aprofundada nas páginas seguintes), mas que há diversos pesquisadores e inventores que também fazem parte desse processo, que desencadeia não apenas uma revolução material tecnológica, mas que inaugura uma nova forma de cultura, transpassada pelas inovações tecnológicas, que ficou conhecida como "cibercultura", principalmente nos anos 1990, após algumas publicações clássicas da área, que iremos apresentar neste tópico.

Carreira (2017) e Rüdiger (2011) refazem cronologicamente o caminho da tecnologia até o início do século XXI. Os autores relembram que o conceito de cibernética nasce em meio a um grupo no qual matemáticos, físicos, biofísicos, psicólogos, engenheiros e cientistas discutiam a respeito de sistemas de telecomunicações e desenvolvimento de redes nos Estados

²⁵ Abordaremos a Teoria matemática da comunicação com maior ênfase no capítulo II.

Unidos. Esse grupo de intelectuais que participou das Conferências *Macy*²⁶ ficou conhecido como "cibernético", termo escolhido por Norbert Wiener, e cuja origem vem do neologismo grego *kubernetes*, que significa "timoneiro" ou "piloto". Segundo Rüdiger (2011), Wiener "criou o termo Cibernética após a II Guerra Mundial (1948) para definir, em seus termos, a ciência do controle das relações entre máquinas e seres vivos" (p.106). Já Carreira (2017), acrescenta que a tese de Wiener, à época, "era de que tanto os seres vivos como as máquinas faziam o mesmo esforço para controlar a informação vinda do mundo exterior e torná-la acessível a partir da transformação interna." (CARREIRA, 2017, p. 39).

No capítulo anterior, explicamos o surgimento dos *bits*, cujo precursor foi Claude Shannon. Esse conceito corresponde à ideia de que uma estrutura lógica pode dar conta de reproduzir o pensamento humano por meio de códigos. Na era cibernética, alguns conceitos preconizados por Shannon se complexificam, conforme aponta Rüdiger (2011), quando revela que, na cibernética,

a sociedade pode ser vista como um conjunto de circuitos e canais, mantido pelas redes técnicas de informação. Para ele [Wiener], a comunicação constitui a base de todos os fenômenos, sejam naturais ou artificiais. O propósito da cibernética é o de desenvolver uma linguagem e técnicas que nos permitam, de fato, enfrentar o problema do controle e da comunicação. (p. 108).

A cibernética, nesse caso, pode ser entendida como uma forma de controlar os meios de comunicação. Para a sociedade capitalista, a cibernética surge, então, como uma poderosa arma de controle social, político e econômico. Aqui, vale o parêntese para lembrarmos as proposições de Vieira Pinto (2005), que de forma irredutível, atrelava o avanço da tecnologia aos movimentos de classes dominantes, numa abordagem criticista do fenômeno tecnológico com forte influência do pensamento marxista. Assim, não é surpreendente que as maiores potências capitalistas do mundo detenham a maior parcela de empresas de tecnologia, figurando como polos de inovação não apenas nas instâncias profissionais e econômicas, mas também de armamento e controle bélico (veremos a seguir que o nascimento da cibernética coincide com os avanços em tecnologias militares).

O pensamento cibernético, conforme apontam os autores citados acima, está intimamente ligado aos ideais capitalistas do século XX, desencadeando não apenas o desenvolvimento de inovação, mas também, como aponta Lemos (2002), o surgimento de problemas como a poluição, maiores desigualdades sociais, econômicas e políticas, o caos

²⁶ O termo cibernética tem origem nas dez conferências Macy que aconteceram entre 1946 e 1953 e foram marcadas por discussões interdisciplinares.

urbano em determinados locais, o aumento da violência, entre outros efeitos que se correlacionam com a ideação de uma sociedade em rede, hiperconectada, ubíqua e aparelhada.

Faz-se necessário compreender que as teorias de informação e comunicação são permeadas por todas essas facetas. O conceito de "informação" é central nas teorias cibernéticas e, como aponta Rüdiger (2011), talvez a "informação" de Shannon seja "a peça mais essencial de todo o pensamento cibernético" (p. 108). Destacamos esse ponto porque ele é central também no que estamos propondo neste estudo. Tanto para os cibernéticos, quanto para outros pesquisadores, matemáticos e engenheiros que participaram da expansão das redes no século XX, a informação seria o ponto central da comunicação e, uma vez que a informação fosse concebida de maneira correta, a comunicação seria possível de ser efetuada de um ponto a outro, caso não houvesse interferência. Acontece, contudo, que compreendemos a comunicação como um processo social dinâmico e interativo, que se complexifica em diversas dimensões. Em nosso tempo, com os avanços vertiginosos das redes de comunicação digital e a instantaneidade das trocas de mensagens e compartilhamento de dados, elevamos a discussão sobre "o que é a comunicação" para outros patamares. Entram em cena outros conceitos que não serão discutidos diretamente neste estudo, mas que devem ser levados em consideração, como as teorias da comunicação da aldeia global, a ideia de que o meio é a mensagem, de McLuhann, as teorias da cultura de massas, os estudos culturais, as teorias relacionadas aos ativismos digitais, da obsolescência programada, entre outras que discutem e analisam, efetivamente, o que podemos considerar como "comunicação" na sociedade contemporânea.

Vale ressaltar que a comunicação como é concebida hoje não pode ser explicada por apenas uma teoria, ou uma disciplina, ou um campo do conhecimento científico. Defendemos, aqui, que a comunicação é permeada por diversas disciplinas e saberes, uma área multidisciplinar que flerta com conceitos teórico-metodológicos de outras áreas: correlatas, como é o caso das relações públicas e da publicidade; exatas, como a ciência da computação, matemática, física e estatística (como é o caso do estudo que aqui estamos propondo); e humanas, como a sociologia, filosofia e as ciências políticas. Compreender o conceito de informação e como ele é o ponto central da comunicação e do pensamento cibernético é essencial para adentrarmos, no próximo capítulo, na teoria matemática da comunicação e nos conceitos que fizeram de Shannon o "pai" da Teoria da Informação.

3.4 KUBERNETES DO SÉCULO XX

Como vimos acima, a cibernética é um projeto que une ideais capitalistas, militares e de controle social, e é na imbricação entre inovação e progresso que ela encontrou um terreno fértil para desenvolver-se. Mesmo que em períodos anteriores a cibernética já estivesse ganhando espaço, Rüdiger (2011) aponta para o momento em que as redes começam a ser estudadas e incorporadas na sociedade como o pontapé inicial para o que viria a ser a cibernética no século XXI, mais ligada às conexões entre máquinas e dispositivos.

Segundo o autor, "do ponto de vista mais imediato, os fundamentos do fenômeno podem ser ligados, porém, à criação de um sistema de comunicação por meio de computadores" (RÜDIGER, 2011, p. 17). Embora redes de comunicação entre máquinas já fossem projetos instaurados em diversas instituições de pesquisa (especialmente aquelas ligadas a setores do governo e militares, tanto nos EUA, quanto em outros locais da Europa), o primeiro registro de uma rede de comunicação (que permitisse não apenas realizar a ligação entre computadores, mas que tolerasse também a troca de mensagens) criada por técnicos e pesquisadores em informática, vinculados à secretaria de defesa do governo dos Estados Unidos, foi realizado em 1969. Vivendo num período de Guerra Fria, o projeto norte-americano era o de manter a comunicação entre suas máquinas de forma segura, mesmo em situação de guerra. O que conhecemos hoje como a rede mundial de computadores, a internet, teve seu início datado em 1983 e, depois dela, várias outras redes semelhantes foram sendo lançadas por pesquisadores e empresas interessadas no avanço das telecomunicações. Essas redes deram o *start* a um fenômeno que viria a ser concebido, anos mais tarde, e daria origem à internet. Várias redes foram sendo lançadas simultaneamente, porém, com diversas limitações, tanto de conectividade, quanto de infraestrutura e segurança. Foi somente em 1990 (mais precisamente no Natal daquele ano, em 25 de dezembro) que Tim-Berners Lee, ao lado de Robert Cailliau e um estudante de informática, implementou a primeira comunicação bem-sucedida entre um computador e um servidor, por meio da rede que seria chamada de internet.

FIGURA 4 – MARCOS DA HISTÓRIA DA ENGENHARIA ELETRÔNICA E DA INFORMÁTICA NO SÉCULO XX

- 1946** Eniac, primeiro computador eletrônico, começa a operar, nos Estados Unidos.
- 1947** A Bell Company começa a substituição das válvulas elétrica pelos transistores.
- 1950** Alan Turing lança os princípios filosóficos da inteligência artificial.
- 1953/1957** Desenvolvimento das redes de transmissão de dados entre computadores.
- 1964** Começa a era dos microcomputadores, para uso profissional no trabalho.
- 1969** Início da Arpanet, primeira rede de comunicação por meio de computadores.
- 1970** A Intel introduz os primeiros circuitos integrados (chips).
- 1974** Começa a era dos PCs (computadores pessoais), com o Altair 8080.
- 1975** Bill Gates e Paul Allen começam a criar programas para computadores pessoais.
- 1976** Steve Jobs e Steve Wosniak, visando ao mercado não corporativo, criam a Apple.
- 1981** O IBM PC começa a era da computação pessoal profissional.
- 1982** Início da internet, protocolo de ligação entre as redes de computadores.
- 1984** Lançamento da ideia do computador como bem de consumo para uso doméstico, com o Apple Macintosh.
- 1991/1992** Tim Berners-Lee coordena a criação da world wide web, ambiente para compartilhamento de documentos multimídia via internet.

Fonte: Francisco Rüdiger (2013)

Tim-Berners Lee é um físico e cientista britânico da computação, conhecido no mundo todo pela criação da World Wide Web, conhecida atualmente como rede mundial de computadores. O professor, que atualmente está vinculado ao Massachusetts Institute of Technology (MIT), coleciona, em seu currículo, diversos prêmios, incluindo menções honrosas e o prêmio Medalha Real, mais conhecido como "A Medalha da Rainha", uma medalha de prata concedida pela Família Real a figuras britânicas que contribuem com o desenvolvimento da ciência e do conhecimento intelectual. A medalha é entregue desde 1826 e grandes nomes ligados a áreas como física, astronomia, química, anatomia, botânica, engenharia, medicina, matemática e, mais recentemente, neurociência, ciência da computação, informática, genética, entre outros. Outro prêmio que Tim-Berners Lee recebeu, em 2016, o Prêmio Turing, considerado como o Prêmio Nobel da Computação, por inventar a World Wide Web, o primeiro navegador para a internet e os protocolos e algoritmos fundamentais que permitiram a escalabilidade da rede.

É no terreno da internet que a cibernética, finalmente, começa a potencializar-se como uma nova forma cultural. Se antes a cibernética via-se intimamente ligada aos maquinismos, autômatos e dispositivos, agora, com os avanços da internet e a socialização dessa tecnologia, a cibernética passou a assumir um novo papel na vida das pessoas. Surge, simultaneamente, o

conceito de cibercultura: quando os conhecimentos sobre informática e computação passam a ser comercializados de forma aberta para a sociedade, saindo dos muros cercados dos institutos de pesquisa e das paredes blindadas das organizações governamentais e militares. A popularização da internet, nos anos 1990, fez surgir um novo modelo de sociedade. Há quem denomine o fenômeno como "a sociedade em rede"²⁷ e outros que indicam o nascimento de uma nova cultura, baseada nas redes informáticas, que mudou o comportamento e a atuação em sociedade, chamada de "cibercultura"²⁸. Entretanto, nem mesmo os pensadores considerados clássicos na literatura intelectual sobre o fenômeno foram capazes de quantificar a velocidade vertiginosa com que a internet adentrou nas casas, empresas e instituições ao redor do mundo.

Se a cibernética avançou de forma irremediável no campo da engenharia, da eletricidade, das máquinas, dos computadores (e do imaginário também, com o surgimento de ficções científicas e ideias de pós-humanismo²⁹), a internet deu as bases para a digitalização dos saberes, das relações, da economia, da política e da cultura, inaugurando uma nova era: a era da informação e da comunicação. Para Mortensen (1980), a era da informação lança sobre a sociedade uma forma suplementar para a comunicação (elementar) humana: aquela da língua falada, escrita ou impressa. O autor recorre à teoria de Claude Shannon (que já vimos anteriormente e vamos aprofundar no próximo tópico) para exemplificar como a teoria matemática da comunicação revolucionou a forma como um ser humano pode transmitir ideias, mensagens e símbolos à medida em que adquire a capacidade de manipular códigos matemáticos.

Recorremos, novamente, à teoria de Claude Shannon para aprofundar a reflexão e a compreensão a respeito de como a teoria da informação nasce e se desenvolve até chegar na contemporaneidade, operando com sistemas complexos de automação completamente conectados e operando por meio da internet. Essas teorizações são necessárias para embasarmos nossa discussão acerca do modelo matemático que construiremos para o jornalismo, uma vez que entendemos que essa teoria emerge no núcleo de uma sociedade extremamente cruzada pelas marcas da digitalização do processo da comunicação, pela *datificação* do sujeito e pela ubiquidade dos sistemas de inteligência artificial. O exercício de recorrer a teorias clássicas nos

²⁷ Livro do sociólogo espanhol Manuel Castells, publicado em 1999, no qual o autor busca esclarecer a dinâmica econômica e social da nova era da informação.

²⁸ Livro do filósofo e pesquisador francês Pierre Lévy, publicado em 1997, no qual o autor discute os impactos da tecnologia na sociedade, cultura, economia e política, além de refletir sobre conceitos como real *versus* virtual, virtualização dos saberes, universalidade *versus* totalidade, e tece críticas ao fenômeno originado pelo que o autor chama de "cibercultura".

²⁹ Livro do sociólogo brasileiro Francisco Rüdiger, "Cibercultura e Pós-Humanismo", no qual o autor reflete, em nove ensaios, sobre a descentralização do humanismo em prol de uma sociedade onde o humano é superado no cenário da cibercultura.

auxilia a compreender a base na qual se erguem as fortalezas da tecnologia e os impérios intelectuais que deram origem à sociedade pós-humanista na qual vivemos.

4 THE INFORMATION AGE E AS CONTRIBUIÇÕES DE CLAUDE SHANNON

Claude Shannon, como citamos acima, é considerado o "pai da teoria da informação" em decorrência da publicação de seu trabalho acerca da teoria matemática da comunicação. Segundo Isaacson (2014),

Houve outro avanço teórico seminal em 1937, semelhante ao de Turing no fato de também ser um experimento puramente mental. Trata-se do trabalho de um estudante de pós-graduação do MIT chamado Claude Shannon, que entregou naquele ano a dissertação de mestrado mais influente de todos os tempos, um artigo que a *Scientific American* mais tarde chamou de “a Constituição da Era da Informação. (ISAACSON, 2014, p. 59).

Destacamos a citação de Isaacson (2014) na qual ele celebra o trabalho de Shannon, considerando-o como *o mais influente de todos os tempos* e que deu origem à era da informação. Vamos abordar em profundidade a obra de Shannon nas próximas páginas. Mas, antes de falarmos especificamente sobre sua dissertação de mestrado, vamos apresentar, brevemente, algumas informações sobre a vida de Shannon que são importantes, inclusive, para compreendermos como foi realizado o percurso teórico do pesquisador até a publicação do seu mais aclamado trabalho.

Shannon nasceu nos EUA, em 1916, e cresceu em uma pequena cidade de Michigan. Na infância e, em parte de sua adolescência, construiu alguns aeromodelos e equipamentos de radioamadorismo. Entretanto, Soni & Goodman (2017) contam que, se compararmos à história de outros grandes gênios como Beethoven, John Stuart Mill ou Norbert Wiener, por exemplo, que apresentaram sua genialidade desde muito novos (e isso se deve, também, aos incentivos que recebiam, vale ressaltar), a infância de Shannon pode ser considerada normal, regular. Segundo os autores,

Comparada à infância destes gênios, a de Shannon era comum. Não havia, por exemplo, nenhuma indicação, nos primeiros anos de Claude, de pressão paternal dominadora e, se ele mostrasse quaisquer sinais precocemente, eles não foram memoráveis o suficiente para terem sido escritos ou notados pela imprensa local. (SONI & GOODMAN, 2017, p. 220 - tradução nossa³⁰).

³⁰ Trecho original: "Compared to those childhoods, Shannon's was ordinary. There was, for instance, no indication in Claude's earliest years of overbearing parental pressure, and if he showed any signs of early precocity, they were not memorable enough to have been written down or noted in the local press" (SONI & GOODMAN, 2017, p. 220).

A menção à imprensa local chama a atenção porque, segundo os autores, na cidade onde Shannon cresceu, Gaylord, no Michigan, absolutamente qualquer evento poderia virar capa do jornal. "Gaylord era o tipo de lugar em que praticamente qualquer evento era digno de notícia" (SONI & GOODMAN, 2017, p. 206 - tradução nossa³¹). Os autores, inclusive, destacam algumas das notícias da época: como quando uma garota matou um lobo com uma vassoura, ou quando alguns cidadãos se reuniram para falar sobre alcachofras no centro da cidade.

Na vida escolar, durante a adolescência, Shannon já demonstrava uma grande aptidão para a matemática e lógica: "ele amava a ciência e não gostava de fatos. Ou melhor, não gostava do tipo de fatos que não conseguia submeter a uma regra e abstrair. A química em particular testou sua paciência." (SONI & GOODMAN, 2017, p. 240 - tradução nossa³²). Formou-se em Engenharia Elétrica, anos depois, pela Universidade de Michigan e "no seu último ano, ele respondeu a um anúncio de emprego no quadro de avisos da universidade que oferecia uma vaga no MIT para trabalhar com Vannevar Bush e ajudar a operar o Analisador Diferencial" (ISAACSON, 2014, p. 59). Na época, o jovem pesquisador conseguiu o emprego e consta que ficou maravilhado pela máquina "não tanto pelas varetas, polias e rodas que formavam os componentes analógicos quanto pelos relés eletromagnéticos que operavam como interruptores e que eram parte de seu circuito de controle (idem).

Observando a máquina, Shannon começou a entender seu funcionamento e foi anotando algumas de suas ideias. Ele percebeu que, à medida que os sinais elétricos faziam os relés se abrirem e fecharem, os interruptores da máquina criavam uma sequência de padrões de circuito. No verão de 1937, o pesquisador foi trabalhar nos Laboratórios Bell (local onde trabalhou por mais de 30 anos).

Localizado em Manhattan à beira do rio Hudson, no Greenwich Village, o laboratório era um refúgio onde ideias eram transformadas em invenções. Lá, teorias abstratas se encontravam com problemas práticos, e nos corredores e nas lanchonetes teóricos excêntricos se misturavam com engenheiros que botavam a mão na massa, mecânicos de dedos nodosos e administradores do tipo solucionadores de problemas, incentivando a fertilização da teoria pela engenharia e vice-versa. Isso tornava os Laboratórios Bell o arquétipo de uma das bases mais importantes da inovação na era digital, aquilo que o historiador da ciência Peter Galison, de Harvard, chamou de "zona de intercâmbio". Quando esses profissionais tão diferentes se reuniam, eles aprendiam a encontrar uma linguagem comum para trocar ideias e informações. Nos Laboratórios Bell, Shannon viu de perto a complexidade dos circuitos de sistemas telefônicos, que usavam comutadores elétricos para direcionar as chamadas e equilibrar cargas. (ISAACSON, 2014, p. 59-60)

³¹ Trecho original: "Gaylord was the kind of place in which just about any event was newsworthy" (SONI & GOODMAN, 2017, p. 206)

³² Trecho original: "He loved science and disliked facts. Or rather, he disliked the kind of facts that he couldn't bring under a rule and abstract his way out of. Chemistry in particular tested his patience" (SONI & GOODMAN, 2017, p. 240)

No local, além de ter contato com os mais diversos pesquisadores e com projetos de pesquisa de alto interesse, após compreender como os circuitos telefônicos funcionavam, Shannon começou a criar conexões (mentalmente) entre o funcionamento daqueles circuitos com a álgebra Booliana (outro tema que o fascinava e que havia estudado recentemente). Vale apresentarmos, de forma breve, o que é a álgebra Booleana, pois ela ajuda a explicar como Shannon chegou às suas conclusões sobre os circuitos do Analisador Diferencial. De forma bastante resumida, a álgebra Booliana foi formulada pelo matemático britânico George Boole, que descobriu que era possível expressar afirmações lógicas utilizando símbolos algébricos e equações. Shannon finalmente descobria o *insight* que o levaria a fundamentar todo o seu pensamento sobre os circuitos elétricos.

Empregando um sistema binário, ele deu a proposições verdadeiras o valor 1 e proposições falsas o valor 0. Um conjunto de operações lógicas básicas — tais como e, ou, não, ou/ou e se/então — podia então ser realizado a partir dessas proposições, como se se tratasse de equações matemáticas. Shannon imaginou que circuitos elétricos podiam realizar essas operações lógicas de álgebra booliana usando-se uma combinação de interruptores. (ISAACSON, 2014, p. 60)

Pouco tempo depois, o pesquisador voltou ao MIT, conversou com seu orientador, que sugeriu que Shannon incluísse aquelas observações em sua dissertação de mestrado.

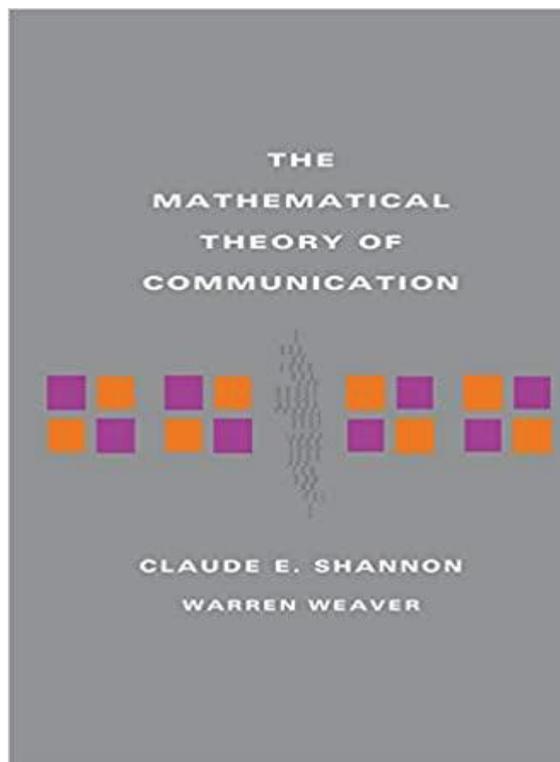
Intitulada “*A Symbolic Analysis of Relay and Switching Circuits*” [Uma análise simbólica do relé e da comutação de circuitos], ela demonstrava como cada uma das muitas funções da álgebra de Boole podia ser executada. “É possível realizar operações matemáticas complexas usando circuitos de relés”, ele resumia ao fim do trabalho. Esse se tornou o conceito básico por trás de todos os computadores digitais. As ideias de Shannon intrigaram Turing por estarem relacionadas de perto ao conceito que ele mesmo havia acabado de publicar de uma máquina universal que pudesse usar instruções simples, expressas em código binário, para resolver problemas não apenas de matemática, mas de lógica. Além disso, uma vez que a lógica se relacionava com o modo como os seres humanos pensam, uma máquina que desempenhasse tarefas lógicas podia, em teoria, imitar o modo de pensamento humano. (ISAACSON, 2014, p. 60).

O trabalho foi apresentado em 1940 e Shannon continuou estudando a ciência dos circuitos até publicar seu artigo sobre a teoria da comunicação, alguns anos mais tarde. O primeiro registro público dessa teoria na comunidade científica data de 1948, momento em que Shannon publica seu artigo *A Mathematical Theory of Communication* (Uma teoria matemática de comunicação) no *Bell System Technical Journal*, revista científica ligada ao *Nokia Bell Labs*, fundada em 1922 e publicada sob esse nome até 1983. Vale ressaltar aqui que, para reconstruir a história de Shannon, além dos livros **Os Inovadores**, de Walter Isaacson (2014) e **A Mind at**

Play: How Claude Shannon invented the digital age, dos biógrafos Jimmy Soni e Rod Goodman (2017), também buscamos ter contato com a obra original do autor.

Para tanto, consultamos o livro publicado em 1963, em coautoria com Warren Weaver, no qual Shannon revisa sua obra original e avança em suas discussões acerca da teoria matemática da comunicação, adicionando (e aprofundando) algumas reflexões acerca do conceito de fonte de informação, ruído/entropia da comunicação, probabilidade dentro de um sistema comunicacional, liberdade de escolha, e apresenta alguns teoremas que explicam suas proposições em termos matemáticos. Além do livro em coautoria com Warren, para este capítulo da tese, utilizamos como referência o ensaio publicado apenas por Warren Weaver no livro **Communication Theory**, organizado por C. David Mortensen, professor emérito da Universidade de Wisconsin (EUA), no qual o autor expõe de forma didática e resumida os princípios da teoria matemática da comunicação, de Shannon.

FIGURA 5 – CAPA DO LIVRO THE MATHEMATICAL THEORY OF COMMUNICATION, PUBLICADO EM 1963, POR CLAUDE SHANNON EM COAUTORIA COM WARREN WEAVER.



Fonte: Shannon & Weaver (1963)

Para compreender como a teoria matemática da comunicação, proposta por Shannon e ampliada após a publicação em parceria com Warren (Figura 5) funciona, recorreremos primeiramente à compreensão do que os autores consideram como "comunicação":

A palavra comunicação será usada aqui em um sentido amplo para incluir todos os procedimentos pelos quais uma mente pode afetar outra. Isso, é claro, envolve não apenas a fala escrita e oral, mas também a música, as artes pictóricas, o teatro, o balé e, de fato, todo o comportamento humano. Em algumas conexões, pode ser desejável usar uma definição ainda mais ampla de comunicação, a saber, uma que inclua os procedimentos por meio dos quais um mecanismo (digamos, equipamento automático para rastrear um avião e calcular suas prováveis posições futuras) afeta outro mecanismo (digamos um míssil guiado perseguindo este avião)³³. (SHANNON & WEAVER, 1963, p. 02-03).

Neste caso, o conceito de comunicação, cuja abordagem é central na teoria matemática da comunicação, dá conta de todo processo que envolve a emissão de um código através de um sinal de um ponto a outro, sendo decodificado no final do processo. Entende-se, portanto, a comunicação de forma ampla, mas que sempre envolve o processo de transmissão de algo para alguma outra coisa e, nesse processo, em algum grau "algo" afeta o outro "algo". Os autores consideram este um modelo ideal tanto para a comunicação humana, quanto para a comunicação entre máquinas, como veremos na sequência. Outro conceito central na teoria de Shannon é o de *informação*³⁴ e o autor aponta para um possível (e quase que inevitável) equívoco na interpretação do termo, indicando que a informação não deve ser confundida com "significado", sentido mais comum da palavra. Para o autor, informação não tem tanta relação com *o que você diz*, mas sim com o que você *poderia* dizer. Essa distinção é necessária porque Shannon utiliza o conceito de informação como uma "medida da liberdade de escolha da pessoa ao selecionar uma mensagem"³⁵ (SHANNON & WEAVER, 1963, p. 08).

Voltando à discussão introdutória que dá origem à teoria matemática de Shannon, primeiro vamos compreender como o conceito de comunicação se comporta dentro dos

³³ Texto original: "word communication will be used here in a very broad sense to include all of the procedures by which one mind may affect another. This, of course, involves not only written and oral speech, but also music, the pictorial arts, the theatre, the ballet, and in fact all human behavior. In some connections it may be desirable to use a still broader definition of communication, namely, one which would include the procedures by means of which one mechanism (say automatic equipment to track an airplane and to compute its probable future positions) affects another mechanism (say a guided missile chasing this airplane)." (SHANNON & WEAVER, 1963, p. 02-03)

³⁴ Texto original: "The word information, in this theory, is used in a special sense that must not be confused with its ordinary usage. In particular, information must not be confused with meaning". (SHANNON & WEAVER, 1963, p. 08)

³⁵ Texto original: "This word information in communication theory relates not so much to what you do say, as to what you could say. That is, information is a measure of one's freedom choice when one selects a message". (SHANNON & WEAVER, 1963, p. 08).

preceitos dessa teoria. Para os autores, um aspecto relevante no conceito de comunicação é a reflexão sobre as intercorrências que podem acontecer durante o processo comunicacional. Se pensarmos nas intercorrências dentro de um processo comunicacional, hoje, nos parecem óbvias as formulações de Shannon. Se colocarmos esse tema dentro de uma perspectiva contemporânea, é possível pensar que "evidentemente existem problemas na comunicação, por exemplo, se estou numa videochamada e ocorrem problemas na internet, é uma intercorrência bastante comum, certo?". De fato, podemos pensar que esse tipo de intercorrência é muito comum se analisarmos à luz do século XXI.

Porém, à época, a grande inovação foi justamente sistematizar essas possíveis falhas na comunicação de forma científica, apresentando como esse processo ocorre. Essas intercorrências foram divididas pelos autores, a fim de identificar problemas que podem afetar a comunicação dentro daquele sistema proposto. Além disso, os autores enfatizam o fato de que pesquisar aspectos sobre a comunicação é um grande desafio, imposto justamente pelo fato de que a comunicação é uma área ampla, sendo indispensável a conceituação que apresentamos acima, pois ela serve sobremaneira para delimitar *o que* os autores consideram como *comunicação* (e também como *informação*). Delimitar conceitos na ciência é uma tarefa de extrema relevância, por vários motivos, entre eles, para evitar ambiguidades e equívocos, para delimitar as fronteiras dos temas abordados, para estabelecer critérios de análise e para classificar, de maneira semântica, as perspectivas do pesquisador.

Nesse sentido, em relação às intercorrências que podem ser identificadas no processo comunicacional, Shannon & Weaver ressaltam que a comunicação, por si só, evoca, a priori, três problemas principais: o primeiro, de ordem técnica; o segundo, de ordem semântica; e o terceiro, de eficiência³⁶. Os autores dividem esses problemas em níveis (A, B e C³⁷), utilizando-se de perguntas para conduzir seu raciocínio:

O nível A corresponde ao seguinte questionamento: com que precisão os símbolos de comunicação podem ser transmitidos?

No nível B, a pergunta que dá origem ao problema semântico corresponde ao questionamento: com que precisão os símbolos transmitidos comunicam o significado desejado?

³⁶ Neste caso, a tradução do termo original "*effectiveness problem*" para o português pode conduzir o leitor a uma confusão, já que na tradução livre podemos considerar como um "problema de eficiência" e em alguns textos clássicos (MORTENSEN, 1979) o termo foi traduzido como "problema de influência".

³⁷ Texto original: "*Level A: How accurately can the symbols of communication be transmitted? Level B: How precisely do the transmitted symbols convey the desired meaning? Level C: How effectively does the received meaning affect conduct in the desired way?*" (SHANNON & WEAVER, 1963, p. 04).

No nível C, que procura compreender a eficiência da comunicação, os autores questionam: com que eficiência o significado recebido (pelo emissor) afeta a conduta (do receptor) da maneira desejada?

QUADRO 1: Apresentação dos problemas da comunicação

Nível A	Nível B	Nível C
Ordem Técnica	Ordem Semântica	Ordem de Eficiência
Diz respeito à precisão de transferência de informação do emissor para o receptor. São inerentes a todas as formas de comunicação, seja por meio de séries ou de símbolos discretos (fala escrita), por um sinal variável (transmissão por rádio ou telefone de voz ou música) ou por um padrão bidimensional variável (televisão).	Os problemas semânticos referem-se à interpretação de significado pelo receptor em relação ao significado pretendido pelo emissor. Por exemplo, se percebe-se que Sr. X não entende o que o Sr. Y diz, então é possível que, não se permitindo ao Sr. Y mais nada a não ser continuar a conversa, esta situação jamais venha a se esclarecer. Se o Sr. Y pergunta: "Agora você já entende o que eu digo?", ao que o Sr. X responde: "Claro que sim!", isto não é necessariamente uma certeza de que foi conseguido o entendimento.	Os problemas de influência ou de eficiência dizem respeito ao grau de êxito com que o significado transmitido ao receptor leva ao comportamento desejado de sua parte. À primeira vista pode parecer desagradavelmente limitado concluir que o objetivo de toda comunicação seja influenciar o comportamento do receptor. Mas, tomando-se qualquer definição razoavelmente ampla de comportamento, fica evidente que a comunicação ou influencia ou não possui qualquer efeito discernível ou demonstrável.

Fonte: adaptação do texto de Weaver, 1979, p. 39-40.

Podemos inferir que os problemas de ordem técnica, então, dizem respeito ao aspecto instrumental da comunicação e são influenciados de forma direta pelo meio, enquanto os outros dois níveis fazem alusão aos aspectos interpretativos do processo comunicacional, sendo o problema de ordem semântica o processo de interpretação da mensagem, e o problema de ordem de eficiência o objetivo comunicacional sob o ponto de vista da emissão.

Segundo os autores:

Matematicamente, a primeira envolve a transmissão de um conjunto finito de símbolos discretos, a segunda, a transmissão de uma função contínua do tempo e a terceira, a transmissão de muitas funções contínuas do tempo ou de uma função contínua do tempo e de duas coordenadas espaciais³⁸. (tradução nossa, SHANNON & WEAVER, 1963, p. 04)

³⁸ Texto original: "Mathematically, the first involves transmission of a finite set of discrete symbols, the second the transmission of one continuous function of time, and the third the transmission of many continuous functions of time or of one continuous function of time and of two space coordinates". (SHANNON & WEAVER, 1963, p. 04)

Isso quer dizer que, sob o olhar da teoria matemática da comunicação, os autores consideram que os problemas A, B e C se relacionam mutuamente durante o processo de transmissão de símbolos (que resulta no processo de comunicação ideal, defendido na teoria original). Os problemas técnicos, ou de nível A, dizem respeito à precisão da transferência de informação do emissor para o receptor, sob uma perspectiva instrumental. Colocando esse problema dentro de um quadro teórico imaginado, a comunicação básica de ordem técnica pode ser resumida e ilustrada da seguinte forma:

FIGURA 6 – SISTEMA DE COMUNICAÇÃO BÁSICO DE ORDEM TÉCNICA



Fonte: Shannon & Weaver (1963)

Dentro dos problemas de nível A, estão incluídas todas as formas de comunicação, como descrito acima: seja por meio de símbolos escritos, como a fala escrita, ou como a fala oral por um sinal variável, como é o caso da transmissão de telefone, ou por um sinal bidimensional, como por exemplo a transmissão pela televisão. Vale ressaltar que esses são exemplos dados à época em que a teoria foi desenvolvida, entre os anos 1948 e 1963 (ano em que o livro de Shannon, em parceria com Weaver, fora publicado - e que utilizamos neste estudo).

Atualmente, destacamos que esse quadro ganha ainda maior complexidade e outros ingredientes podem ser adicionados ao processo comunicacional, já que na época ainda não existia a internet e os sistemas de automação, por exemplo. Percebam que este estudo que estamos propondo, por exemplo, pode ser considerado como uma ampliação da visão a respeito da teoria matemática da comunicação, aplicada ao jornalismo. Assim como Shannon, que percebeu uma estrutura lógica para a transmissão de uma mensagem, nesta tese, propomos uma estrutura lógica, utilizando a matemática, que corresponda ao processo jornalístico. O modelo que propomos neste estudo, e que iremos demonstrar em seguida, possui, além de inspiração em autores como Claude Shannon e Ada Lovelace, correspondência direta com as contribuições desses pesquisadores. Mas, antes de entrarmos no nosso modelo matemático para o jornalismo, voltemos aos preceitos de Shannon.

Falávamos dos problemas comunicacionais apontados por Shannon. Enquanto o primeiro acontece no âmbito da técnica, os problemas semânticos de segundo nível se referem

à interpretação da mensagem pelo receptor em relação ao significado pretendido pelo emissor. Os problemas semânticos adicionam maior complexidade ao modelo proposto por Shannon, uma vez que o significado da mensagem emitida (significado *pretendido*) pode não ser completamente compreendido pelo receptor (significado *interpretado*). Os problemas de interpretação e significado das mensagens podem variar muito, e se tomarmos como exemplo um acontecimento comum na contemporaneidade, poderemos compreender a relevância dos problemas de segundo nível. Vejamos um exemplo: a mesma mensagem, escrita numa rede social digital como o Instagram³⁹, pode ganhar dezenas, até centenas de significados, de acordo com a interpretação de cada receptor dentro do sistema. Isso porque, numa rede social digital como o Instagram, um mesmo emissor poderá falar para uma massa quase que infinita de receptores. Cada receptor, no entanto, tem suas particularidades. Um mesmo emissor pode falar com pessoas do mundo inteiro, por exemplo. A língua pode ser um problema semântico, a legenda, outro, a comunicação visual da mensagem, outro. Em relação a esse assunto, a seguir falaremos sobre os sistemas complexos (abertos e adaptativos, como é o caso das redes sociais digitais), sua interferência no processo comunicacional, atualmente, e sua relevância nos sistemas de comunicação na contemporaneidade.

Por fim, o terceiro nível (C) dos problemas da comunicação, segundo Shannon, são os problemas de ordem de eficiência da mensagem, isto é, dizem respeito ao grau de êxito com que o significado transmitido ao receptor o leva ao comportamento desejado pelo emissor. Nesse aspecto, os autores argumentam que todo processo de comunicação carrega em si uma intencionalidade e que, em primeira instância, a intenção do emissor é, em grande medida, a de influenciar na tomada de decisão do receptor. Como vimos na figura acima, os autores consideram que o objetivo de toda comunicação é, no final das contas, o de influenciar o receptor, por meio de símbolos (mensagem), já que toda mensagem carrega em si o desejo de influenciar no comportamento do outro.

Para exemplificarmos esse terceiro e último nível dos problemas comunicacionais, temos a seguinte situação: pense que você gostaria de pedir um favor a alguém, por exemplo: "Fulano, gostaria que me alcançasse o laranja". Tomemos como hipótese que você estava se referindo ao copo com suco de laranja. Neste caso, o objetivo da sua fala seria o de pedir, com a expectativa de receber em troca aquilo que pediu. No entanto, é possível que o seu interlocutor não tenha apreendido o objetivo da sua mensagem e te entregue um objeto laranja em vez do

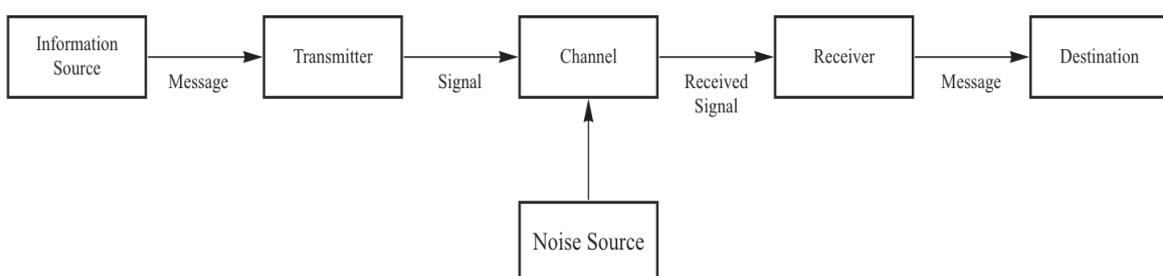
³⁹ Rede social de compartilhamento de fotos e vídeos, que teve sua origem em 2010 e é uma das redes sociais mais difundidas no mundo. No Brasil, em 2022, a rede social já tem mais de 122 milhões de usuários, segundo o Data Digital Report 2022 <<https://datareportal.com/reports/digital-2022-brazil>>

copo de suco (que era o seu objetivo). Assim, o interlocutor te entregou "o laranja", mas provavelmente o objetivo comunicativo pretendido pelo emissor não tenha sido atingido, certo?

Para Weaver (1979), "tomando-se qualquer definição razoavelmente ampla de comportamento, fica evidente que a comunicação ou influencia o comportamento ou não possui qualquer efeito discernível ou demonstrável." (in MORTENSEN, 1979, p. 40). Ou seja, todo processo comunicacional estabelece, por si próprio, uma intencionalidade. Embora não esteja sempre explícita, a intencionalidade frequentemente deverá imperar numa comunicação ideal, mesmo em mensagens nas quais há ambiguidade - como no exemplo acima (problema de ordem da eficiência).

Tendo como premissa que um processo comunicacional pode (e frequentemente irá) receber interferência de alguns problemas (níveis A, B e C, como mostramos acima), e partindo da ideia de comunicação proposta por Shannon & Weaver (1963) é possível visualizar como esta pode acontecer quando colocada dentro de um sistema de emissão e recepção de mensagens, cujo trajeto é interpelado por variáveis, como o sinal e o ruído, conceitos de grande relevância dentro da teoria matemática da comunicação. Para os autores, um sistema de comunicação ideal deve levar em consideração as possíveis intercorrências e pode ser resumido a estes elementos fundamentais: fonte de informação, emissor, canal, receptor, fonte de ruído, sinal e mensagem.

FIGURA 7 – MODELO MATEMÁTICO DA COMUNICAÇÃO

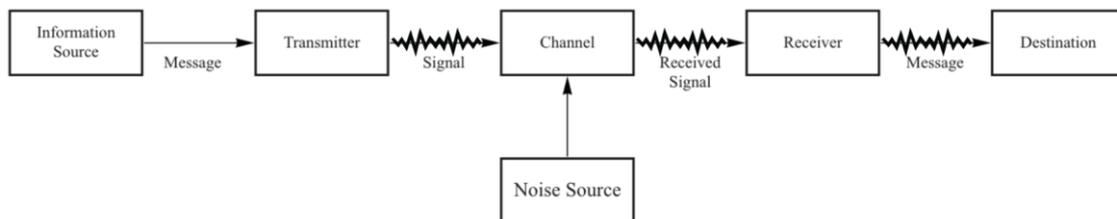


Fonte: Shannon & Weaver, 1963, p. 07)

No esquema proposto por Shannon & Weaver (1963), a fonte de informação seleciona as mensagens desejadas, enquanto o transmissor transforma a mensagem desejada em um sinal, que é enviado pelo canal de comunicação desde o transmissor até o receptor. De acordo com os autores, o receptor constitui-se como uma espécie de transmissor inverso, que tem a capacidade de transformar o sinal transmitido de volta em uma mensagem, entregando essa mensagem ao

destinatário⁴⁰. Ainda, no processo de transmissão da mensagem, é esperado que sejam adicionados aspectos que não eram pretendidos pela fonte de informação, é o caso dos ruídos no sistema comunicacional, que são adicionados durante o processo, sem a intenção da fonte, e podem modificar parcial ou completamente a recepção da mensagem. O ruído pode estabelecer problemas mais complexos ao modelo comunicacional proposto pelos autores⁴¹, já que tem o potencial de causar interferência em uma ou em todas as etapas do processo comunicacional. Por exemplo, se ocorrer uma falha no sinal entre o transmissor e o canal (Figura 8), todo o processo adiante poderá ser afetado, modificando a mensagem original e, por consequência, interferindo no resultado final da mensagem.

FIGURA 8 – MODELO MATEMÁTICO DA COMUNICAÇÃO COM INTERFERÊNCIA NA BASE



Fonte: A autora, adaptação de Shannon & Weaver, 1963.

Sendo assim, a mensagem pretendida pelo emissor pode ser impactada, resultando em outra conotação ao chegar no receptor, ou em seu destino final. Ou seja, se o emissor queria falar "X" e durante o processo de transmissão da mensagem, por algum motivo, houve uma fonte de ruído, é possível que o receptor tenha entendido "Y", logo, a comunicação foi realizada, embora seu objetivo comunicativo não tenha sido alcançado com êxito. O ruído, por sua vez, pode ocorrer em qualquer etapa do processo de transmissão de informações, podendo impactar não apenas o significado da mensagem, mas também o objetivo comunicacional, que acaba por não se completar. Por esta razão, Shannon dedicou sua concentração a compreender, de forma

⁴⁰ Texto original: "The information source selects a desired messages. The transmitter changes this message into the signal which is actually sent over the communication channel from the transmitter to the receiver. The receiver is a sort of inverse transmitter, changing the transmitted signal back into a message, and handing this message on to the destination."(Shannon & Weaver, 1963, p. 07-08).

⁴¹ Texto original: "In the process of being transmitted, it is unfortunately characteristic that certain things are added to the signal which were not intended by the information source. These unwanted additions may be distortions (...). All of these changes in the transmitted signal are called noise". (Shannon & Weaver, 1963, p. 07-08).

mais aprofundada, os impactos dos ruídos no processo de transmissão informacional, procurando evitar ou diminuir a sua incidência dentro do processo de comunicação, como veremos a seguir.

4.1 NO MUNDO DA ENTROPIA, O RUÍDO É REI

No one knows what entropy really is [...]. It has been the inability of a steam engine to do work; it has been the dissipation of heat and energy, the unalterable tendency of every part of closed system to lapse toward lukewarm muck; it has been more roughly but also more resonantly, the trend toward disorder, chaos. (Soni & Goodman, 2017, p. 254⁴²)

O ruído foi um dos conceitos centrais da teoria matemática proposta por Shannon e viria a ser conhecido, nas décadas seguintes, como uma das principais abordagens para compreender o conceito de entropia, uma medida muito presente na Ciência da Informação e que rege princípios da termodinâmica, física e estudos moleculares. É facilmente percebido que as diversas contribuições de Shannon ressoam até este momento, no qual estamos elaborando esta tese. Para Weaver (1979), o conceito de entropia já é muito conhecido entre os pesquisadores da termodinâmica e da física, mas ainda não era utilizado de maneira aprofundada no campo da comunicação e da informação. Weaver, que foi um importante matemático e coautor do livro **The Mathematical Theory of Communication** ao lado de Shannon, relata, em ensaio publicado no livro **Communication Theory**, publicado em 1979 por C. David Mortensen, que a entropia pode ser conhecida como o grau de *casualidade* ou *embaralhamento* de uma situação dentro do sistema comunicacional.

Segundo ele, "isto é expresso em termos das várias probabilidades em jogo" (WEAVER, in MORTENSEN, 1979, p. 43). O conceito de *probabilidade* é outro termo que também se faz presente na teoria matemática da comunicação, originalmente proposta por Shannon. Weaver (1979) afirma que os símbolos (palavras), por exemplo, dentro de um sistema comunicacional cuja língua oficial é o inglês podem ser organizados de acordo com uma lógica de probabilidade estatística, já que aproximadamente metade de tudo aquilo que é escrito na língua inglesa é determinado pela probabilidade da estrutura da linguagem, sendo apenas metade da mensagem uma escolha livre do emissor. Isto porque, no caso da língua inglesa,

⁴² Neste trecho, em específico, optamos por reproduzir a citação original. O motivo é que entendemos que, na língua original, a conotação da frase de Soni & Goodman ressoa com maior potência. A seguir, a tradução: "Ninguém sabe o que realmente é a entropia [...]. Tem sido a incapacidade de uma máquina a vapor realizar seu trabalho; tem sido a dissipação de calor e energia, a tendência inalterável de cada parte do sistema fechado de cair na lama morna; tem sido mais rude, mas também mais ressonante, a tendência à desordem, ao caos. (Soni & Goodman, 2017, p. 254).

metade do que se escreve pode ser considerado como uma redundância imposta justamente pela estrutura daquele idioma. Para Weaver (1979),

É evidente que a probabilidade desempenha importante papel na criação da mensagem, e que as escolhas dos símbolos sucessivos dependem de escolhas anteriores. Assim, se tomarmos a língua inglesa como exemplo, e tivermos como o último símbolo escolhido "*the*" ("o/a"), então a probabilidade de que a próxima palavra venha a ser um artigo ou uma forma verbal que não um substantivo verbal é muito pequena. Após as três palavras "*in the event*" ("no caso de"), a probabilidade de "*that*" ("que") como a próxima palavra é grande, enquanto a probabilidade de "*elephant*" é muito pequena. (WEAVER in MORTENSEN, 1979, p. 42).

Sendo assim, é importante destacar que a probabilidade é um fator essencial para compreender um sistema comunicacional, uma vez que essa variável desempenha "importante papel na criação da mensagem, e que as escolhas dos símbolos sucessivos dependem das escolhas anteriores" (WEAVER in MORTENSEN, 1979, p. 42). O conceito de probabilidade - que interfere, assim como o ruído, dentro da estrutura de um sistema comunicacional - é assunto central para compreendermos como os sistemas de automação atuais adquirem a capacidade de prever acontecimentos por meio do processamento dos dados pré-existentes, gerando notícias em tempo real, de acordo com a estrutura do software de coleta, tratamento dos dados e produção da narrativa. Isso porque a probabilidade, ou seja, as chances de determinado evento ocorrer, dadas as circunstâncias e dentro de um contexto específico, constitui grande parte da organização de um sistema autônomo atual. A respeito da probabilidade, veremos adiante, no capítulo metodológico, como seu conceito pode interferir num sistema complexo, como é o caso dos sistemas de geração de notícias automatizadas.

Segundo Weaver (1979), "Shannon demonstrou que quando as letras ou palavras escolhidas aleatoriamente são dispostas em sequências regidas apenas pelas considerações de probabilidade, há a tendência de se organizarem em palavras e orações com significado." (WEAVER in MORTENSEN, 1979, p. 43). As diversas probabilidades dentro de um sistema comunicacional podem ocasionar certo grau de casualidade, ou entropia. Carreira (2017) afirma que Shannon se "preocupava com o sistema de comunicação que suportava a troca de mensagens, com a capacidade do canal e com a correção e eliminação de ruídos que afetavam a qualidade do sinal." (CARREIRA, 2017, p. 35). Segundo a autora, a teoria de Shannon buscava, além de propor um sistema comunicacional efetivo, diminuir efetivamente o grau de desordem (incidência de ruídos) dentro do sistema.

Shannon passou a observar seu sistema de comunicação levando em consideração o fato de que a probabilidade da incidência de ruídos era tão alta que, iminentemente, o ruído

integraria parte do sistema, ou seja, "a questão, portanto, era como passar da entropia à informação" (CARREIRA, 2017, p. 36). Por isso, vale ressaltar essa que foi uma das grandes constatações do pesquisador: o ruído não só faz "parte" do processo, como ele "é", de fato, o processo. Isso significa que no mundo da entropia, o ruído reina. E, mais do que isso, no mundo da informação, a entropia é quase que uma lei.

No caso das teorias das Ciências da Informação, a entropia está ligada a determinado grau de desordem no processo de transmissão da mensagem (o que, invariavelmente, afeta o processo comunicacional, seja ele humano ou mecânico, no caso das máquinas). Segundo Carreira (2017), "é importante ressaltar que Shannon, naquela época, já assinalava que tanto a fonte como o destinatário poderiam ser máquinas" (p. 35). Essa citação merece destaque, pois, assim como Ada, em meados de 1830, falava que as máquinas poderiam ser programadas para servir a propósitos gerais (vislumbrando o computador moderno, que não apenas calcula dados, mas serve para uma infinidade de outras finalidades), Shannon, quase que 100 anos depois, também vislumbrava, com pioneirismo, o conceito de comunicação entre computadores (o ideal que foi conquistado com o advento da internet - e que inaugura uma nova fase nos avanços tecnológicos, como veremos a seguir).

De acordo com o esquema proposto por Shannon, o processo comunicacional poderia ser considerado como um sistema capaz de ser reproduzido não apenas no âmbito da comunicação oral (como o exemplo da fala, cuja fonte de informação é o cérebro do emissor, passando pela mensagem falada e pela interpretação e audição do destinatário), mas também por máquinas, via sistemas de transferência de informação. É nesse ponto que encontramos um grande avanço no que diz respeito ao modo como era realizada a comunicação entre máquinas. A teoria de Shannon ficou mundialmente conhecida justamente porque buscava resolver um problema já largamente estudado por organizações e grupos de pesquisa vinculados à universidade e, especialmente, aos setores militares.

Segundo Weaver (1979), um sistema comunicacional como o proposto por Shannon pode ser estudado de acordo com quatro matrizes: "a quantidade de informação, a capacidade do canal de comunicação, ao processo de codificação que pode ser utilizado para transformar uma mensagem em um sinal e aos efeitos dos ruídos" (in MORTENSEN, 1979, p. 41). Vale ressaltar que, para a teoria matemática da comunicação, proposta por Shannon, pouco importa o significado da mensagem, sendo mais relevante observar o funcionamento do sistema comunicacional pelas unidades probabilísticas da fonte de informação, que serão determinantes para estabelecer a possibilidade de troca de mensagens dentro do sistema.

A natureza estatística da fonte de informação é o que irá determinar a capacidade de um canal de comunicação, como aponta Weaver (1979), quando defende que a quantidade de comunicação de um canal só pode ser mensurada a partir da quantidade de informação que está disponível na fonte. É neste ponto da teoria de Shannon que encontramos convergência com o que estamos propondo neste estudo, pois, após constatar que a fonte de informação é o ponto mais importante dentro do sistema comunicativo, o pesquisador inicia uma primeira aproximação entre o desenho conceitual do sistema comunicativo e o teorema básico que reduz o sistema a uma equação.

O emissor pode tomar uma mensagem escrita e utilizar algum código para cifrá-la em, digamos, uma sequência de números que, por sua vez, seriam enviados pelo canal como o sinal. É por isso que se costuma dizer que, em geral, a função do emissor é codificar e a do receptor decodificar a mensagem. (WEAVER, in MORTENSEN, 1979, p. 44)

A teoria matemática, enfim, começava a ganhar forma após as constatações de que o emissor teria a responsabilidade de codificar uma mensagem e o receptor, de decodificá-la. Nesse sentido, a teoria também abrangeria emissores e receptores considerados bastante sofisticados para a época, como por exemplo, aqueles que possuíam "memórias", de forma que a "maneira como codificam um determinado símbolo da mensagem não depende exclusivamente deste único símbolo, mas também de símbolos anteriores da mensagem e da forma como foram codificados." (WEAVER, in MORTENSEN, 1979, p. 44).

Essas constatações levaram os autores à formulação de um primeiro teorema a fim de desdobrar sua teoria em uma estrutura lógica. O objetivo era determinar um teorema básico que representasse um canal transmissor de símbolos sem a incidência de ruídos. Segundo Weaver (1979), "esse teorema é relativo a um canal de comunicação que possui capacidade de C unidades por segundo, recebendo sinais de uma fonte de informação de H unidades por segundos." (in MORTENSEN, 1979, p. 44). A ideia do teorema é demonstrar que, ao imaginar-se processos codificadores adequados para o emissor, seria possível emitir símbolos (a mensagem pretendida) pelo canal a uma razão média aproximadamente igual a C/H . O teorema trata, resumidamente, da intenção de extinguir os ruídos dentro do processo. Neste caso, Shannon & Weaver (1979) pretendem conceder um teorema básico, dentro de um cenário ideal e não modificado, onde o ruído seria inexistente. Entretanto, os próprios pesquisadores reconhecem que o teorema em sua constituição inicial é insuficiente para dar conta de um sistema comunicacional que se complexifica quando levamos em consideração a possibilidade de ruídos em alguma (ou em várias) etapa do processo comunicacional.

Sendo assim, num sistema que opera com a presença de ruídos, é importante equiparar a capacidade do canal à quantidade de entropia para tornar o resultado da mensagem confiável. Os autores explicam que o "grande teorema" da teoria matemática da comunicação é justamente este, criado a partir da hipótese que leva em consideração não apenas um sistema de comunicação básico, fechado e determinante (como o apresentado acima), mas que trate também da interferência da entropia (desordem) no sistema. O teorema pode ser explicado da seguinte forma:

Suponhamos que um canal ruidoso de capacidade C esteja recebendo informação de uma fonte de entropia H , entropia esta correspondente ao número de mensagens possíveis a partir da fonte. Se a capacidade C do canal for igual ou maior que H , então, pela utilização de sistemas codificadores adequados, o output da fonte pode ser emitido pelo canal com o mínimo de erro que se desejar. Porém, se a capacidade C do canal for menor que H (a entropia da fonte), então torna-se impossível escolher códigos que reduzam a frequência do erro tanto quanto se deseje. (WEAVER in MORTENSEN, 1979, p. 46).

Segundo Weaver, esse teorema tem a capacidade de oferecer uma descrição precisa (ainda que simples) da "extrema confiança que se pode sempre depositar em um canal de comunicação que opera com a interferência de ruído", pois, tendo calculado a entropia de uma determinada fonte de informação, é possível compará-la ao valor máximo que esta pode ter (sujeito apenas ao fato de que a fonte continue a operar de acordo com as mesmas condições da entropia calculada). Nesse sentido, os autores consideram que o máximo de entropia é facilmente calculado pelo máximo de capacidade de transmissão do canal, sendo representado pelas seguintes fórmulas:

QUADRO 2: Esquema matemático da capacidade de entropia em sistemas

Redução de Entropia	Aumento de Entropia
$C \geq H$ logo, menor grau de entropia.	$C < H$, logo, maior grau de entropia.

Fonte: Teorema geral da entropia, de Shannon & Weaver (1963)

Essas representações podem indicar o valor da entropia dentro de um sistema e a razão entre a entropia real e a entropia máxima, dentro deste esquema, é denominada por entropia relativa da fonte. De acordo com Weaver (1979), considerando que a entropia relativa de uma determinada fonte de informação seja de oito décimos "isto vem a determinar, grosso modo, que essa fonte é, em sua escolha de símbolos para formar uma mensagem, aproximadamente 80% livre do que poderia possivelmente ser com estes mesmos símbolos" (WEAVER in

MORTENSEN, 1979, p. 46). Para exemplificar o teorema e tornar mais compreensível o raciocínio dessa proposição, podemos tomar como exemplo a estrutura da língua inglesa, como citamos acima, cuja redundância é de 50%, ou seja, apenas metade das palavras que os falantes da língua inglesa escolhem para comunicarem-se são determinadas pela "livre escolha" dos sujeitos, já que a outra metade delas é "na realidade controlada pela estrutura estatística da língua, embora não estejamos habitualmente cientes deste fato" (idem).

Todos os conceitos apresentados neste capítulo são essenciais para desenvolver também os próximos tópicos deste estudo, especialmente porque a teoria matemática da comunicação, criada por Shannon, inaugura um período de grandes avanços não apenas no que tange aos estudos de estatística e probabilidade aplicados a sistemas de transmissão, mas, em especial, nos estudos sobre cibernética, computação e sistemas digitais. O que podemos concluir é que Shannon concedeu alicerces para o desenvolvimento de diversas áreas, estabelecendo bases sólidas sobre o pensamento de sistemas de comunicação para que outros pesquisadores pudessem avançar em suas conceituações teórico-metodológicas. É o caso da teoria geral dos sistemas complexos, que embora "especialistas no tema colocam a história da teoria dos sistemas com início entre os Sumérios na Mesopotâmia (anterior a 2500 a. C.)", traz relevantes contribuições à sombra da teoria matemática da comunicação, já que mesmo com certo grau de entropia, Shannon considerava o sistema de comunicação como um sistema fechado. A teoria geral dos sistemas, segundo Bertocchi (2016), nos auxilia a compreender melhor o funcionamento de sistemas complexos em constante estado de adaptação, mutabilidade e interferências externas, como é o caso dos softwares de elaboração de narrativas digitais (objeto de estudo de Bertocchi) e dos complexos sistemas de automação de notícias (Carreira, 2017).

Para Carreira (2017), a contribuição da teoria de Shannon oferece uma abordagem pertinente para reflexões sobre o jornalismo contemporâneo, especialmente para compreender como sistemas de automação de notícias operam, codificando e decodificando símbolos (mensagens), com o objetivo de reduzir ao máximo a entropia, gerando notícias em uma fração de segundo. Em seu estudo, Carreira (2017) relata que

a teoria matemática da comunicação e a cibernética oferecem abordagens pertinentes ao estudo dos sistemas automatizados. A abordagem de Shannon, de forma resumida, nos interessa pelo fato de que os algoritmos empregados na automatização da notícia trabalham a partir de dados codificados, seguem processos lógicos e estatísticos, procuram extrair ordem dos dados (informação), e evitar ruídos. Com a cibernética, entendemos a automatização a partir dos fluxos de dados que alimentam constantemente os sistemas (inputs) para gerar feedbacks (outputs). (p. 42).

Além de Carreira (2017), Bertocchi (2016) também considera relevantes os legados da teoria matemática da comunicação e da cibernética, especialmente no que tange aos formatos narrativos digitais que performam o jornalismo na contemporaneidade. Para a autora,

quando defendemos que a narrativa é sistema, estamos recorrendo à moderna teoria dos sistemas – ou pensamento sistêmico holístico – para estruturar e sustentar tal entendimento. Isso ocorre, em grande medida, porque necessitamos de uma perspectiva que nos ajude a pensar em fenômenos complexos em constante estado de adaptação, incerteza e abertura ao meio externo – assim como o fenômeno da narrativa digital. (BERTOCCHI, 2016, p. 40).

Interessa-nos, aqui, estabelecer essa linha de raciocínio que desponta com a teoria matemática de Shannon, passando pela história da cibernética até chegar à modernidade, com a complexificação dos sistemas de comunicação e transmissão de dados. Isso porque, e somente assim, poderemos explicar como um software de automação de notícias consegue, a partir de uma fonte de informação rica em dados, criar narrativas jornalísticas digitais tão apuradas que chegam a superar a acurácia de um jornalista humano. Não há, sob o nosso ponto de vista, como separar esses fenômenos, já que a História os fez despontar juntos, em convergência, sob a luz do desenvolvimento tecnológico em prol de um ideal tecnicista. Ideal este cujos objetivos principais giram em torno da eficiência e da produtividade, nem que isso cause a obsolescência do humano durante o processo.

Trouxemos, até aqui, a história da teoria matemática da comunicação porque alguns de seus conceitos, como os problemas de comunicação (técnicos, semânticos e de eficiência), a entropia e o conceito de probabilidade serão agregados ao nosso modelo matemático para o jornalismo. Mas, para entendermos sobre esse modelo, gostaríamos de realizar uma incursão, ainda, em outras vertentes de pensamento. A seguir, falaremos sobre a teoria dos sistemas complexos, para, então, adentrar na filosofia matemática de Russell. Juntas, as teorias aqui apresentadas serão alicerces para o proposto nesta tese: criar um modelo matemático capaz de dar conta do processo jornalístico.

Entretanto, não podemos falar de processo jornalístico sem antes compreender o que é o jornalismo, como ele se desenvolve ao longo dos anos e, principalmente, qual é o contexto do jornalismo nos dias atuais: permeado pelos sistemas de automação, no que Carreira (2017) considera como a quinta fase do jornalismo (Cap. 4). Destacamos que é sob essa perspectiva que criaremos o modelo matemático para o jornalismo: levando em consideração o funcionamento dos sistemas de automação de notícias, pois essa é a nossa realidade atual. E estes sistemas nada mais são do que o que veremos a seguir.

4. 2 SISTEMAS COMPLEXOS: DE SHANNON A BERTOCCHI

A teoria dos sistemas, segundo Bertocchi (2016), tem seu início há muito tempo (antes de 2.500 a.C.), na Mesopotâmia Antiga, entre os Sumérios, e foi evoluindo mediante ao desenvolvimento da própria sociedade. Em quase cinco mil anos de existência, no entanto, a noção de sistema foi usada quase que de forma intuitiva, tendo sua gênese documentada apenas nos anos 1940, quando os EUA conduziam sua participação na Guerra Mundial. Segundo a autora, para viabilizar os esforços de guerra entre dois oceanos, era necessário formalizar determinados procedimentos altamente perpassados pelos conceitos matemáticos para "prever" o futuro, por meio da interceptação de mensagens e comunicação à distância com suas bases militares e políticas. Assim, à época, os EUA buscavam que os sistemas de comunicação pudessem, com o auxílio de uma equação matemática, descrever o funcionamento de estruturas inimigas ou que se comunicassem com outros sistemas aliados⁴³.

Do mesmo modo, Araújo & Gouveia (2016) relatam que a noção de sistema, de fato, não é uma descoberta recente, e que há muito "os únicos instrumentos úteis à disposição eram os velhos preceitos de René Descartes: dividir e examinar o problema por parcelas e começar metodicamente pelas menos complexas, ascendendo, por partes menores, até as questões mais complexas (p. 02). Os autores relatam que a abordagem cartesiana foi facilmente aplicada em diversas áreas consagradas cientificamente como a biologia e a física, por exemplo. Entretanto, a abordagem cartesiana tornou-se insuficiente "quando na Física se descobriu partícula menor que o átomo — os quarks — que eram difíceis de descrever, o que trouxe uma crise epistemológica à forma como a ciência estuda e se desenvolve" (idem). Nesse momento é que a abordagem clássica cartesiana se mostrou falha, tornando-se necessária a invenção de uma nova abordagem que desse conta de explicar como partes (ou partículas) menores poderiam encaixar-se dentro de um sistema. Segundo os autores, "o paradigma da visão do sistema como um todo indivisível, com relacionamentos interdependentes, conhecido por Abordagem Sistêmica foi a solução do dilema, e gradativamente substituiu a abordagem cartesiana." (ARAÚJO & GOUVEIA, 2016, p. 03).

O novo paradigma sistêmico trouxe um desafio à ciência: olhar para todas as partes de um determinado contexto para compreender como funcionam seus elementos individualmente

⁴³ Ressaltamos que os esforços militares têm grande impacto nos avanços tecnológicos. Para aprofundar o entendimento sobre a participação dos militares especialmente nos períodos entre Guerras, na tecnologia, sugerimos a leitura do capítulo sobre "Programação" do livro "Os Inovadores" de Walter Isaacson (2014) e o capítulo "O Ocidente e a Técnica: estágios reflexivos do pensamento tecnológico" do livro "Cibercultura" de Francisco Rüdiger (2011). Além disso, vale também assistir ao filme "O Jogo da Imitação" que conta a história de Alan Turing e demonstra seu trabalho durante o período da Segunda Guerra Mundial.

e em interação com o meio circundante. Segundo Araújo & Gouveia (2016) enquanto que a ciência tradicional pautada na abordagem clássica focalizava o "funcionamento isolado dos diversos sistemas que formam o sistema em estudo, a abordagem sistêmica aborda a interação desses sistemas menores entre eles e com o funcionamento do sistema em estudo como um todo." (p. 03). Sendo assim, passou-se a estudar a relação entre cada elemento que compõe um sistema, concentrando-se nas suas interações com outros sistemas individuais e com o sistema como um todo. Aqui, vale ressaltar que a palavra sistema adquire um significado amplo, podendo ser aplicada, por exemplo, a um sistema ligado à sociologia, como as dinâmicas dentro de um grupo social X e as características de cada membro em relação tanto ao sistema, quanto aos membros que o compõem; ou a um sistema informático de geração de notícias, podendo ser aplicado ao sistema em si, mas também aos comandos, etapas e agentes envolvidos no processo de funcionamento desse sistema; ou, ainda, a um sistema biológico, como a identificação de uma célula e de todos os elementos que a compõem. Por fim, o sistema pode ser compreendido como um conjunto de elementos, concretos ou abstratos, organizados.

Basta fazer uma rápida pesquisa aos dicionários (online e físicos) para identificarmos que estamos falando de uma palavra cujo significado é amplo, variável e pode ser interpretado sob diversas lentes do conhecimento intelectual humano. Para exemplificar, pesquisamos o significado da palavra "sistema" no Dicionário Online Priberam⁴⁴ e, como resultado, pudemos identificar que o vocábulo possui 16 significados diferentes, cada um ligado a uma área diferente do conhecimento, como anatomia, astronomia, filosofia, história natural e até música:

⁴⁴ O Dicionário Priberam da Língua Portuguesa (DPLP) é um dicionário online de português. Compreende o vocabulário geral e os termos das principais áreas científicas e técnicas. O DPLP contém informações sobre as diferenças ortográficas e de uso entre o português europeu e o português do Brasil.

FIGURA 9: CAPTURA DE TELA DO RESULTADO DA BUSCA PELA PALAVRA "SISTEMA" NO DICIONÁRIO ONLINE PRIBERAM

sistema

sistema | *n. m.*

sis·te·ma lêl

(latim *systema*, -atis, do grego *sústema*, -atos, conjunto composto de várias partes)
nome masculino

1. Conjunto de princípios verdadeiros ou falsos reunidos de modo que formem um corpo de doutrina.
2. Combinação de partes reunidas para concorrerem para um resultado, ou de modo a formarem um conjunto.
3. Modo de organização (ex.: *sistema capitalista*).
4. Modo de governo, de administração, de rotação (ex.: *os diferentes sistemas eleitorais*).
5. Conjunto de meios e processos empregados para alcançar determinado fim.
6. Conjunto de métodos ou processos didáticos.
7. Método, modo, forma.
8. [Anatomia] Conjunto de órgãos compostos pelos mesmos tecidos e destinados a funções análogas (ex.: *sistema digestivo*, *sistema respiratório*). = APARELHO
9. [Astronomia] Grupo de corpos celestes associados e que giram segundo as leis de gravitação.
10. [Filosofia] Conjunto de ideias científicas ou filosóficas logicamente solidárias, consideradas mais na sua coerência que na sua verdade.
11. [Geologia] Períodos que separam as eras (ex.: *sistema devoniano*).
12. Modo de distribuição ou disposição das cordilheiras quando se apresentam mais ou menos no mesmo sentido.
13. [História natural] Método de classificação no emprego de um só ou de um pequeno número de caracteres (ex.: *sistema de Lineu*).
14. Modo de distribuição dos seres da natureza.
15. Classificação dos seres que tem somente por fim facilitar o estudo dos mesmos seres.
16. [Música] Reunião dos intervalos musicais elementares compreendidos entre dois limites sonoros extremos e agradáveis ao ouvido.

Fonte: Sistema <<https://dicionario.priberam.org/sistema>>

Pode-se perceber que, em muitos dos significados, a palavra *conjunto* está atrelada a *sistema*. Segundo o dicionário consultado, portanto, podemos dizer que um sistema é um conjunto (seja ele de princípios, de partes, ideias ou mesmo órgãos) que se caracteriza por ser formado por várias partes, ou seja, adquire também o significado de coletivo. Um sistema, portanto, só pode ser considerado como tal, quando identificadas as suas estratificações e relações entre seus elementos. Tendo em vista a variedade de significados atrelados ao termo "sistema", faz-se necessário refletir sobre quais aspectos são relevantes para esta pesquisa, pois mesmo que seja considerado como um termo de ampla significação, nosso intuito, aqui, é delimitar e contextualizar sua compreensão.

Entendemos, portanto, que a palavra sistema designa um conjunto de aspectos que, quando colocados em relação entre si, formam determinado contexto. Aqui, tomamos como pressuposto os sistemas de informação, quaisquer que sejam, desde que atendam a alguns critérios: a) sejam criados a partir de uma linguagem de programação; b) tenham uma estrutura algorítmica lógica capaz de ser compreendida por meio de estatística ou equações; c) funcione

de forma a transmitir uma informação (comando, sinal, mensagem) de um ponto a outro. O sistema comunicacional de Shannon é o exemplo ideal, que atende a todos os pontos acima mencionados e os aspectos que são colocados em relação dentro do sistema são as etapas do processo (fonte de informação, canal, ruído, mensagem, emissor, receptor). Cada etapa pode ser estudada de forma individual, tendo em si suas próprias características e seu próprio contexto, mas também podemos estudar o sistema de Shannon de forma completa, buscando compreender as complexas relações desempenhadas por cada etapa do sistema como um todo. Neste caso, o sistema de Shannon pode ser analisado: como um todo, como cada uma de suas partes, e como um todo e suas partes de forma relacional.

Bertocchi (2016) explica que, atualmente, os sistemas podem ser classificados como abertos ou fechados, "sendo que tal classificação considera a interação do sistema com o meio ambiente (sistema-entorno)" (p. 40). Vale ressaltar que os sistemas fechados podem ser considerados como sistemas mecânicos e "possuem pouca ou quase nenhuma interação com o meio", diferente dos sistemas considerados abertos, ou adaptativos, orgânicos, que "buscam a preservação de uma estrutura dentro de certos limites (processo de homeostase)" (p.41). A autora oferece como exemplo de sistema aberto o que ela denomina como "sistema cibernético", cuja principal característica é sua complexidade e a morfogênese (recriação de estruturas), ou seja, "ao contrário dos demais, os sistemas cibernéticos têm características adaptativas, nas quais a criação, a elaboração e a modificação das estruturas são tidas como pré-requisitos para permanecerem viáveis como sistemas operantes." (idem).

A entropia é um conceito central, também, na teoria dos sistemas. Segundo a autora, "o ponto crucial na distinção entre sistemas abertos e fechados é a questão da entropia" (Idem). Mas, ao contrário do que relatamos acima, na teoria matemática da comunicação, proposta por Shannon (que relaciona a entropia ao ruído), aqui, a autora destaca outra característica da entropia, adicionando a ela um novo significado: o de desordem. Embora a autora não aprofunde a discussão acerca da entropia, é relevante destacar que ela defende que o "pensamento sistêmico objetiva enxergar o todo, detectar padrões e correlações, abrindo oportunidades de reestruturações mais inteligentes do sistema" (p.42). A grande contribuição de Bertocchi (2016), para nosso estudo, reside no fato de que a autora resgata a teoria geral dos sistemas correlacionando-o com o jornalismo, isto é, considerando-o sob um ponto de vista sistêmico, levando em consideração o fato de que o jornalismo é formado por pequenas estruturas que o compõem (e que só fazem sentido quando relacionadas umas com as outras, compartilhando de um mesmo objetivo e de um mesmo *ethos* a respeito de suas práticas e valores).

Bertocchi (2016), relata que, nos anos 1980, outro pesquisador brasileiro já havia atentado para uma perspectiva sistêmica do jornalismo, ao propor um modelo didático de abordagem do jornalismo impresso a partir da Teoria Geral dos Sistemas: Edvaldo Pereira Lima (USP). A autora cita, ainda, Edgar Morin, invocando seu conceito de complexidade, para apresentar a teoria de Ronaldo Henn, autor que "coloca o jornalismo como sistema em outros termos: ele restringe a delimitação do sistema jornalístico ao que ele possui de mais essencial, ou seja, a produção do noticiário" (p.43). Mesmo recorrendo a vários autores, Bertocchi (2016) adquire um posicionamento próprio com uma visão holística em relação à teoria dos sistemas, colocando o produto do jornalismo digital (a narrativa jornalística digital) como um sistema aberto, adaptativo e complexo, "uma vez que sua sobrevivência depende da adaptabilidade de sua estrutura em relação aos demais sistemas em seu entorno, com os quais interage" (idem). Isso quer dizer que, além de organizar-se internamente enquanto sistema, a narrativa jornalística digital está sujeita a alterações de acordo com o nível de interação com outros sistemas que, naturalmente, estão envolvidos com ela. É o caso, por exemplo, de uma narrativa (notícia) publicada em uma rede social digital, como o *Facebook*⁴⁵, cujas interações levam em consideração não apenas outros grupos de mídia (outros jornais com perfis online), mas também a participação dos leitores, que podem interferir na narrativa fazendo comentários, compartilhando e adicionando fatos à informação originalmente publicada.

Além disso, também deve-se levar em consideração os sistemas de internet, já que são eles que possibilitam a conexão entre os jornais e os leitores, dos jornais com as fontes (sejam elas usuários, ou fontes oficiais do governo, dados públicos, etc.), dos jornais com a concorrência, dos jornais com outras plataformas (como o *WordPress*⁴⁶), e outras interações que podem influenciar no sistema narrativo do jornalismo digital.

Defendendo a narrativa jornalística digital como um sistema adaptativo complexo, a autora oferece um panorama sobre as características desses sistemas: os sistemas abertos, por exemplo, têm características adaptativas, nas quais "a criação, elaboração e a modificação das estruturas são tidas como pré-requisitos para permanecerem viáveis como sistemas operantes." (BERTOCCHI, 2016 p.43). De acordo com a autora, os sistemas abertos constituem-se de

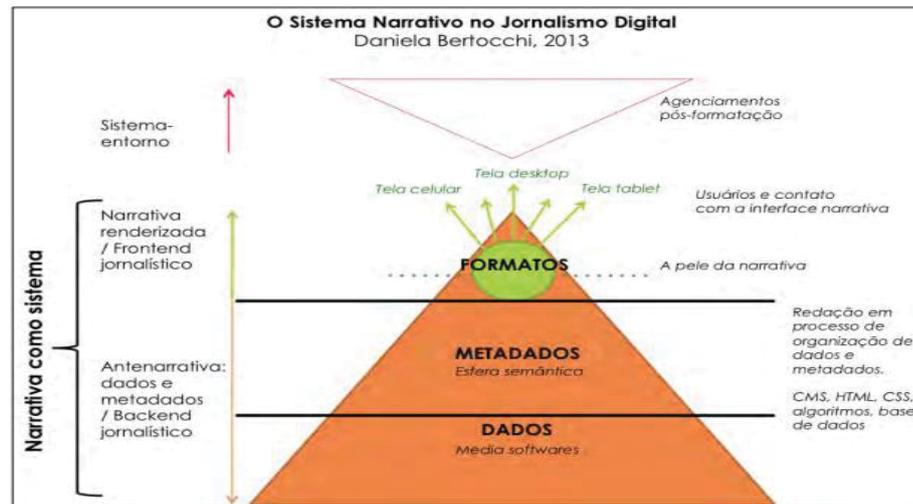
⁴⁵ Facebook é uma rede social vinculada ao Grupo Meta Inc. (grupo proprietário de outras plataformas bastante utilizadas mundialmente, como WhatsApp, Instagram e Messenger) considera a rede social mais utilizada pela população mundial: com mais de 2,91 bilhões de contas ativas em 2022, segundo o report *We Are Social* <<https://wearesocial.com/uk/blog/2022/01/digital-2022-another-year-of-bumper-growth-2/>>

⁴⁶ Lançamento em 2003, o *WordPress* é um dos mais utilizados CMSs (*Content Management Systems*) do mundo para criação e divulgação de conteúdo na web.

partes, que podem ser agenciadas entre si e que tendem a perseguir um propósito dentro do sistema.

Em 2013, ano em que defendeu sua tese de doutorado, Daniela Bertocchi apresentou um quadro com uma proposta de modelo de sistema narrativo no jornalismo digital, demonstrando as diversas camadas pertencentes ao sistema analisado:

FIGURA 10 – MODELO PROPOSTO POR DANIELA BERTOCCHI EM 2013



Fonte: Daniela Bertocchi (2013)

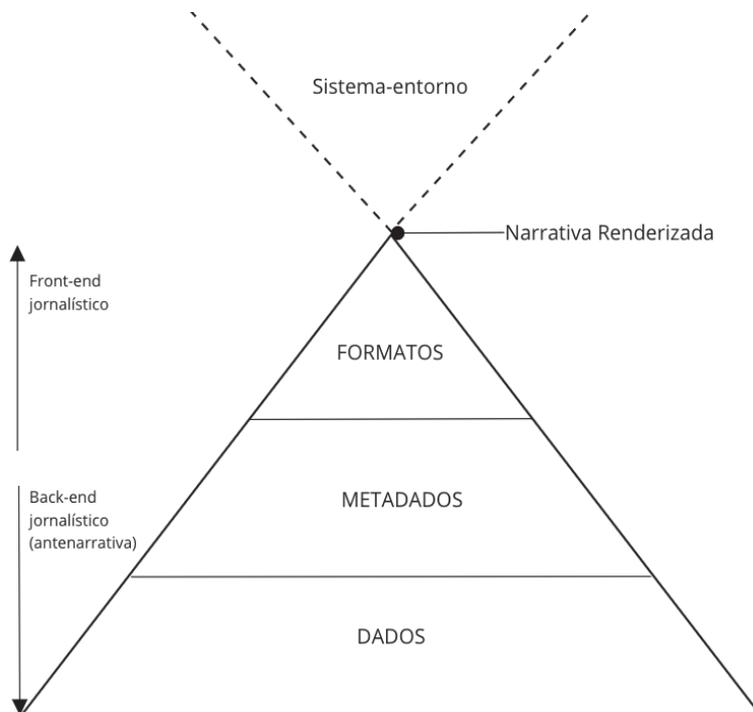
Alguns anos mais tarde, a autora reeditou sua tese apresentada em 2013 e a transformou em um livro, no qual trouxe um modelo de sistema narrativo no jornalismo digital atualizado (2016). Nesse novo modelo, o esquema conta com o mesmo número de camadas, porém com menos variáveis que interferem no funcionamento do modelo sistêmico proposto originalmente pela autora. Para ela, como outros sistemas abertos (adaptativos), o sistema jornalístico digital compõe-se de partes agenciadas entre si e sua característica principal é a complexidade e a capacidade de morfogênese, ou seja, o potencial de recriar suas próprias estruturas de forma adaptativa, de acordo com a necessidade.

No novo modelo apresentado pela autora, em seu livro publicado em 2016, o sistema narrativo no jornalismo digital é também representado por uma pirâmide, na qual as duas primeiras camadas, de baixo para cima, assinaladas como *Dados e Metadados*, compõem o que a autora considera como "antenarrativa", ou seja, todo o processo sistêmico ao qual a notícia é submetida, antes mesmo de tornar uma narrativa propriamente dita (ou uma narrativa visualmente adequada para o consumo do público, como defende a autora). Já a camada de cima

designa os *Formatos*, cuja representação é o que a autora considera como "o momento da narração ou formatação narrativa", ou seja, o momento em que a narrativa se torna visível para os usuários.

No ponto mais alto da pirâmide, Bertocchi insere o processo de renderização da narrativa. Vale ressaltar que o objeto de análise da autora é a narrativa *digital* do jornalismo, ou seja, a notícia publicada em portais/sites, e replicada de diversas formas no ambiente digital. No topo da pirâmide do jornalismo digital, consideramos a renderização da narrativa o exato momento em que ela é transformada e reproduzida no formato narrativo digital (a notícia "desenhada" na tela do usuário, "diagramada" de acordo com a interface que o usuário utiliza - computador, smartphone, tablet - e como ela aparece de fato na aplicação digital). Após a formatação e renderização da notícia, a autora destaca que há ainda a existência de todo um sistema-entorno, característica muito comum dos sistemas de natureza adaptativa, como é possível verificar a seguir:

FIGURA 11- MODELO SISTEMA NARRATIVO NO JORNALISMO DIGITAL, PROPOSTO POR DANIELA BERTOCCHI EM SEU LIVRO.



Fonte: Bertocchi (2016)

No esquema apresentado acima, a autora simplifica seu modelo de sistema narrativo, na mesma medida em que aprofunda conceitos. A camada dos dados, segundo a autora, refere-

se ao momento de levantamento e seleção das informações em bases de dados (e como veremos no capítulo seguinte, nos sistemas de automação de notícias esses dados podem ser estruturados ou não estruturados e podem ser extraídos de diversas bases de dados em formatos variados, como textos, tabelas, gráficos, áudios, imagens e códigos). Fazendo um paralelo com o sistema de Shannon (1963), este é o momento em que a mensagem está, ainda, na etapa de fonte de informação.

O conceito de metadados, que ocupam a camada intermediária dentro do sistema narrativo proposto por Bertocchi (2016) "refere-se à semantificação desses dados para a compreensão de softwares e algoritmos, para a apresentação de narrativas nas interfaces digitais" (p. 65). Shannon & Weaver (1963) também previram, no sistema comunicacional, a etapa de semantificação da mensagem, considerando-o como uma possível intercorrência dentro do sistema.

No topo da pirâmide, a última camada do sistema (dentro do software) é a dos formatos, que diz respeito ao momento da "organização corporificada da narrativa na interface digital para o acesso dos usuários finais" (BERTOCCHI, 2016, p. 65), ou seja, é o momento em que a notícia, já produzida pelo software, é disponibilizada para os leitores no formato de uma matéria, dentro de um layout específico que deve corresponder às dimensões da tela do usuário (e esse layout deve adaptar-se ao dispositivo do usuário, caso ele esteja vendo a notícia num celular, ou num computador, ou num tablet). Este é o momento em que o receptor, para usar os conceitos de Shannon & Weaver (1963), tem acesso à mensagem codificada pelo destinatário e passa a decodificá-la. Na última camada da pirâmide de Bertocchi (2016), é que o objetivo comunicacional do sistema acaba completando-se.

Vale destacar que a autora considera a narrativa jornalística digital como um sistema aberto e adaptativo, "isso é, a narrativa como sistema é um conjunto complexo, artificial, no qual atuam distintos atores humanos e não humanos e que produz um todo maior que suas partes" (BERTOCCHI, 2016, p. 45). Sendo assim, torna-se relevante enfatizar que a narrativa digital jornalística, objeto de análise da autora, pode ser compreendida como um sistema complexo que, para além de si próprio, agrega também subsistemas com regras singulares, como é o caso de uma notícia que não obteve o engajamento esperado quando publicada no portal de um jornal. A autora destaca que

Apesar de serem familiares, e mesmo com uma compreensão prática do que seja um sistema (um animal, uma cidade, uma floresta), este se torna um ponto de atenção à medida que carrega em si a mudança da mentalidade cartesiana reducionista para o pensamento sistêmico holístico. Assim, em vez de um olhar que traça caminhos diretos da causa para o efeito e em vez de uma observação por partes de um problema,

o pensamento sistêmico coloca os problemas do mundo como problemas complexos e interligados e para os quais devemos olhar, colocando todos seus agentes como responsáveis pelo sistema. Quando se indaga o porquê de uma notícia online não ter gerado tantos *pageviews* quanto o imaginado, responde-se: pessoas não têm mais interesse em ler notícias; ou houve um problema de SEO e a página não está sendo encontrada via buscador; ou a usabilidade está comprometida; ou o site demora a ser carregado. (BERTOCCHI, 2016, p. 46-47).

Esse ponto de vista que a autora traz para seus estudos é bastante fértil também para colocarmos a teoria de Shannon em contraponto: se, de um lado, Shannon visualizava o sistema comunicacional de forma fechada, levando em consideração apenas alguns níveis de intercorrência, do outro lado, Bertocchi amplia a discussão sobre as características dos sistemas, trazendo para a discussão o fato de que um sistema, por si próprio, pode criar relações complexas com outros sistemas e, inclusive, dentro dele próprio (obrigando-o a recorrer ao que a autora chama de morfogênese, ou seja, sua capacidade de readaptar-se de acordo com a necessidade do sistema).

Sendo assim, o pensamento sistêmico ao qual Bertocchi (2016) se refere, acaba levando nosso olhar para um lado que está nas antípodas da simplificação meramente lógica, defendida por Shannon (1949). Esse ponto de vista, que leva em conta outras interferências (inclusive de outros sistemas) traz também para o centro da discussão a questão da complexidade, não apenas da adaptabilidade e da retro-organização do próprio sistema em si mesmo, mas também do sistema em relação a outros sistemas em seu entorno, assim "em vez de um mundo de causas e efeitos, temos um mundo de sistemas capazes de provocar seus comportamentos. Os processos de causa-efeito, em um contexto sistêmico, funcionam como pontos de influência circulares (*feedback loopings*)" (BERTOCCHI, 2016, p. 46). Partimos dessa discussão para apresentar, no capítulo metodológico, um modelo de sistema aberto, adaptativo e complexo, para o jornalismo, que pode ser compreendido dentro de um modelo lógico matemático.

5 OS JORNALISMOS: DA TRADIÇÃO ORAL AO REPÓRTER-ROBÔ

O jornalismo não tem uma data de aniversário. Isso porque é praticamente impossível determinar quando ele nasceu e há quanto tempo existe na sociedade, de fato, já que há registros de que os povos antigos, mesmo antes da invenção da escrita, costumavam relatar acontecimentos da vida por meio da cultura oral. Os fatos da época eram comunicados através da fala. Ainda, alguns autores, como Souza (2008), relatam que mesmo algumas pinturas rupestres podem ser consideradas como reproduções de fatos do cotidiano do homem na antiguidade. Mesmo que não possamos defini-las como jornalismo, certamente podemos tomar como exemplares rudimentares do início da reprodução dos fatos do cotidiano. Se não é jornalismo, faz parte da história do homem. E isso já é o suficiente para iniciarmos nossa incursão genealógica do jornalismo, que inicia na tradição oral, e tem um grande avanço cerca de 3.500 a.C., quando a escrita era inventada na antiga Mesopotâmia⁴⁷. Para Souza (2008), falar da história do jornalismo é fazer uma viagem aos mais remotos registros da humanidade. Segundo o autor,

A gênese do jornalismo situa-se na Antiguidade Clássica, havendo uma retoma na Idade Moderna, graças ao Renascimento, ao desenvolvimento do espírito iluminista da Ilustração e à satisfação das necessárias condições técnicas (tipografia de Gutenberg, fábricas de papel) e socioeconômicas (alfabetização, capital, iniciativa privada e empreendedorismo). (SOUZA, 2008, p. 03).

O fato é que o jornalismo não assopra velinhas quando faz aniversário. Ele é tão antigo, que acompanha o desenvolvimento da sociedade e das técnicas humanas, além de ser um instrumento a favor das necessidades de trocas de informações entre os povos. Ele pode ser considerado intrínseco à história da humanidade, já que não podemos conceber uma sociedade na qual ele não exista, já que a cultura do homem, desde seus primórdios, é a de troca de informações. Em determinado momento, essa troca se dava pelo instinto de sobrevivência; em outro, pela troca de informações quando os povos deixaram de ser nômades e passaram a cultivar suas próprias terras, passando pelos primeiros registros administrativos, pelo entretenimento, pelo jornal impresso, folhetins, rádio, televisão, até chegar numa era em que o jornalismo é feito por robôs, como apresentaremos a seguir. Por isso, embora empreguemos esforços para encontrar a origem do jornalismo, sempre vamos falhar na busca por uma data específica. Ele sempre esteve ali.

⁴⁷ Costuma-se delegar a invenção da escrita aos povos Mesopotâmicos, mas vale ressaltar que a escrita também vinha sendo desenvolvida, de forma simultânea, em outras partes do mundo. Como a criação do sânscrito, na Índia, e dos hieróglifos, no Egito.

Já que não podemos precisar com eficiência seu surgimento, vamos empreender esforços em agrupar algumas visões sobre o que é o jornalismo, afinal, e sua relevância para a sociedade.

5.1 O QUE É, O QUE É?

Jornalismo é: técnica.

Não parece tarefa fácil definir algo tão amplo e que, como vimos, existe na sociedade há tanto tempo. Defendemos que o jornalismo assume vários significados de acordo com o tempo e de acordo com o desenvolvimento natural da sociedade, inclusive, como veremos na sequência, há de ressaltar-se alguns movimentos que foram importantes para a construção da sociedade, como a Revolução Industrial, o avanço do capitalismo e das burguesias, a invenção da fotografia, do rádio, até chegarmos ao século XX, com as vertiginosas inovações na tecnologia e no sistema de redes. Traquina (2020) posiciona-se e relata que "é absurdo pensar que possamos responder à pergunta: o que é o jornalismo? numa frase ou até mesmo em um livro" (p. 14). De fato, conceituar um termo como o jornalismo não é uma função para completar-se rapidamente e em poucas linhas. Não porque o jornalismo seja difícil de compreender, mas porque ele se desenvolve junto com a história da sociedade. E essa história, sim, é bastante complexa.

De qualquer forma, vamos empreender esforços em trazer alguns conceitos e apresentar, de forma resumida, algumas perspectivas sobre o conceito, tentando não cair no generalismo de pensar que o jornalismo é apenas uma narrativa do real, mas tentando apropriarmos de outros aspectos que também contribuem para o entendimento do jornalismo e suas relações com as grandes evoluções da humanidade. O que gostaríamos, no entanto, de destacar, é que, neste estudo, vamos trazer apenas alguns conceitos sobre o tema, sem a intencionalidade de esgotá-lo, já que compreendemos que existem diversas linhas de pensamento que podem divergir, em maior ou menor grau, sobre a conceituação em relação ao jornalismo.

E essa decisão é ancorada naquela frase de Traquina (2020), pois é insensato pensar que em apenas um capítulo poderíamos responder com precisão à pergunta: "o que é o jornalismo?". Logo, na impossibilidade de trazer todas as visões sobre o assunto, vamos focar em alguns pontos: definição do conceito, relacionamento entre o jornalismo e o capitalismo industrial, jornalismo na era digital e automação no jornalismo. Outro ponto de destaque é que nossa perspectiva, nesta tese, está centrada no jornalismo no século XXI, nas limitações e

desafios impostos à prática jornalística na contemporaneidade. Por essa razão, a seguir incluímos algumas conceituações sobre o jornalismo, mas buscamos aprofundarmo-nos, de fato, no jornalismo de automação, ou quinto jornalismo, como defende Carreira (2017), fase na qual estamos inseridos e na qual este estudo está sendo desenvolvido.

Para Charron & Bonville (2016) "o jornalismo é, por definição, uma prática discursiva realista sobre um referente real, em oposição a outras formas de expressão, como a literatura, a pintura, cujos referentes são ou podem ser fictícios ou imaginários." (CHARRON & BONVILLE, 2016, p. 187-188). Já Souza (2008), aponta que o jornalismo nada mais é do que "uma representação discursiva de factos e ideias da vida do homem, construída para se contar ou mostrar a outrem. Por outras palavras, o jornalismo é uma representação discursiva da vida humana na sua diversidade de vivências e ideias." (SOUZA, 2008, p. 05). Traquina (2020) concede-nos outra definição, mais poética, quando relata que "o jornalismo é a vida, tal como é contada nas notícias de nascimentos e de mortes" (p. 14) e que, além disso, "poder-se-ia dizer que o jornalismo é um conjunto de 'estórias', 'estórias' da vida, 'estórias' das estrelas, 'estórias' de triunfo e tragédia." (TRAQUINA, 2020, p. 15).

Entre as três referências que trouxemos acima, podemos situar Traquina (2020) entre aqueles pesquisadores que observam o jornalismo com uma perspectiva mais poética e inspirada, enquanto Charron & Bonville (2016) buscam uma conceituação focada na literalidade e objetividade do jornalismo. Souza (2008) traz uma visão realista, mas que também flerta com a poesia em determinados momentos. E é sobre essa colcha de retalhos que enxergamos o trabalho científico: buscar, em fontes diferentes, suas percepções e visões (sejam elas apocalípticas e integradas, fascinadas ou obscurantistas) acerca de um mesmo assunto. Embora Traquina (2020) nos presenteie com uma visão poética sobre o conceito, assim como Genro Filho (2004), ele também discute a importância do jornalismo nas sociedades democráticas e o situa dentro de uma dinâmica capitalista, como veremos na sequência.

Para Charron & Bonville (2016), ainda, o jornalismo diferencia-se de outras atividades como a literatura, por exemplo, porque busca retratar exclusivamente acontecimentos do real. Enquanto na literatura o escritor tem liberdade para criar cenários imaginários e acontecimentos, no jornalismo essa é uma conduta considerada inadmissível, pois,

diferentemente dessas formas de expressão (*literatura, arte*), que estão livres da obrigação de representar fielmente o real, o jornalismo não pode, sem deixar de existir, escapar dessa imposição essencial: representar o real de maneira que dê a todos os agentes sociais engajados em sua produção, jornalistas, fontes de informação, anunciantes e leitores, a convicção do real. (CHARRON & BONVILLE, 2016, p. 188 – *grifo nosso*).

Os autores não descartam, no entanto, o fato de que o jornalista, ao contar uma história real, possa usar de sua criatividade e subjetividade para que a narrativa se torne mais atrativa ao público, mas atentam para o fato de que o principal objetivo do jornalismo é reproduzir a realidade, isto é, "por mais criativo que o jornalismo seja, [...] ele sempre precisa, para não correr o risco de se negar como prática discursiva específica, invocar o real" (CHARRON & BONVILLE, 2016, p. 188). Essa citação nos dá a abertura para falarmos sobre uma das características mais importantes do jornalismo: a objetividade. Possivelmente, essa é a primeira palavra que estudantes nas faculdades de jornalismo escutam, já no primeiro dia de aula. O jornalismo deve ser objetivo, logo, a subjetividade do jornalista deve ser deixada (ao máximo) de lado na hora de relatar os fatos. Isso porque entende-se que, quanto mais objetiva é a notícia, mais fidedigna à realidade. No entanto, essa afirmação gera alguns debates férteis a respeito da impossibilidade de separar-se o sujeito (e sua subjetividade) daquilo que está sendo contado (o fato). Outra contribuição de Traquina (2020) para o conceito, mostra que o jornalismo é uma reprodução da vida em suas diversas dimensões, como em uma "enciclopédia" na qual podemos consultar diversos assuntos. Para o autor,

Uma breve passagem pelos jornais diários vê a vida dividida em seções que vão da sociedade, a economia, a ciência e o ambiente, à educação, à cultura, à arte, aos livros, aos media, à televisão, e cobre o planeta com a divisão do mundo em local, regional, nacional (onde está essencialmente a política do país) e internacional. (TRAQUINA, 2020, p. 14).

Está claro que o jornalismo encontra na realidade o seu motivo de ser; nos acontecimentos cotidianos, sua inspiração maior. Traquina (2020) ainda nos auxilia na compreensão sobre o termo evidenciando que o jornalismo pode ser entendido como a frase "que é a resposta à pergunta que muita gente se faz todos os dias – o que é que aconteceu/está acontecendo no mundo? no Timor? no meu país? na minha “terra”?" (p. 14). De forma resumida, o jornalismo é a narrativa criada a partir daquilo que aconteceu. É, portanto, um recorte da realidade.

Para avançar no entendimento sobre o conceito de jornalismo, Souza (2008), em sua obra **Uma história breve do jornalismo no Ocidente**, traz relevantes contribuições para o compreendermos. Aqui, nesta tese, falamos dos jornalismo, no plural, pois, especialmente no século XX, o jornalismo toma diversos significados diferentes. Para Souza (2008), coexistiram (e alguns ainda coexistem) diversos modelos de jornalismo na sociedade ocidental: modelo autoritário de jornalismo, modelo comunista de jornalismo, modelo desenvolvimentista,

revolucionário, fotojornalismo, o jornalismo televisivo, o jornalismo participativo, o novo jornalismo, o jornalismo gonzo, o jornalismo cívico, o jornalismo de precisão, entre outros. Para compreender, de forma cronológica, as fases do jornalismo e suas derivadas conceituações, sugerimos a leitura da obra completa de Souza (2008), que traz um panorama plural acerca do assunto.

5. 2 JORNALISMO É CAPITAL, CAPITAL É PODER

Jornalismo é: indústria.

Traquina evidencia que o jornalismo como o conhecemos hoje tem suas raízes no século XIX, com a expansão dos jornais impressos e a criação de grandes veículos de comunicação, o que permitiu a geração de cada vez mais empregos formais sob a denominação de "jornalistas". O autor destaca que "embora houvesse pessoas que, por exemplo, fizeram negócio com a venda de jornais durante a revolução francesa no fim do século XVIII, os jornais eram sobretudo armas na luta política, estreitamente identificados com causas políticas." (TRAQUINA, 2020, p. 25). Assim, no século XIX, os veículos de comunicação, que começavam a surgir com maior presença nas sociedades democráticas, tinham um objetivo maior em mente: o de informar. Genro Filho (2004), no entanto, destaca também o caráter ideológico que se instaurou no jornalismo nessa época. Para o autor, além de informar, o jornalismo assumiu uma posição de instrumento para a perpetuação de determinadas linhas ideológicas, bastante alinhadas com as necessidades da burguesia. Segundo Genro Filho (2004),

O jornalismo, como forma específica de transmissão de informações requer, naturalmente, um meio técnico apropriado capaz de multiplicar uma mesma informação. A imprensa de Gutemberg, como meio técnico que origina o jornalismo, ou seja, que origina a transmissão sistemática de informações de determinada natureza, tem seu desenvolvimento intimamente ligado ao desenvolvimento do capitalismo. (GENRO FILHO, 2004, p. 160).

Sob essa perspectiva, o jornalismo passa a ser considerado como uma indústria e seu produto, a notícia, adquire as feições de uma mercadoria. É no capitalismo, então, que o jornalismo encontra sua base material de expansão, não apenas como um propagador de ideologias dominantes, mas também como um veículo de propaganda, conforme aponta Genro Filho (2004), pois "a formação do modo de produção capitalista e o surgimento da burguesia, que precisava ampliar infinitamente o mercado e conseqüentemente criar novas necessidades em todas as pessoas em condições de comprar" foram as condições que permitiram a expansão

rápida do jornalismo após a Revolução Industrial, principalmente tomando força entre os séculos XIX e XX. De fato, no século XX, o jornalismo passa a tomar um ritmo industrial, especialmente com os avanços das tecnologias como o rádio, a televisão, a invenção dos computadores e da internet, que formaram as bases técnicas e materiais para a evolução do jornalismo de maneira acelerada.

Se antes uma notícia demorava determinado tempo para ser escrita, revisada, impressa e enviada aos assinantes, com o rádio e a televisão a notícia era dada em tempo quase real e em uma velocidade jamais vista antes. Com a chegada da internet, outra grande revolução impunha desafios ao jornalismo, já que esse cenário potencializa-se e torna-se ainda mais complexo, quando, além da velocidade no compartilhamento das informações, os próprios usuários/leitores passam a fazer parte do processo de produção jornalística, momento que alguns autores convencionaram chamar de jornalismo participativo⁴⁸.

Genro Filho (2004) descreve que, aos poucos, "o conceito de jornalismo vai sendo delineado a partir de uma teorização mais profunda e da crítica do conceito empírico." (GENRO FILHO, 2004, p. 160), o que possibilitou pensar sobre o jornalismo de forma teórica mais aprofundada. Para o autor, o jornalismo de fato "não é os meios técnicos, embora deles necessite. E não é, também, apenas a transmissão sistemática de informações ao meio social, na medida que implica necessariamente num determinado tipo de informação, necessário em condições históricas." (Idem p. 161). As "condições históricas" às quais o autor se refere são as condições criadas e compartilhadas dentro de uma lógica do sistema capitalista, como os modos de produção, a visão de jornalismo como mercadoria e as ferramentas técnicas da época.

Genro Filho (2004) destaca que o jornalismo não é o meio técnico, nem a transmissão de informação. Mas, então, o que seria o jornalismo sob o seu ponto de vista? É nesse ponto que o autor destaca a necessidade de pensar-se o jornalismo para além dos conceitos meramente descritivos, trazendo a inevitável reflexão sobre uma teoria geral do jornalismo, visualizando-o não como uma prática ou como um conceito, mas como um campo de estudos, atravessado por complexidades epistemológicas. Embora estejamos avançados no jornalismo de forma técnica e sistemática (com regras, pressupostos e um código de conduta), destaca o autor, ainda estávamos, à época em que foi publicado o artigo de Genro Filho, na pré-história de sua compreensão enquanto fenômeno empírico amplo, permeado de ligações com a cultura, sociedade, política, economia, sociologia. É por isso que Genro Filho (2004) defende a

⁴⁸ Ver Souza (2008), HOLANDA et. al (2008) e RODRIGUES (2013);

necessidade de olhar para o jornalismo com uma lente ampliada, a qual denomina de teoria geral do jornalismo. Isto é, uma teoria que

explique o fenômeno historicamente, que o reconheça enquanto estrutura específica de comunicação, situando nessa totalidade uma série de conceitos que hoje são operacionalizados, parece uma proposição estritamente válida. Isto implicaria numa redefinição de conceitos como “notícia”, “reportagem”, “editorial”, etc. e no questionamento de seus aspectos estruturais e do próprio conteúdo. Desde os chamados “atributos da notícia”, como atualidade, veracidade, curiosidade, proximidade e outros, até as questões estruturais do “lead” precisariam ser rediscutidos em função de uma compreensão globalizante. (GENRO FILHO, 2004, p. 162).

Os estudos sobre as teorias do jornalismo começam a aparecer, no Brasil, a partir de 1950, mas "só ganharam o respaldo definitivo da universidade na década de 1970, graças ao trabalho do professor José Marques de Melo (PENA, 2005, p. 1). Naquele momento (até metade do século XX) onde os estudos em jornalismo ainda eram incipientes, o que se entendia como jornalismo era basicamente o uso dos jornais como forma de divulgar opiniões, fazer propagandas, difundir discursos ideológicos e posicionamentos políticos. É nessa época, também, que o jornalismo passa a ser visto como uma atividade profissional remunerada. Nascia, finalmente, a imprensa. Genro Filho (2004) atenta para a importante distinção entre jornalismo e imprensa. Para o autor, definir esses conceitos é uma tarefa fundamental e, segundo seu entendimento,

A imprensa é o corpo material do jornalismo, é o processo técnico e mecânico que resulta num produto final (no caso da imprensa escrita) composto de papel e tinta, o jornalismo é a natureza da informação que surge em função destes meios e das necessidades sócio-políticas de um período histórico. (GENRO FILHO, 2004, p. 160).

Traquina relata que "o surgimento do jornalismo enquanto atividade remunerada está ligado à emergência dum dispositivo tecnológico, à emergência do primeiro *mass media*, a imprensa." (2020, p. 29). Nesse momento, os grandes conglomerados de mídia estavam estruturando-se, tornando os jornais cada vez mais populares, com tiragens maiores, com mais jornais sendo criados. Enfim, o jornalismo entrava numa era industrial. É possível verificar a vertiginosa expansão da imprensa pelo número de jornais franceses publicados entre 1830 até 1914:

O número de jornais franceses aumentou de 49, em 1830, para 73 em 1867, para 220 em 1881, e 322 na véspera da Primeira Guerra Mundial, em 1914. As tiragens dos jornais também sofreram um aumento notável durante o século: 34.000 em 1815, 1.000.000 em 1867, 2.500.000 em 1880, e 9.500.000 em 1914. Nos Estados Unidos, o número de jornais publicados a nível nacional duplicou entre 1830 e 1840. (TRAQUINA, 2020, p. 26).

Em menos de 100 anos, 273 novos jornais foram criados na França. Foi na efervescência do final do século XIX e início do século XX que o jornalismo instaurou-se não apenas como uma atividade remunerada, mas também como uma profissão cada vez mais alinhada com o objetivo de informar os fatos. As invenções que ganhavam notoriedade naquele período, como a máquina fotográfica, o rádio, o telégrafo⁴⁹, a impressão colorida, entre outras, contribuíram para profissionalizar cada vez mais a prática jornalística. Traquina destaca que "em particular, a invenção da máquina fotográfica iria [...] inspirar o jornalismo no seu objetivo de ser as "lentes" da sociedade, reproduzindo *ipsis verbis* a realidade. (Idem, p. 29). Para o autor, o telégrafo, inclusive, que trouxe relevantes avanços na transmissão de informações à distância, tornaria o jornalismo cada vez mais global, inaugurando uma nova era na profissão, que até então atuava de forma local e (com muitas dificuldades) de forma regional.

A imprensa, assim, avançava rapidamente em ritmo industrial em diversas partes do mundo, adquirindo, inclusive, cada vez mais notoriedade e importância. Para Traquina (2020), a imprensa acabou tornando-se um pilar importante da sociedade, principalmente como um instrumento de defesa da democracia. O autor afirma que

Existe uma relação simbiótica entre jornalismo e democracia na teoria democrática. Mas a teoria democrática define claramente um papel adversarial entre o poder político e o jornalismo, historicamente desde o século XIX chamado o "Quarto Poder", talvez porque séculos de domínio autocrático e por vezes despótico criaram um legado de desconfiança, suspeita e medo em relação ao poder político. (TRAQUINA, 2020, p. 17).

A influência da democracia no jornalismo é o que o torna tão imprescindível em uma sociedade permeada por disputas de poder, especialmente ligadas ao capitalismo. Segundo Souza (2008), o modelo ocidental do jornalismo está bastante alicerçado na liberdade de expressão e na independência do jornalismo em relação ao estado e aos setores do poder. Sendo assim, o jornalismo, em tese, teria

o direito a reportar, comentar, interpretar e criticar as atividades dos agentes de poder, inclusive dos agentes institucionais, sem repressão ou ameaça de repressão. Teoricamente, os jornalistas são apenas limitados pela lei (tida por justa), pela ética e pela deontologia. O campo jornalístico configura-se, assim, como uma espécie de espaço público, um mercado livre de ideias, onde se ouvem e, por vezes, se digladiam as diferentes correntes de opinião. (SOUZA, 2008, p. 195).

⁴⁹ Inventado por Samuel Morse em 1837, o telégrafo revolucionou a forma como as notícias eram transmitidas. O telégrafo permitiu a transmissão de informações em tempo real a longas distâncias, o que possibilitou a criação de agências de notícias como a Associated Press, em 1846.

Entendido como o "Quarto Poder", por muito tempo o jornalismo usufruiu de prestígio justamente porque ele defendia os aspectos da realidade, com isenção, objetividade e liberdade de expressão. Sendo assim, aquele que tinha a possibilidade de contar ou reconstruir a realidade por meio da narrativa jornalística tinha, de fato, poder sobre a sociedade, pois "os jornalistas são participantes ativos na definição e na construção das notícias, e, por efeito, na construção da realidade." (TRAQUINA, 2020, p. 19).

À medida que as notícias passaram a ser vistas, na sociedade capitalista, como uma mercadoria, emergiu uma nova forma de fazer jornalismo e "as notícias tornaram-se crescentemente estandardizadas ao tomarem a forma a que chamamos hoje 'pirâmide invertida'" (TRAQUINA, 2020, p. 47). De certo modo, não apenas a velocidade de transmissão das informações havia mudado, mas também a forma como essas informações eram coletadas, organizadas e descritas. A pirâmide invertida invocava um novo modo de estruturar as notícias, padronizando a maneira como eram organizadas as informações dentro de um texto jornalístico.

Nessa época, além da pirâmide invertida, "os jornalistas inventaram novas formas, novas práticas, novas técnicas, e uma série de valores que contribuíram para uma identidade profissional" (Idem, p. 60). O próprio conceito de entrevista é cunhado durante esse período, quando também começam a surgir os manuais de redação jornalística, o lide como primeiro parágrafo, entre outras.

5.3 UMA HISTÓRIA CENTENÁRIA: CATEGORIZANDO AS FASES DO JORNALISMO ATÉ A CONTEMPORANEIDADE

Jornalismo é: ciência.

Marcondes Filho (2002) realiza uma espécie de sistematização na qual separa a história do jornalismo por épocas, sendo a primeira fase, ou como o autor sugere "*o primeiro jornalismo*", aquele realizado entre 1789 até metade do século XIX, influenciado pelo movimento iluminista da época, que surgiu na Europa. Não apenas o jornalismo sofreu influências dos pensadores dessa corrente filosófica, mas outros setores como as artes, a política e, principalmente, as ciências, quando a sociedade passa a defender a racionalidade como principal motor de mudanças. Os pensadores iluministas acreditavam nas liberdades individuais, na ciência e no método científico, na igualdade perante a lei, na tolerância religiosa e na separação entre os poderes do Estado. Essa onda de pensamento, inclusive, teve grande

impacto em movimentos sociais e políticos da época, como a Revolução Francesa, por exemplo. Segundo Marcondes Filho (2002),

O primeiro jornalismo, de 1789 à metade do século 19, foi, assim, o da "iluminação", tanto no sentido de exposição do obscurantismo à luz quanto de esclarecimento político e ideológico. O controle do saber e da informação funcionava como forma de dominação, de manutenção da autoridade e do poder, assim como facilitava a submissão e a servidão. Enquanto eu não sei que o poder é algo dos homens, eu acredito que ele é natural, que Deus e a natureza criam homens para mandar e outros para servir; (MARCONDES FILHO, 2002, p. 11).

O período do segundo jornalismo, de acordo com a concepção de Marcondes Filho (2002), é a fase que compreende o período dos grandes avanços tecnológicos no século XIX. Esta é fortemente marcada pelo capitalismo, no qual os veículos de imprensa são considerados como empresas capitalistas e a notícia, como mercadoria, além das inovações que chegavam aos processos de produção do jornal. Para o autor, a transformação tecnológica da época exigia que as empresas fossem regulamentadas de acordo com uma lógica de mercado, por isso a empresa jornalística deveria desenvolver a capacidade de sustentar-se por si própria, o que iria "transformar uma atividade praticamente livre de pensar e de fazer política em uma operação que precisará se autofinanciar." (MARCONDES FILHO, 2002, p. 13).

FIGURA 12 – CARACTERÍSTICAS DE CADA FASE DO JORNALISMO

	Tipo	Época	Valores jornalísticos dominantes	Aspectos funcionais e tecnológicos	Agentes	Economia
Pré-história	artesanal	1631 a 1789	Espetacular, singularmente novo (desastres, mortes, seres deformados, reis, etc)	Jornal ainda semelhante ao livro, poucas páginas	Empreendedor isolado	Elementar
Primeiro Jornalismo	Político-literário	1789 a 1830	Razão (verdade, transparência); questionamento da autoridade; crítica da política; confiança no progresso	Profissionalização; surge a redação; diretor separa-se do editor; artigo de fundo; autonomia da redação	Políticos; escritores; críticos; cientistas	Economia deficitária
Segundo Jornalismo	Imprensa de massa	1830 a +- 1900	O furo; a atualidade; a neutralidade; criam-se a reportagem, as enquetes, as entrevistas, as manchetes; investe-se nas capas, logo e chamadas de la. página	Rotativas e composição mecânica por linotipos (1890); telégrafo e telefone; cria-se a agência Havas; mais publicidade e menor o peso de editores e redatores; títulos passam a ser feitos pelo editor	Jornalistas profissionais	Economia de empresa: jornal tem que dar lucro; aumento das tiragens: 35 para 200mil
Terceiro Jornalismo	Imprensa-monopolista	de +- 1900 a +- 1960	Grandes rubricas políticas ou literárias; páginas-magazines: esporte, cinema, rádio, teatro, turismo, infantil, feminina	Influência da indústria publicitária e das relações públicas; uso da fotografia	Jornalistas, publicitários e relações públicas promovem indústria da consciência	Grupos monopolistas dominam a imprensa; época de tiragens-monstro
Quarto jornalismo	Informação eletrônica e interatividade	de +- 1970 até o presente	Impactos visuais; velocidade; transparência	Implantações tecnológicas (barateamento da produção); alteração das funções do jornalista; toda a sociedade produz informação	Redes/sistemas informatizados; pessoas em interface; jornalistas prestadores de serviço	Financiamentos migram para a TV e a Internet; crise da imprensa escrita

Fonte: Marcondes Filho (2002)

Com o passar do tempo, já no século XX, o desenvolvimento acelerado da sociedade industrial e "o crescimento das empresas jornalísticas desembocam na constituição do terceiro jornalismo, o de monopólios, cuja sobrevivência só será ameaçada pelas guerras e pelos governos totalitários do período." (MARCONDES FILHO, 2002, p. 14). Nesse período, marcado por guerras e regimes totalitários, o jornalismo perde um pouco de sua força, mas em paralelo, a imprensa, enquanto negócio, acaba evoluindo e tornando-se cada vez mais lucrativa. O aumento fantástico da produção significou uma total reorientação da indústria jornalística no sentido de render lucros e tornar-se economicamente autossustentável (Idem, p. 32). O autor considera, como consequência que, em sua terceira fase, o jornalismo deixa de ser tão livre e comprometido com a realidade exclusivamente, já que agora era condicionado pelas leis de negócio e pelas ideologias de uma burguesia que detinha poder sobre a máquina jornalística, enquanto a "audácia e a criatividade jornalística perdem terreno em relação ao conformismo e à repetitividade mercadológica." (Idem, p. 33).

Após a década de 1970, é que o jornalismo entra em sua quarta fase, na era tecnológica, na qual a informação passa a ser operacionalizada de forma eletrônica e digital. Nesse cenário, a virtualização do trabalho do jornalista é colocada como um grande desafio ao exercício da profissão e a televisão surge como uma potência industrial no jornalismo. Esse é o momento em que o telejornalismo supera qualquer outra atividade jornalística: trucidando o radiojornalismo e o jornalismo impresso.

O quarto jornalismo é o último estágio apresentado por Marcondes Filho (2002) em sua obra **Comunicação e jornalismo: a saga dos cães perdidos**, que compreende desde os anos de 1970 até meados dos anos 2000, quando o autor concluiu o texto que viria a tornar-se livro. É um período marcado pela televisão, mas principalmente pela invenção da internet e da web. O autor define essa como a era tecnológica, fortemente marcada pelas inovações no ramo da tecnologia e eletrônica, pelo aumento na velocidade de transmissão de informações, pela virtualização dos saberes e pela crise da imprensa escrita. Nessa etapa, Marcondes Filho já inicia algumas incursões sobre o papel do jornalista na era tecnológica, embora não o faça de modo a contemplar a realidade do jornalismo no século XXI. Em suas inferências, já sinaliza alguns dos desafios impostos à profissão pelo avanço da virtualização e dos sistemas de automação dentro das redações. Evidentemente, por tratar-se de uma obra que foi lançada no início dos anos 2000, o autor não chega a abordar de maneira aprofundada alguns temas que ganharam destaque em anos seguintes.

Na impossibilidade de avançar nas fases do jornalismo após os anos 2000, recorreremos, então, a outros autores que pesquisaram sobre o jornalismo na contemporaneidade. Uma delas

é Krishma Anaísa Coura Carreira, pesquisadora brasileira que, em 2017, defendeu sua dissertação de mestrado intitulada **Notícias Automatizadas: a evolução que levou o jornalismo a ser feito por não humanos**, no programa de Pós-Graduação em Comunicação Social, da Universidade Metodista de São Paulo (Umesp). No texto, a autora sinaliza o surgimento de uma nova fase no jornalismo, que chamou de "quinto jornalismo", trazendo informações sobre o período pós anos 2000 que, de forma bastante alinhada com as proposições de Marcondes Filho (2002), podem servir como uma atualização de sua obra.

5.4 O QUINTO JORNALISMO

Neste tópico, trazemos importantes contribuições de Carreira (2017) para pensarmos o jornalismo em uma fase que vai além daquelas descritas por Marcondes Filho (2002). Em seu texto **Notícias Automatizadas: A evolução que levou o jornalismo a ser feito por não humanos** (2017), a autora desdobra conceitos que convergem com as proposições que levantamos neste estudo. Além disso, não faltam pesquisadores, no Brasil, que também se dedicaram a observar e buscar compreender o fenômeno da "era tecnológica". A seguir, abordaremos algumas dessas visões até chegarmos ao que convencionou-se chamar de "jornalismo de automação".

Para Carreira (2017), no período em que vivemos, as fronteiras entre o que é jornalismo, ciência da computação e inteligência artificial foram diluídas. Isso porque, agora, a tecnologia funde-se com o jornalismo de tal forma, que é possível criar sistemas que produzem o que até os anos 2000 era algo impensado: "a eliminação da necessidade da presença humana para apurar, analisar, redigir e distribuir alguns tipos de notícia" (CARREIRA, 2017, p. 95). A autora destaca que, na era corrente, o jornalismo passou a não mais ser feito por agente humanos, e sim por sistemas de automação que "imitam" a atuação de um jornalista humano de maneira quase que perfeita. Além disso, traz exemplos de experimentos nos quais notícias redigidas por jornalistas humanos e jornalistas robôs eram apresentadas para grupos distintos de leitores. Os resultados demonstram que, mesmo os leitores mais atentos, não conseguiam distinguir diferenças entre as notícias criadas por sistemas em comparação às criadas por humanos (CARREIRA, 2017, p. 169-172).

Os experimentos demonstram que as funções e atividades que eram feitas exclusivamente por jornalistas e repórteres, agora, são facilmente executadas por programas de computador que, além de trabalhar de forma ininterrupta e em alta velocidade, conseguem, segundo a autora, elaborar até milhares de narrativas jornalísticas por segundo e, inclusive,

disponibilizando conteúdos de forma personalizada, de acordo com o perfil de cada usuário que irá consumir a notícia. O nível de credibilidade e de competência também são fatores determinantes nos experimentos apresentados pela autora, pois os resultados das pesquisas demonstram que há pouca (ou nenhuma) diferença entre as percepções do público (e, em alguns casos, os resultados mostraram pontos mais favoráveis ao software em relação ao nível de credibilidade e competência das notícias apresentadas).

Assim, as funções que eram feitas exclusivamente por um ser humano podem ser, agora, executadas por um programa de computador, que trabalha de forma ininterrupta e em alta velocidade. Esses algoritmos conseguem elaborar centenas e até milhares de versões das mesmas histórias em pouco tempo e de acordo com o perfil de cada consumidor de notícias. O cenário da automação no jornalismo traz, por consequência, um movimento de obsolescência do repórter humano, causando uma ruptura na profissão. Em sua dissertação, a autora defende que "esta ruptura é tão drástica que justifica a defesa de uma nova fase, aqui denominada de Quinto Jornalismo." (CARREIRA, 2017, p. 95).

No quinto jornalismo, os sistemas de automação de notícias são o ponto central e a característica mais proeminente dessa nova fase é a incerteza sobre o futuro do jornalismo que, em épocas passadas, obteve glória e foi considerado como "O Quarto Poder". Mas ela não se limita a apenas incertezas: há também que se levar em consideração aspectos como as relações entre os públicos envolvidos no ecossistema jornalístico (empresas, jornalistas, leitores), os avanços técnicos (como o lançamento de novas ferramentas de inteligência artificial, que ocorre de forma vertiginosa, como é o caso do novíssimo ChatGPT, que foi lançado em novembro de 2022 e, em abril de 2023, já conta com 4 novas versões atualizadas), os aspectos econômicos e sociais que se relacionam com essa nova fase, o ensino do jornalismo levando em consideração suas rápidas mudanças, entre outras questões que complexificam ainda mais a atividade jornalística na era contemporânea. Para Carreira (2017), "neste novo período, outras inovações tecnológicas podem alterar a forma do jornalista fazer matérias, a estrutura do conteúdo e das organizações, assim como as relações com todos os públicos envolvidos no ecossistema jornalístico" (p. 95)

Nas etapas anteriores, o jornalismo enfrentava desafios como a concorrência entre conglomerados de mídia, a difusão das ideologias dominantes e outros atores como veículos independentes. No quinto jornalismo, afirma Carreira (2017), a concorrência é com as máquinas, que "pensam" e produzem notícias de maneira completamente automatizada, utilizando de recursos da inteligência artificial para se desenvolverem. A autora enfatiza que, neste cenário, que mais parece um roteiro de filme futurista,

As funções que eram feitas exclusivamente por um ser humano podem ser, agora, executadas por um programa de computador, que trabalha de forma ininterrupta e em alta velocidade. Esses algoritmos conseguem elaborar centenas e até milhares de versões das mesmas histórias em pouco tempo e de acordo com o perfil de cada consumidor de notícias. Eles têm sido usados em grandes empresas de jornalismo como a Associated Press, Forbes e Le Monde na produção de notícias essencialmente estruturadas em dados. (CARREIRA, 2017, p. 97).

Parece que, de fato, os ideais cibernéticos plantados pelos tecnófilos dos anos 1960 encontraram, nos anos 2000, sua época de colheita. Mais de 50 anos após as primeiras reuniões cibernéticas, estamos colhendo os frutos dos anseios de um grupo de pesquisadores que idealizava uma sociedade permeada por sistemas inteligentes.

Os avanços no ramo da automação devem-se, em grande medida, aos esforços de programadores e cientistas da computação na criação de linguagens inteligentes e que se retroalimentam, como veremos no capítulo a seguir. Para Carreira (2017), "com a automação, o software passou a fazer funções até então feitas apenas por um ser humano, como a redação de notícias." (p. 97). Sendo assim, a automação

Está sendo usada para enfrentar as consequências da expansão das tecnologias digitais interconectadas, como a crise do modelo de negócios do jornalismo, que causou o fechamento de publicações tradicionais e que continua longe de uma solução. E também como forma de atender as crescentes exigências de customização, velocidade na transmissão de notícias, de atualização constante e ampliação da cobertura. Os algoritmos que geram textos automatizados resultam de diversos avanços tecnológicos, como a Inteligência Artificial e a Geração de linguagem natural. (Idem).

A utilização dos sistemas de produção de notícias dentro das redações, no entanto, implica em novos desafios tanto para os profissionais, quanto para as empresas jornalísticas. Do lado dos profissionais, surgem cada vez mais questionamentos sobre o futuro da profissão: quem nunca se perguntou se a profissão de jornalista ainda existirá daqui a algum tempo? Evidentemente, os sistemas de automação já mostraram que são capazes de fazer o trabalho que antes era feito apenas por jornalistas e repórteres. Já a visão do lado da empresa jornalística, que há muito tempo vinha enfrentando problemas financeiros, é a de que provavelmente adquirir uma licença de um software é mais barato do que manter meia dúzia de jornalistas empregados, assinando carteira de trabalho e pagando benefícios aos seus colaboradores. É viável, ainda, manter redações funcionando? Realizar pagamentos de recursos humanos, estrutura física, equipamentos? Um software não seria mais barato, eficiente e rápido?

Essas são indagações que surgem quando paramos para refletir sobre esses aspectos do "quinto jornalismo", proposto por Carreira (2017). Embora sejam questionamentos extremamente úteis, não vamos, nesta tese de doutoramento, ousar buscar as respostas. Como

qualquer movimento ou fenômeno, pensamos que as respostas se farão no caminho. E possivelmente, será papel da academia buscar não apenas mais respostas, mas também novos problemas de pesquisa relacionados ao tema.

Carreira não denuncia a extinção da figura do jornalista humano, mas evoca que, nesta nova era, jornalistas humanos e não-humanos (robôs) coexistirão dentro das redações, cada um com uma responsabilidade diferente:

A introdução dos softwares de IA sinaliza uma reestruturação nas rotinas produtivas das redações e uma divisão entre as matérias que são próprias para os algoritmos e outras que demandam jornalistas em função de narrativas mais sofisticadas, que exigem contextualização, sensibilidade e, porque não dizer, de humanidade. (CARREIRA, 2017, p. 98).

A autora, como indicamos no início deste tópico, buscou criar uma nova conceituação para a fase atual do jornalismo, cunhando o termo "quinto jornalismo" na tentativa de avançar e contribuir para o conhecimento já publicado por Marcondes Filho (2002) no início do século XXI. Dessa forma, podemos observar que Carreira (2017) buscou, efetivamente, contribuir para o avanço na construção de conhecimento sobre teoria do jornalismo, quando, na verdade, poderia apenas exercer um papel de crítica em relação ao conhecimento ultrapassado apresentado no quadro de referência.

Por essa razão, a autora traz os mesmos elementos indicados por Marcondes Filho (2002) em seu quadro com as "Características de cada fase do jornalismo" (Figura 12). Para tanto, Carreira (2017) adiciona mais uma linha ao quadro original, inicialmente proposto por Marcondes Filho (2002). Em seu acréscimo, a autora enfatiza que, na quinta fase, o jornalismo é permeado pela participação dos sistemas de automação no processo de produção das notícias. Quanto aos valores jornalísticos dominantes, a autora destaca os seguintes:

- Personalização excessiva;
- Interação;
- Maior velocidade na produção de notícias;
- Eficiência como valor central;
- Atualização constante;
- Ampliação da cobertura.

Sua maior contribuição é no que diz respeito aos aspectos funcionais e tecnológicos e à economia dessa nova fase do jornalismo. Para a autora, os aspectos funcionais e tecnológicos que merecem maior destaque nesta quinta fase são os algoritmos de inteligência artificial, que

atuam na apuração, redação e distribuição das notícias, num processo totalmente automatizado e sem a intervenção do jornalista humano. Além disso, destaca também as novas rotinas produtivas e sua relação com as novas tecnologias emergentes (como a realidade virtual, câmeras 360°, etc.).

FIGURA 13 - CARACTERÍSTICAS DE CADA FASE DO JORNALISMO COM ACRÉSCIMO DO QUINTO JORNALISMO

	Tipo	Época	Valores jornalísticos dominantes	Aspectos funcionais e tecnológicos	Agentes	Economia
Pré-história	artesanal	1631 a 1789	Espectacular, singularmente novo (desastres, mortes, seres deformados, reis, etc)	Jornal ainda semelhante ao livro, poucas páginas	Empreendedor isolado	Elementar
Primeiro Jornalismo	Político- literário	1789 a 1830	Razão (verdade, transparência); questionamento da autoridade; crítica da política; confiança no progresso	Profissionalização; surge a redação; diretor separa-se do editor; artigo de fundo; autonomia da redação	Políticos; escritores; críticos; cientistas	Economia deficitária
Segundo Jornalismo	Imprensa de massa	1830 a +- 1900	O furo; a atualidade; a neutralidade; criam-se a reportagem, as enquetes, as entrevistas, as manchetes; investe-se nas capas, logo e chamadas de la. página	Rotativas e composição mecânica por linotipos (1890); telégrafo e telefone; cria-se a agência Havas; mais publicidade e menor o peso de editores e redatores; títulos passam a ser feitos pelo editor	Jornalistas profissionais	Economia de empresa: jornal tem que dar lucro; aumento das tiragens: 35 para 200mil
Terceiro Jornalismo	Imprensa-monopolista	de +- 1900 a +- 1960	Grandes rubricas políticas ou literárias; páginas-magazines: esporte, cinema, rádio, teatro, turismo, infantil, feminina	Influência da indústria publicitária e das relações públicas; uso da fotografia	Jornalistas, publicitários e relações públicas promovem indústria da consciência	Grupos monopolistas dominam a imprensa; época de tiragens-monstro
Quarto jornalismo	Informação eletrônica e interatividade	de +- 1970 até o presente	Impactos visuais; velocidade; transparência	Implantações tecnológicas (barateamento da produção); alteração das funções do jornalista; toda a sociedade produz informação	Redes/sistemas informatizados; pessoas em interface; jornalistas prestadores de serviço	Financiamentos migram para a TV e a Internet; crise da imprensa escrita
Quinto Jornalismo	Jornalismo de Imersão e Notícia Automatizada	Entre + ou - 2010 até os dias atuais	Personalização excessiva, interação, maior velocidade na produção de notícias, eficiência como valor central, atualização constante, ampliação da cobertura.	Algoritmos de IA apuram, redigem e distribuem notícias (sem jornalistas), divisão de notícias (feitas por algoritmos e/ou por jornalistas), novas funções de jornalistas (preparo dos algoritmos), novas rotinas produtivas, realidade aumentada, realidade virtual (exigem novas habilidades dos jornalistas), câmera 360o	Jornalistas, softwares de automação, jornalistas independentes, amadores, ampla presença da área de TI nas empresas jornalísticas, novos empreendedores.	Automação como estratégia para enfrentar a crise do modelo de negócios.

Fonte: Marcondes Filho (2002) com contribuições de Carreira (2017)

Outra observação pertinente é sobre a Economia nesta nova fase do jornalismo, destacada por Carreira (2017) como a fase dos investimentos em automação como estratégia para enfrentar a crise do modelo de negócios do jornalismo. Charron (2019) também discute essa questão quando afirma que a crise que vivemos na contemporaneidade não é do jornalismo como prática (embora a automação apresente reais desafios em relação a isso), mas sim, de seu financiamento.

Em entrevista concedida à Lívia Guilhermano ao periódico Intexto, publicado em maio de 2019, o pesquisador canadense Jean Charron relata que

a crise do jornalismo é uma crise real no sentido de que o jornalismo é uma instituição e que não sabemos mais como fazê-la funcionar, como financiá-la. Isso provoca uma

situação de crise. Mas não é o antijornalismo. Não é que não sabemos fazer o jornalismo. Nós sabemos como fazer e o público ainda quer. A pergunta é como organizá-lo para que seja viável financeiramente. (GUILHERMANO, 2019, p. 10).

De fato, no quinto jornalismo surgem diversos aspectos que contribuem para um novo pensamento em relação não apenas à prática profissional, mas ao financiamento do jornalismo enquanto instituição capitalista e a outros aspectos, como o ensino do jornalismo, a formação de novos profissionais, o papel do repórter humano, a qualidade das notícias, às *fake News*, etc. Mas, mais do que isso, o quinto jornalismo invoca, acima de tudo, que admitamos uma nova era, na qual preceitos utilizados à exaustão no passado caem por terra. Uma era jamais imaginada: A era dos repórteres-robôs.

6 CAN YOU TELL WHICH ONE OF THESE STORIES WAS WRITTEN BY A ROBOT?

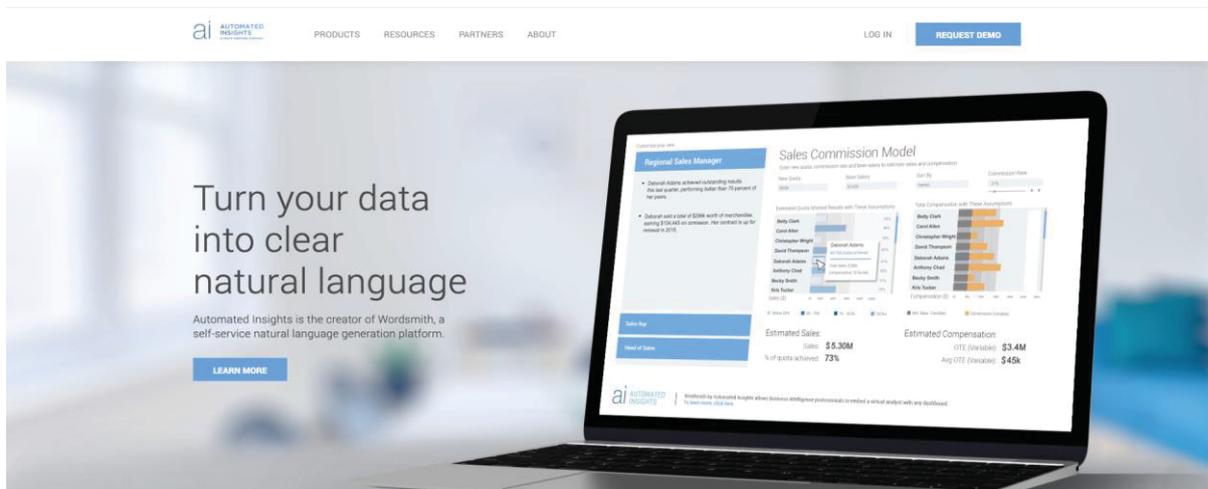
Jornalismo é: automação.

O jornalista Michael Nunes inicia seu artigo "*Can You Tell Which One of These Baseball Stories Was Written By a Robot?*", para a *Gizmodo*, fazendo uma constatação: "é apenas uma questão de tempo até que os robôs substituam todos os empregos, certo? Bom, hoje eles estão vindo para o meu." (Gizmodo, on-line⁵⁰). No texto, o repórter informa que, naquele dia, 30 de junho de 2016, a Associated Press estava iniciando o uso de um software de automação para a cobertura dos jogos de beisebol. Nunes apresenta, no artigo, dois trechos de notícias: um criado pelo software e outro, por um jornalista humano. Discute, logo em seguida, sobre a qualidade das duas narrativas, enfatizando que ambas atingiram o seu objetivo: informar sobre uma partida de jogo.

Algum tempo depois, em fevereiro de 2019, a jornalista Jaclyn Peiser publicou, no portal do *The New York Times*, um artigo intitulado "*The Rise of the Robot Reporter*"⁵¹. Já no primeiro parágrafo do texto, a autora alerta para o fato de que jornalistas vêm perdendo espaço no mercado de trabalho, enquanto as máquinas e sistemas parecem entrar em ascensão nesse contexto profissional. No texto, a jornalista revela alguns números relevantes para a discussão sobre o tema: à época, cerca de um terço de todo o conteúdo publicado pela *Bloomberg News* passava por algum procedimento de automação, e a *Associated Press* (A.P.) já automatizava editorias inteiras de seu conteúdo, como as notícias sobre a bolsa de valores, os relatórios financeiros (com cotações, mercado de ações, previsões financeiras, fechamento da bolsa de valores, etc.), os resultados esportivos (campeonatos, vitórias, dados sobre times, novos jogos de acordo com os resultados anteriores, escalações, etc.) e os relatórios de previsão do tempo (mapas meteorológicos, previsão para a semana, alertas de mudanças abruptas de temperatura, etc.) com o auxílio de sistemas desenvolvidos por uma empresa americana de desenvolvimento de softwares de geração de linguagem natural, a *Automated Insights*:

⁵⁰ Artigo na íntegra: <https://gizmodo.com/can-you-tell-which-one-of-these-baseball-stories-was-wr-1782892240>

⁵¹ "The Rise of the Robot Reporter", by Jaclyn Peiser. Disponível em: <<http://bit.ly/NYTimesJP>>

FIGURA 14 - CAPTURA DE TELA DO SITE DA *AUTOMATED INSIGHTS*

Fonte: <https://automatedinsights.com/>

Com a frase "*turn your data into clear natural language*", que em português pode ser traduzida para "transforme seus dados em linguagem natural clara" é uma alusão crítica aos sistemas antigos de geração de linguagem natural que, em vez de criarem narrativas automatizadas de forma clara e coerente, acabavam retornando para os usuários textos sem nexos, com diversos problemas de coerência em sua estrutura narrativa. Em todo o mundo, empresas como a Automated Insights estão desenvolvendo sistemas com a utilização de inteligência artificial para criar textos de maneira totalmente automatizada, com base em dados (não linguísticos, estruturados e não estruturados) disponíveis na web, transformando-os em textos contextualizados e no formato estilístico do jornalismo (obedecendo a estrutura do *lead* e a consulta a fontes oficiais, por exemplo).

No artigo do *New York Times*, que citamos acima, Jaclyn Paiser destaca a fala de uma de suas entrevistadas, a gerente geral de *machine learning* da Cloudera (empresa de gerenciamento de dados), Hilary Mason. Segundo Hilary, seu desejo seria de ver a inteligência artificial habilitada na capacidade de encontrar pistas e aprofundar a contextualização das notícias: "Espero ver ferramentas de inteligência artificial tornando-se ferramentas produtivas na prática de relatar e encontrar pistas" (tradução nossa)⁵², relata. Naquele momento, as empresas que já apostavam nas tecnologias de linguagem gerada por software garantiam que os postos de trabalho dos jornalistas manteriam-se inalterados, já que a possibilidade de automatizar alguns processos poderia abrir portas para que jornalistas reservassem tempo e

⁵²Texto original: "I hope we'll see A.I. tools become a productivity tool in the practice of reporting and finding clues" <nytimes.com/2019/02/05/business/media/artificial-intelligence-journalism-robots.html>

fôlego no esforço criativo e investigativo. Esse discurso, no entanto, mais parece uma tentativa de colocar "panos quentes" em cima de uma situação impossível de ser ignorada: o iminente declínio de postos de trabalho dentro das redações jornalísticas.

Não seria necessário, aqui, trazer dados que comprovem essa situação, pois o conhecimento empírico já escancara o problema. Basta conversar com qualquer jornalista, repórter ou colega de redação para percebermos que, há mais de uma década, são cada vez mais escassos os postos de trabalho nesse setor. E, aqueles que tinham seus empregos garantidos, viram, um a um, seus colegas de redação serem demitidos ao longo dos últimos anos. Jornais fechando. Edições impressas deixando de circular. Até mesmo portais de notícias, por não encontrarem uma forma consistente de autofinanciar-se, foram deixando de existir. Rádios foram desligando seus sinais⁵³. E, embora a experiência comprove os fatos, trouxemos alguns dados que corroboram a situação:

Em 2021, segundo dados do portal Poder360, 79 veículos de comunicação nacionais foram fechados no Brasil. Segundo o mesmo mapeamento realizado pelo portal, 7 veículos internacionais de grande porte fecharam suas sedes no Brasil nos últimos 4 anos, como The Huffington Post, BuzzFeed Brasil e El País.

⁵³ Relato pessoal da autora: o próprio jornal em que a autora desta tese trabalhava, em 31 de agosto de 2022, emitiu nota oficial comunicando o fechamento de sua edição impressa, portal de notícias e rádio FM, demitindo, de uma só vez, cerca de 30 jornalistas sem aviso prévio (sem contar outros colaboradores, como administradores, diagramadores, setor comercial, etc.). O veículo em questão estava atuando desde 28 de novembro de 1935 no Rio Grande do Sul e chegou a ser um Grupo de Comunicação com jornais impressos diários e redações em Chapecó, Pelotas e Carazinho, além da cobertura de suas estações de rádio, que chegava a toda a região norte e noroeste do Rio Grande do Sul, e sul de Santa Catarina. Quase completando nove décadas de existência, o Grupo declarou o fechamento de suas portas.

FIGURA 15 - MAPEAMENTO DOS VEÍCULOS JORNALÍSTICOS QUE FECHARAM NOS ÚLTIMOS 4 ANOS NO BRASIL.

		o que fechou		
		sede no Brasil	versão impressa	versão digital
Cosmopolitan, Jornal Destak, Women's Health, Vice, BuzzFeed News, Huffpost e El País encerraram totalmente as suas atividades em português no país				
2018				
maio	Rolling Stone	●	⊗	●
agosto	Casa Claudia		⊗	
agosto	Elle*		⊗	
agosto	Cosmopolitan	⊗	⊗	⊗
agosto	Boa Forma		⊗	
2019				
maio	Jornal Metro**		⊗	
2020				
março	Jornal Destak	⊗	⊗	⊗
março	Lance!		⊗	
abril	Women's Health	⊗	⊗	⊗
maio	Vice Brasil	⊗		⊗
agosto	Buzzfeed News	⊗		⊗
novembro	Huffpost	⊗		⊗
2021				
março	Diário do Nordeste		⊗	
maio	Época***		⊗	
outubro	Terça Livre			⊗
novembro	Agora		⊗	⊗
dezembro	El País	⊗		⊗

Fonte: Portal Poder360 <https://www.poder360.com.br/midia/ao-menos-17-veiculos-de-midia-fecharam-no-brasil-em-4-anos/>

No artigo que fala sobre os fechamentos desses veículos, o então presidente da Associação Nacional dos Jornalistas, Marcelo Rech, afirma que são um reflexo direto da "crise mundial do jornalismo" e indica grandes corporações de tecnologia, como o Google e o Facebook (hoje, Meta), como bases de fomento para o sucateamento do modelo de negócio jornalístico, especialmente em decorrência de suas plataformas de publicidade on-line (Google Ads e Meta Ads), que "roubaram" anunciantes dos veículos de comunicação tradicionais.

Aqui, vale observar a crise do jornalismo como uma das potencializadoras para a popularização do jornalismo de automação. Como apresentamos acima, na impossibilidade de

encontrar formas rentáveis de financiamento, aos poucos, veículos de comunicação foram fechando suas portas. Aqueles que ainda persistem, reduziram significativamente suas equipes, tornando o trabalho dos jornalistas que restaram ainda mais desafiador. Foi no momento do colapso das empresas jornalísticas que a automação ganhou força.

Consideramos, aqui, o jornalismo de automação como o produto da utilização de sistemas de produção de conteúdo noticioso, sejam eles guiados por *natural language generation* ou outra tecnologia de refinamento de conteúdo. Há, na literatura, autores que consideram o fenômeno sob a perspectiva do jornalismo-robô (LINDÉN, 2018; LINDÉN & DIERICKX, 2019), utilizando-se da metáfora visual de um robô escrevendo em um teclado de computador para demonstrar, figurativamente, como se dá o processo de automação do jornalismo, como se o jornalista "humano" tivesse seu lugar, efetivamente, ocupado por uma espécie de autômato, dotado de inteligência (mesmo que artificial).

FIGURA 16 - CAPA DO ARTIGO DE LINDÉN SOBRE A METÁFORA DOS JORNALISTAS-ROBÔS



Fonte: Revista eletrônica do Programa de Mestrado em Comunicação da Faculdade Cásper Líbero edição 41 <<https://seer.casperlibero.edu.br/index.php/libero/issue/view/51/showToc>>.

Lindén (2018) argumenta que "os robôs produtores de notícias são programas de computador que contêm algoritmos que detalham instruções específicas, em uma ordem determinada." (p. 08). Vale ressaltar que o autor considera esses programas como *sistemas* automáticos de geração de notícias. E destacamos a palavra sistema porque ela permeia nossa discussão neste estudo, invocando desde a teoria matemática da comunicação, proposta por Shannon (1949), passando pela teoria dos sistemas complexos, apresentada por Bertocchi (2016), até chegar à era do quinto jornalismo, na qual esses sistemas de automação das notícias figuram como personagens protagonistas do enredo, como nos apresenta Carreira (2017).

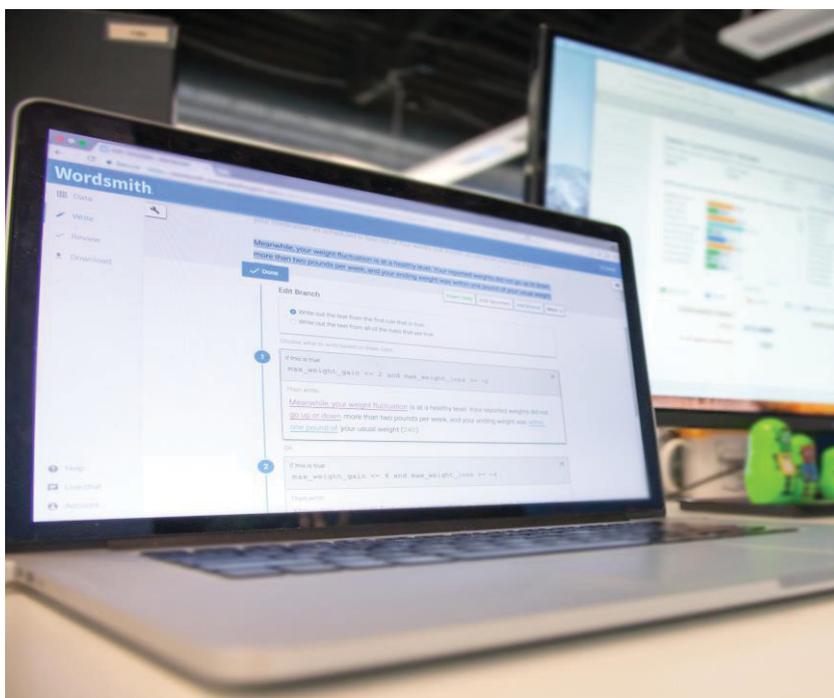
Essa imagem de um autômato trabalhando ou "pensando" por conta própria pode tornar-se uma armadilha. A verdade é que a representação gráfica do que consideramos como "repórter-robô" ou "jornalista-robô" pode ser substituída por uma série de códigos, escritos em linguagem de programação (o que se assemelha muito mais com o que de fato faz o jornalismo de automação) ou pela interface de um sistema de computador (Figura 17). Embora a primeira impressão visual (de um robô em tamanho real e com feições que se assemelham às humanas) seja mais impactante e de fácil assimilação, o verdadeiro visual de um jornalista-robô está mais alinhado com a imagem a seguir: a tela de um computador com o software, alguns botões, um menu de navegação e nada muito "futurístico", como poderia imaginar-se.

Quanto ao funcionamento dos sistemas capazes de coletar dados e gerar notícias automatizadas, Lindén (2018) compara a forma de atuação do sistema com as próprias etapas do trabalho jornalístico: para realizar a produção de uma notícia, o sistema deve ser configurado para desenvolver suas funções de acordo com regras específicas (comandos), assim como o faz um jornalista dentro da redação. Na prática, o dever do desenvolvedor de software é configurar as funções como comandos dentro do sistema:

procure uma informação que atenda a um conjunto predeterminado de regras de produção de notícias, tais como curiosidade, relevância e impacto, procurar três fontes independentes e reconhecidas para a contextualização e comentários, produza um artigo de 200 palavras redigido conforme o manual de estilo da redação e o submeta a um editor ou diretamente ao público. (LINDÉN, 2018, p. 09).

Essas instruções podem ser consideradas como as premissas básicas no trabalho cotidiano dos jornalistas e também podem ser empregadas ao funcionamento dos sistemas de automação de notícias. Assim, Lindén (2018) considera que "o conceito de jornalismo estritamente baseado em regras pode ser aplicado à automação de notícias" (p.09), uma vez que as etapas identificadas dentro do sistema seriam as mesmas efetuadas pelos jornalistas humanos.

FIGURA 17 - INTERFACE DE UM SISTEMA DE AUTOMAÇÃO



Fonte: WordSmith, software de produção de notícias automatizadas da Automated Insights, <<https://automatedinsights.com/wordsmith/>>

Não há, portanto, nenhuma "novidade" no que diz respeito à produção das notícias pelo software, já que ele segue uma série de normativas amplamente utilizadas no mercado jornalístico. São regras e preceitos que nasceram na época do "segundo jornalismo", momento da história em que o jornalismo começou a ser reconhecido como uma profissão, com seus manuais de conduta, regras, sua pirâmide invertida e demais princípios (MARCONDES FILHO, 2002). A grande inovação desses sistemas, no entanto, diz respeito ao uso de *natural language generation* para a criação das narrativas jornalísticas. Como o próprio termo já sugere: são sistemas que buscam imitar a linguagem humana. Se isso não assusta, ao menos, causa curiosidade. A seguir, buscamos compreender como funcionam esses sistemas.

6.1 PODE UM SISTEMA IMITAR A LINGUAGEM HUMANA?

Lindén é um pesquisador finlandês que coordena o Grupo de Pesquisas "*Immersive Automation*" (I.A.), projeto originado na *Swedish School of Social Science* e na *University of Helsinki*, com colaboração de jornalistas e empresas jornalísticas, com o objetivo de realizar pesquisas sobre o fenômeno do jornalismo de automação. O grupo I.A., sob comando de Lindén, foi formado em 2018 e, desde então, diversos estudos vêm sendo realizados no âmbito do jornalismo de automação e publicados em periódicos científicos com impacto internacional. O texto "*Unboxing news automation: Exploring the imagined affordances of automation in news journalism*"⁵⁴, publicado pelo grupo, em 2020, é resultado dessa colaboração entre pesquisadores e apresenta conceituações relevantes sobre o tema. De acordo com Sirén-Heikel et.al. (2020), o conceito de "*automated journalism*" ou "*news automation*" refere-se ao uso de métodos de *natural language generation* (NLG) no domínio das notícias.

Essa explicação é essencial para compreendermos como os autores conceituam o jornalismo de automação por meio de sua aplicabilidade técnica, ou seja, só consideram um produto do jornalismo de automação aquelas notícias que passaram, dentro do sistema jornalístico, pelas etapas referentes ao processamento de linguagem natural. Os sistemas de NLG serão detalhados no tópico a seguir, entretanto, cabe adiantar que esses sistemas baseiam-se em duas formas de *input*⁵⁵ (entrada) e *setup* (configuração) de dados distintas: sistemas com configuração via *rule based*, quando o conteúdo gerado pelo software obedece a determinadas regras pré-estabelecidas no momento do desenvolvimento da ferramenta (e que poderíamos equiparar a um sistema fechado, segundo a conceituação de Bertocchi, 2016); e/ou sistemas com configuração por meio de *machine learning*, quando o software tem a capacidade de reorganizar-se automaticamente e criar uma curva de aprendizado à medida em que é utilizado (que, segundo Bertocchi (2016), poderia encaixar-se no conceito de sistema aberto, adaptável e complexo). Nesse segundo modelo, o sistema guiado por *machine learning* adquire a capacidade de morfogênese, conceito apresentado pela pesquisadora Daniela Bertocchi (2016) em seu estudo sobre os formatos das narrativas jornalísticas digitais. Há, ainda, aqueles sistemas que podem utilizar-se dessas duas aplicações, segundo Sirén-Heikel et.al (2020), e ambas as formas são capazes de gerar notícias automaticamente, tendo como base a configuração da aplicação com princípios de *natural language generation* e pela web semântica.

⁵⁴ "*Unboxing news automation: Explorando as possibilidades imaginadas da automação no jornalismo*", tradução da autora.

⁵⁵ A expressão "input" é largamente utilizada dentro a comunidade de desenvolvedores de software, trata-se de um termo originário do inglês que significa "entrada".

A web semântica, conceito que aparece muito nos estudos sobre jornalismo de automação, pode ser caracterizada como a tentativa de agrupar todos os dados não linguísticos disponíveis na internet (textos, imagens, sons, vídeos e códigos), atribuindo-lhes significados legíveis, tanto por máquinas, quanto por seres humanos. Segundo informações do W3C⁵⁶, o termo “Web Semântica” concede aos usuários "a capacidade de criarem repositórios de dados na Web, construírem vocabulários e escreverem regras para interoperarem com esses dados" (on-line). O W3C considera a web semântica como a “web dos dados”, uma espécie de sistematização de dados dos mais diferentes formatos e tamanhos com o objetivo de disponibilizá-los de forma organizada para que usuários e outros sistemas possam utilizá-los, conforme trecho retirado do site do consórcio internacional W3C: “a Web Semântica é a Web de Dados – de datas, títulos, números, propriedades químicas e qualquer outro dado que se possa conceber. O RDF é a base para a publicação e linkagem de dados” (on-line). Outros conceitos que se somam à web semântica, segundo as diretrizes do W3C, são:

Inferências: o raciocínio sobre os dados a partir das regras, com foco na tradução entre linguagens de regras e intercâmbio entre diferentes sistemas;

Vocabulários e Ontologias: Usando OWL (para construir vocabulários, ou "ontologias") e SKOS (para projetar sistemas de gestão do conhecimento), é possível enriquecer os dados adicionando significados, permitindo que mais pessoas (e mais máquinas) possam reutilizá-los;

Aplicações Verticais: possibilidade de aplicar a web semântica em outras áreas do conhecimento, como Saúde, Energia e Governos;

Assim, a web semântica pode ser conceituada como uma sistematização de dados, atribuindo a eles determinados "valores" que podem ser utilizados para diversas funções na web, especialmente em etapas de cooperação entre máquinas e humanos. Isso possibilita aos sistemas criarem conteúdos mais apropriados para o consumo, já que, com o auxílio da sistematização dos dados, torna-se menos oneroso o trabalho de mineração e validação dos dados quando entram num sistema, por exemplo. Segundo Lammel e Mielniczuk (2012), de maneira sucinta, o que a web semântica realiza é a identificação dos significados presentes na rede [...], utilizando-as para criar relação entre as informações presentes na web a coleções de

⁵⁶ O W3C é o World Wide Web (W3C): consórcio internacional no qual organizações filiadas, uma equipe em tempo integral e o público trabalham juntos para desenvolver padrões para a Web. Liderado pelo inventor da web, Tim Berners-Lee e o CEO Jeffrey Jaffe, o W3C tem como missão Conduzir a World Wide Web para que atinja todo seu potencial, desenvolvendo protocolos e diretrizes que garantam seu crescimento de longo prazo. Dados do site oficial do consórcio <<https://www.w3c.br/Sobre/>>

teorias que são disponibilizadas e compartilhadas na própria web (LAMMEL & MIELNICZUK, 2012, p. 182).

Lozano Tello (2020) explica que a ideia da web semântica "é que os dados possam ser utilizados e "entendidos" por computadores sem a necessidade de supervisão humana, de forma que agentes web possam ser projetados para processar informações localizadas em páginas web de forma semiautomática (LOZANO TELLO, 2020, p. 02 - tradução nossa⁵⁷). Para o autor, a web semântica trata da conversão da informação (um dado, que pode ser de qualquer formato), em conhecimento (linguagem acessível tanto para máquinas, quanto para humanos). Aqui, vale lembrar o conceito de *informação* criado por Shannon, que trouxemos no capítulo 4: informação é uma medida da liberdade de escolha no momento da seleção da mensagem. Pode, portanto, ser considerada uma *unidade* de medida dentro de um sistema de automação.

Com o uso da web semântica, portanto, a busca por informações (dados) e o desenvolvimento de aplicações estão facilitados,

uma vez que as anotações das informações seguirão um esquema comum, e os navegadores da web compartilharão os mesmos esquemas com as anotações da web. As empresas que lidam com clientes e fornecedores podem trocar seus dados de produtos seguindo esses esquemas comuns acordados" (LOZANO TELLO, 2020, p.02 - tradução nossa⁵⁸).

Assim, os sistemas de processamento de linguagem natural encontram na web semântica uma aliada, pois ela figura como uma grande sistematização das estruturas de dados encontrados atualmente na web. Para compreender mais a fundo como a web semântica contribui para o jornalismo de automação, é necessário, também, apresentar outros conceitos e o de *natural language generation* possivelmente é o mais importante deles.

⁵⁷ Trecho original: "La idea es que los datos puedan ser utilizados y "comprendidos" por las ordenadores sin necesidad de supervisión humana, de forma que los agentes web puedan ser diseñados para tratar la información situada en las páginas web de manera semiautomática. Se trata de convertir la información en conocimiento, referenciando datos dentro de las páginas web a metadatos con un esquema común consensuado sobre algún dominio. Los metadatos no sólo especificarán el esquema de datos que debe aparecer en cada instancia, sino que además podrán tener información adicional de cómo hacer deducciones con ellos, es decir, axiomas que podrán aplicarse en los diferentes dominios que trate el conocimiento almacenado" (LOZANO TELLO, 2020, p.02)

⁵⁸ Trecho original: "Con ello, se mejorará la búsqueda de información y se potenciará el desarrollo de aplicaciones de comercio electrónico, ya que las anotaciones de información seguirán un esquema común, y los buscadores web compartirán con las anotaciones web los mismos esquemas. Empresas que traten con clientes y proveedores, podrán intercambiar sus datos de productos siguiendo estos esquemas comunes consensuados" (LOZANO TELLO, 2020, p.02).

6.2 NLG: A LINGUAGEM É ARTIFICIAL, MAS É NATURAL TAMBÉM

O conceito de *natural language generation* já apareceu em alguns parágrafos anteriores, entretanto, como consideramos a NLG um princípio fundamental para compreender como os sistemas de automação de notícias funcionam, avaliamos a necessidade de criar um tópico, neste estudo, dedicado exclusivamente à compreensão do termo. Com os avanços nas pesquisas relacionadas à Inteligência Artificial, a *natural language generation* (NLG) surge como um novo campo de estudos que tem como foco criar narrativas cada vez mais refinadas, a ponto de se assemelharem com a escrita humana. Como conceito, podemos dizer que a NLG busca criar narrativas coerentes a partir de um conjunto de dados (muitas vezes não linguísticos). A NLG alimenta-se especialmente da imensurável quantidade de dados existentes e disponíveis na web e realiza operações de tratamento em cima desses dados para transformá-los em textos legíveis, isto é: o sistema agrupa centenas e milhares de dados (textos, imagens, tabelas, de diversos formatos) e, por meio de uma série de etapas, os transforma em uma narrativa textual estruturada.

Segundo Ribeiro (2019), "além da abundância de informação existente, muitos destes dados demandam um conhecimento especializado para serem interpretados, uma vez que se encontram estruturados para que um sistema computacional os consiga manipular facilmente" (RIBEIRO, 2019, p. 01-02). Em outras palavras, com a utilização da NLG, é possível transformar dados (como um balanço financeiro repleto de informações técnicas sobre o mercado de ações, por exemplo) de uma forma compreensível para o consumo de usuários comuns, como leitores de um portal jornalístico, por exemplo. Esses dados, por sua vez, podem ser de diferentes naturezas: imagens, vídeos, gifs, áudios, gráficos, tabelas, cálculos, textos, documentos, etc. Isso quer dizer que, na web, todo arquivo que carrega uma *informação* pode ser considerado como um *dado*. Para compreender como um sistema como esse pode operar dentro de uma lógica jornalística, podemos pensar da seguinte maneira: imagine que as estações meteorológicas emitem, diariamente, relatórios com a previsão do tempo. Esses relatórios geralmente são representações gráficas e mapeamentos complexos que podem ser disponibilizados na web no formato de documento (PDF), de imagens (JPG, PNG), ou, ainda, em tabelas (XML, CSV). Um sistema de NLG irá coletar esse dado (a uma velocidade de milésimos de segundo), identificará a natureza do formato e, em seguida, dependendo da forma como o sistema é organizado, irá criar um texto com todas as informações presentes naquele arquivo. Esse texto poderá ser escrito de acordo com o manual de redação de um veículo jornalístico e o próprio software também poderá publicá-lo automaticamente, sem,

necessariamente, passar pela revisão de um editor ou revisor humano. Em seu estudo, Ribeiro (2019) traz ainda outros exemplos de como sistemas de NLG podem ser utilizados no jornalismo.

Devyatkin et.al. (2019) demonstram que uma das tarefas mais importantes da NLG é a identificação e posterior conversão dos dados em estruturas gráficas para transformá-las em texto. Para os autores, esta etapa "é baseada na conversão dos dados estruturados não linguísticos mantidos na forma de um gráfico (ou seja, estruturas ontológicas ou semânticas, por exemplo, gráficos RDF que combinam RDF-triplos com elementos com referência cruzada) em um texto coerente" (DEVYATKIN et.al., 2019, p. 01 - tradução nossa)⁵⁹. Ou seja, a NLG é capaz de realizar a conversão de um dado (um arquivo que contém informação) em um texto facilmente compreensível aos humanos e, além disso, também atribui ao dado algum nível de significado, combinando-o com outros dados para criar uma espécie de relação entre eles. Este movimento reproduz o modo de funcionamento descrito por Bertocchi (2016) ao descrever como um sistema recebe influência de outros elementos internos do próprio sistema.

Assim, é possível assumir que os textos podem ser gerados a partir do cruzamento de informações variadas (dados de diferentes formatos convertidos em um único arquivo) e, como um sistema complexo e adaptativo, tem a capacidade de se reorganizar internamente de acordo com os elementos presentes em sua estrutura. Segundo as perspectivas dos autores que consultamos neste estudo, parece um consenso que há algumas etapas fundamentais (para não dizer obrigatórias) que caracterizam o funcionamento de um sistema de NLG. De acordo com sistematizações de Devyatkin et.al., (2019) e Ribeiro (2019), é possível identificar seis passos fundamentais para que um software de NLG funcione de forma adequada: a determinação do conteúdo, estruturação do texto, agregação de frases, a lexicalização, seleção de expressões e, finalmente, a realização da sentença. A seguir (Quadro 3), explicitamos quais são as tarefas executadas pelo sistema em cada uma dessas etapas:

QUADRO 3 – Passos fundamentais de NLG aplicada ao jornalismo (continua)

ETAPA	PROCEDIMENTO
Determinação do conteúdo	Entrada do dado no sistema. O sistema identifica a natureza do dado e deve selecionar quais dados serão utilizados para executar aquela função (texto).

⁵⁹ Texto original: "consists in converting structured non-linguistic data kept in a form of a graph (i.e. ontological or semantic structures, e.g. RDF#- graphs which combine single RDF-triples with cross-referenced elements) to a coherent text".

QUADRO 3 – Passos fundamentais de NLG aplicada ao jornalismo (conclusão)

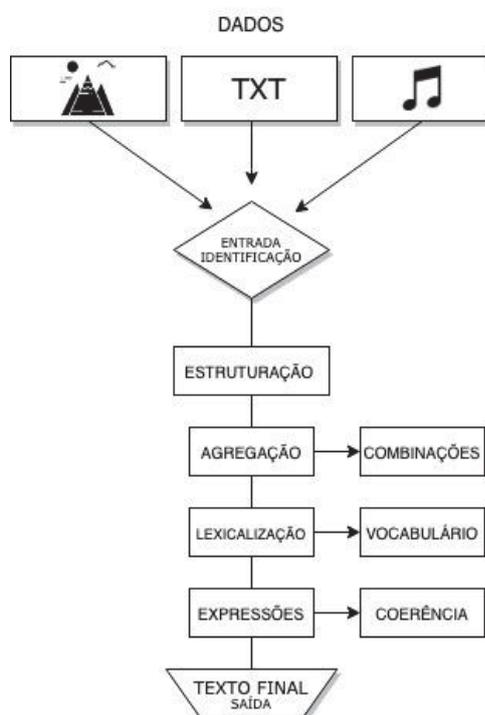
ETAPA	PROCEDIMENTO
Agregação de frases	O sistema decide quais informações serão apresentadas em cada frase.
Lexicalização	Fase na qual o sistema opta por determinadas combinações de vocabulário para tornar o texto compreensível.
Seleção de expressões de referência	Consiste em determinar a maneira apropriada de se referir aos conceitos e objetos contemplados no plano do documento para evitar ambiguidade.
Realização da sentença	O sistema combina as frases para criar textos estruturados.

Fonte: Devyatkin et.al., 2019; Ribeiro, 2019.

Todos esses passos podem ser realizados em sequência ou separadamente, conforme mencionam Devyatkin et.al. (2019) "as abordagens para lidar com essas etapas geralmente são baseadas em uma combinação de regras criadas manualmente e modelos de aprendizado de máquina" (p. 01)⁶⁰. Isso quer dizer que, caso seja necessário, o sistema poderá voltar a uma etapa anterior, ou proceder a uma etapa seguinte, de acordo com a necessidade daquele dado (mas, ainda, não consegue recriar estruturas dentro do sistema utilizando inteligência artificial). Os autores evidenciam que algumas dessas arquiteturas permitem a criação de estruturas completas de NLG de forma não supervisionada, que são capazes de atingir um nível de qualidade relativamente alto nos textos gerados. De acordo com a proposição inicial dos autores, o software de geração de linguagem natural atuaria da seguinte maneira:

⁶⁰ Texto original: "the approaches handling these steps are usually based on a combination of hand-crafted rules and machine learning models" (Devyatkin et.al. 2019, p. 01)

FIGURA 18 - MODELO DE ESTRUTURA DE SISTEMA EM NLG



Fonte: Elaborado pela autora (2022).

Essa arquitetura básica representa um modelo de estrutura de sistema em NLG e surge da adaptação dos conceitos apresentados por Devyatkin et.al. (2019) e por Ribeiro (2019), porém, ela apresenta algumas limitações. Sendo assim, percebemos que ela pode ser aperfeiçoada com o uso de uma cadeia de processos (*pipeline*). De acordo com Ribeiro,

dentro deste quadro, ainda é possível conceber a NLG em termos de um *pipeline*. Um modo alternativo de pensar sobre isso é em termos de uma rede ponderada de múltiplas camadas, onde a geração equivale a um melhor primeiro percurso: em qualquer etapa i , o classificador C_i produz a saída mais provável, que leva à próxima etapa $C_i + 1$ ao longo do caminho mais provável. Essa generalização está conceitualmente relacionada à visão da NLG em termos de políticas na Aprendizagem por Reforço, que define um percurso através de sequências de estados que podem ser organizados hierarquicamente. (2019, p. 20).

Neste novo modelo, que apresentaremos a seguir, visualizamos que a estrutura do sistema não é estática, ou seja, o software pode oferecer diversos caminhos possíveis (etapas) e os dados irão percorrê-los de acordo com a necessidade até se transformarem em uma narrativa. É um exemplo que une as definições de Devyatkin et.al. (2019) e Ribeiro (2019) às proposições de sistemas complexos, abertos e adaptativos de Bertocchi (2016). Em outras palavras, podemos defender que as variações de conteúdo é que irão definir quais caminhos serão

realizados dentro das etapas do sistema. Diferentemente da estrutura apresentada acima, onde os dados devem, obrigatoriamente, passar por todas as etapas estabelecidas no software.

Utilizar uma cadeia de processos dentro do sistema possibilita, ainda, a criação de novas camadas, ou etapas, no processamento dos dados, indicadas como variáveis dentro das funções estabelecidas no sistema (veremos, no capítulo metodológico, o significado dos termos "função" e "variável" dentro de um contexto de programação de software). Por exemplo, entre a entrada e a estruturação do dado, pode-se inserir uma nova etapa de relevância, na qual o sistema deve conferir se os dados selecionados são de fato importantes para aquela narrativa jornalística. Entre a etapa de seleção de expressões e o texto final, é possível inserir as etapas de conferência e validação, para verificar se todas as informações estruturadas são verdadeiras. Em um sistema completo, a etapa de publicação seria o último item, já que o próprio sistema seria capaz de disparar a notícia para o portal da empresa jornalística. Assim, o software seria capaz de oferecer uma solução completa de produção de narrativas jornalísticas: desde a coleta do dado, tratamento, combinação, descarte, estruturação, validação e publicação. A seguir, desenhamos uma possibilidade (das muitas) de arquitetura de software para a produção de notícias utilizando NLG em modo *pipeline*, para exemplificar seu funcionamento:

QUADRO 4 – Passos Fundamentais de NLG em modo *pipeline* aplicada ao jornalismo (continua)

ETAPA	PROCEDIMENTO
Dados	Em seu sentido informacional, um dado é todo o arquivo de informação que é armazenado que pode ser coletado na web;
Entrada	É por onde os dados entram no software, designada, no modelo da teoria matemática do jornalismo proposta neste estudo, como a função " <i>main</i> ". A função <i>main</i> serve como o ponto de partida para a execução de um programa;
Rotina de Coleta	Função que atribui ao sistema o período em que a coleta de dados deve ser realizada (de 1 em 1 minuto, ou a cada 10 minutos, ou a cada 24 horas, dependendo da natureza dos dados coletados);
Classificação do Dado	O sistema irá reconhecer o formato do dado e poderá, inclusive, convertê-lo em outro formato. Por exemplo: um dado em imagem poderá ser lido e transformado em texto;
Validação do Dado	O sistema deverá realizar a checagem do dado a partir de princípios como a veracidade e a relevância do dado;

QUADRO 4 – Passos Fundamentais de NLG em modo *pipeline* aplicada ao jornalismo (conclusão)

ETAPA	PROCEDIMENTO
Lexicalização	Momento em que o sistema deverá executar as funções de estruturação do texto, coerência das frases, lexicalização utilizando as ontologias, a web semântica e <i>machine learning</i> para gerar conexões e combinações entre frases, acontecimentos, palavras, etc. A última etapa da lexicalização pode ser denominada como " <i>HumanFriendlyNews</i> ", conforme apresentaremos no capítulo metodológico, a seguir.
Saída	A narrativa jornalística digital, criada pelo sistema automático, está pronta para ser publicada. Neste momento, o sistema pode disparar a matéria para o editor do jornal, ou publicá-la automaticamente no portal da empresa e compartilhar nas redes sociais digitais.

Fonte: Elaborado pela autora (2022)

O esquema mostrado acima determina que cada etapa do sistema é passível de uma abertura, característica bastante presente nos sistemas abertos adaptativos e complexos, como demonstra Bertocchi (2016). É essa possibilidade que torna o sistema um software de inteligência artificial, pois cada nova etapa faz com que o sistema compreenda melhor seu próprio funcionamento. Nas páginas que se seguem, detalharemos como chegamos a uma proposta de modelo de estruturação de sistema de geração de notícias automatizadas, com base na aplicação de cada uma das etapas em um IDE (*Integrated Development Environment*). Ressaltamos a importância de fazer este caminho teórico até aqui, pois foi necessário compreender as proposições de Ada Lovelace, Claude Shannon, Álvaro Vieira Pinto, Marcondes Filho e todos os autores do jornalismo de automação para compreender o objetivo desta tese: criar um modelo matemático para o jornalismo contemporâneo, que, sob nosso ponto de vista, é o jornalismo fruto da quinta fase proposta por Carreira (2017). Um jornalismo que não se pode comparar ao jornalismo feito há décadas: estamos falando do jornalismo de automação.

Sendo assim, gostaríamos de evidenciar que nosso modelo matemático é, primordialmente, inspirado nos sistemas de automação de notícias. Embora pudéssemos, também, criar teoremas matemáticos apenas tomando como referência o *lead*, por exemplo, optamos por ir além, pois consideramos não apenas uma etapa da produção jornalística, mas

também o fenômeno tecnológico no qual o jornalismo se insere. Por essa razão, sob nosso ponto de vista, não faria sentido criar teoremas para um jornalismo já ultrapassado, então foi realizado este percurso teórico: porque é necessário entender nosso tempo e suas características.

6.3 ELES JÁ ESTÃO ENTRE NÓS

Por ser alimentado pela web semântica e por ontologias, os sistemas que utilizam a *natural language generation* em *pipeline* ainda são capazes de criar textos simples, como o próprio "lead automatizado" proposto por Arce (2009) (que somente informam o leitor de um fato, quanto textos mais complexos), como um texto sobre um episódio de desastre natural, resgatando dados de outros episódios semelhantes (memória) e apresentando relatórios com previsões do tempo (ações futuras e preditivas). Ribeiro (2019) destaca que, "os sistemas NLG podem ser distinguidos em vários tipos dependendo do intuito comunicativo dos textos gerados - textos informativos, textos simplificados, textos persuasivos, sistemas de diálogo, explicações, recomendações, entre outros." (RIBEIRO, 2019, p.07-08).

Os textos informativos seriam aqueles produzidos a partir de dados concretos, como a previsão do tempo ou o resultado de uma partida de futebol, e constituem-se como as narrativas jornalísticas digitais (BERTOCCHI, 2016). Os textos simplificados seriam aqueles destinados a auxiliar usuários da web com problemas de oralidade ou escrita, como é o caso do sistema de *autocomplete* do Gmail, provedor de e-mails do Google, que completa as frases dos usuários de forma automática à medida em que o usuário vai escrevendo o e-mail.

Além disso, textos para auxiliar pessoas com deficiência visual também são gerados por NLG. Já os textos persuasivos, conforme Ribeiro (2019), têm o objetivo de "tentar persuadir ou tirar vantagens do estado emocional daquele que lê, como exemplo: cartas que ajudam a deixar de fumar; informações de saúde personalizadas para ajudar a diminuir a ansiedade de pacientes com câncer" (RIBEIRO, 2019, p. 08). Quando o autor fala de sistemas que geram textos com o objetivo de dar explicações, ele se refere aos passos que o sistema passou para executar um algoritmo, processar uma transação ou resolver um problema matemático. E, por fim, o último modelo de geração de textos elencado pelo autor, seria o de recomendações. Nesse modelo, o sistema cria recomendações através do processamento de informação relacionada com as preferências e opiniões dos usuários. Um exemplo disso são as recomendações de leituras ou viagens personalizadas de acordo com o perfil de cada leitor.

QUADRO 5 – Objetivos comunicativos dos sistemas de NLG

Tipo de Texto	Objetivo do Sistema	Exemplos
Informativos	Gerar textos informativos através de dados concretos	Previsões meteorológicas, relatórios de competências básicas
Simplificados	Ajudar pessoas com problemas de oralidade ou escrita	Produção de texto para pessoas afásicas, produção de texto para pessoas com deficiência visual
Persuasivos	Tentar persuadir ou tirar vantagem do estado emocional do leitor	Cartas para ajudar a deixar de fumar, informações de saúde personalizadas para pacientes com câncer
Diálogo	Melhorar a comunicação entre o utilizador e a máquina	Questões para escrita académica, tutoriais de diálogo para melhorar o conhecimento sobre determinados assuntos
Explicações	Dar explicação dos passos que o sistema passou para executar uma tarefa	Sistema de explicação de provas
Recomendações	Criar recomendações através do processamento de informação relacionada com as preferências e opiniões do utilizador	Sistema de aconselhamento que fornece dietas personalizadas com base no perfil do utilizador

Fonte: Elaborado pela autora, sistematizando conceitos de Ribeiro (2019).

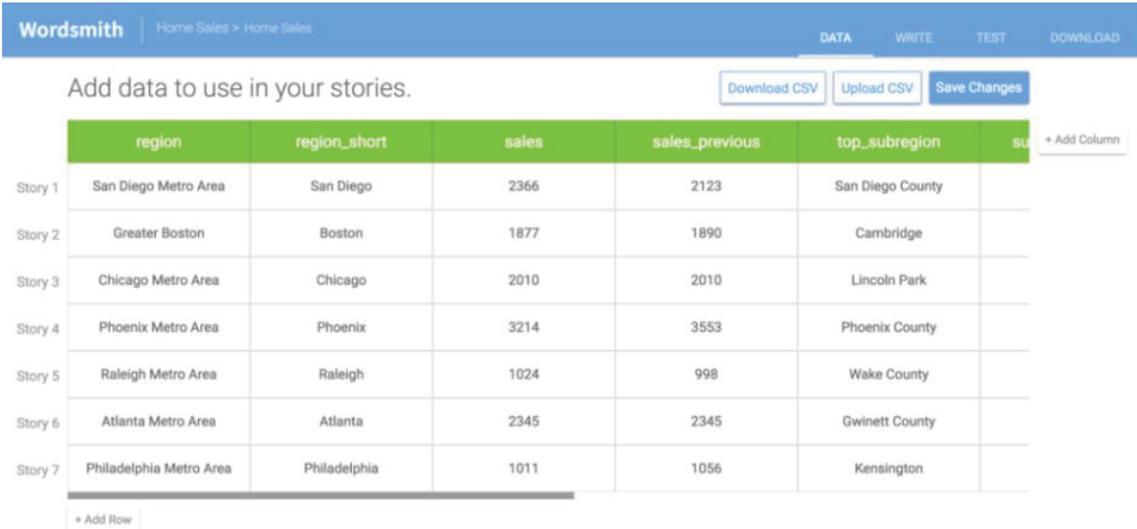
Sendo assim, com os avanços em desenvolvimento de software, e com a ajuda da *natural language generation*, da web semântica e do processamento em camadas (*pipeline*), é possível afirmar que, muito mais do que automatizar o simples *lead* de uma notícia, como sugeria Arce (2009), atualmente os sistemas já possuem a capacidade de criar textos relacionais, contextualizados e verificados (do ponto de vista da validação do conteúdo). No mundo todo, diversas empresas de desenvolvimento de software estão aprimorando seus produtos para entregar soluções cada vez mais refinadas de sistemas em NLG. A seguir, destacamos algumas iniciativas que utilizam NLG atualmente e já obtiveram resultados satisfatórios no mercado jornalístico:

6.3.1. Automated Insights

A *Automated Insights* é uma empresa de tecnologia e desenvolvimento de sistemas norte-americana, que foi criada em 2007 pelo engenheiro Robbie Allen na Carolina do Norte (EUA). Segundo informações do site da empresa, "A *Associated Press* usa NLG para transformar dados brutos de ganhos em milhares de histórias publicáveis, cobrindo centenas de histórias trimestrais a mais do que os esforços manuais anteriores" (*Automated Insights*, online

- tradução nossa)⁶¹. O software utilizado pela *Associated Press* é o Wordsmith, criado pela *Automated Insights*, capaz de gerar centenas de textos jornalísticos em uma fração de segundo. O software foi especialmente configurado para desenvolver textos com o mesmo estilo de redação da *Associated Press*.

FIGURA 18: INTERFACE DO WORDSMITH.



The screenshot shows the Wordsmith interface with a blue header bar containing the logo and navigation tabs: DATA, WRITE, TEST, and DOWNLOAD. Below the header, there is a section titled "Add data to use in your stories." with buttons for "Download CSV", "Upload CSV", and "Save Changes". The main part of the interface is a table with the following data:

	region	region_short	sales	sales_previous	top_subregion	su
Story 1	San Diego Metro Area	San Diego	2366	2123	San Diego County	
Story 2	Greater Boston	Boston	1877	1890	Cambridge	
Story 3	Chicago Metro Area	Chicago	2010	2010	Lincoln Park	
Story 4	Phoenix Metro Area	Phoenix	3214	3553	Phoenix County	
Story 5	Raleigh Metro Area	Raleigh	1024	998	Wake County	
Story 6	Atlanta Metro Area	Atlanta	2345	2345	Gwinnett County	
Story 7	Philadelphia Metro Area	Philadelphia	1011	1056	Kensington	

At the bottom of the table, there is a "+ Add Row" button and a "+ Add Column" button.

Fonte: <<https://automatedinsights.com/blog/new-wordsmith-feature-working-data-easier-ever/>>

O Yahoo!Sports é outro exemplo de empresa jornalística que utiliza o software Wordsmith para a produção de relatórios esportivos personalizados. Esses relatórios são utilizados especialmente no jogo Yahoo! Fantasy, que cria notícias para as centenas de jogadores e times envolvidos nas partidas. Além de gerar softwares para a criação de notícias, a Automated Insights ainda produz softwares para empresas do ramo financeiro, análises de dados e indústrias.

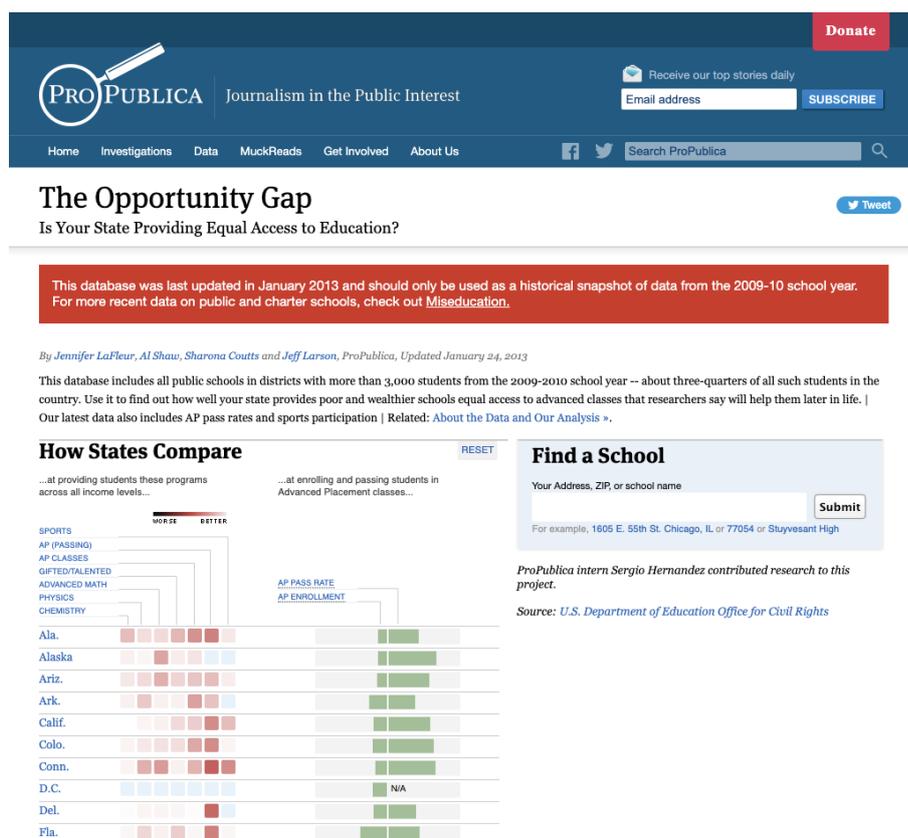
6.3.2. Opportunity Gap

A ProPublica, organização sem fins lucrativos, localizada em Nova Iorque criou uma aplicação capaz de ler e contextualizar dados dos direitos civis do Departamento de Educação

⁶¹ Texto original: "The Associated Press uses NLG to transform raw earnings data into thousands of publishable stories, covering hundreds more quarterly earnings stories than previous manual efforts" (Automated Insights, online)

dos EUA. A aplicação, denominada de *Opportunity Gap*⁶², foi desenvolvida pela própria ProPublica e produziu mais de 52.000 matérias sobre a qualidade das escolas americanas. O site mostrou que alguns estados, como a Flórida, oferecem aos alunos ricos e pobres acesso praticamente igualitário a cursos de alto nível, enquanto outros estados, como Kansas, Maryland e Oklahoma oferecem menos oportunidades às famílias de menor renda.

FIGURA 19 - INTERFACE DO OPPORTUNITY GAP



Fonte: <<https://projects.propublica.org/schools/>>

Ainda, os usuários podem navegar por diversas informações, que são atualizadas em tempo real, com dados de todas as escolas, públicas e privadas, dos Estados Unidos.

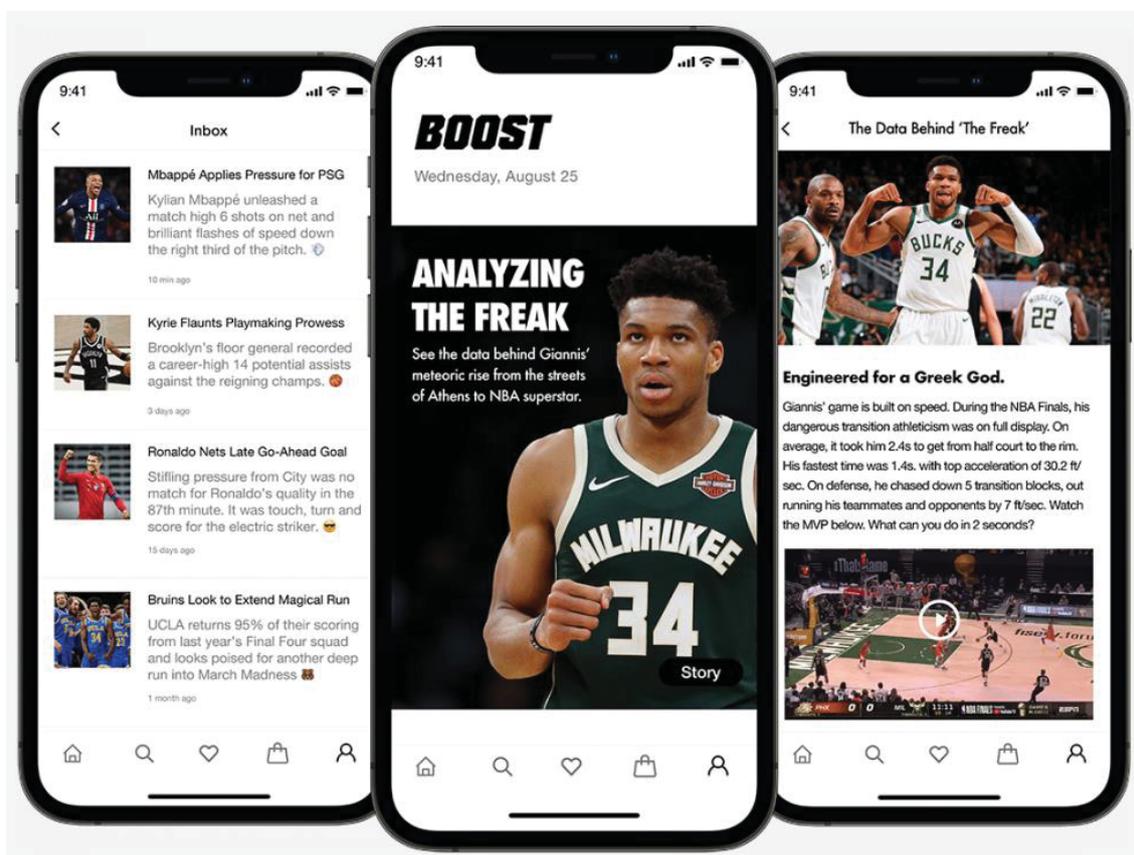
6.3.3. Arria NLG

A Arria NLG é uma das mais populares empresas de desenvolvimento de software de *nature language generation* norte-americanas. Isso porque ela desenvolveu a ideia de criar uma

⁶² ProPublica explica como utilizar o aplicativo: <https://www.propublica.org/article/how-you-can-use-our-opportunity-gap-project-in-your-reporting>

solução completa de NLG para o jornalismo em uma plataforma intuitiva e fácil de manusear. Em 2019, a BBC News contratou a Arria NLG para auxiliar na cobertura jornalística das eleições no Reino Unido, com o objetivo de fornecer aos seus assinantes relatórios completos com estatísticas em tempo real sobre as campanhas. De acordo com a Revista Exame, "com a Arria NLG, a BBC publicou, em dez horas, um montante de 100.000 palavras em 689 textos locais, apresentando em tempo real os resultados de cada um dos 690 distritos eleitorais do país" (EXAME, 2019, online). Segundo Soo Hutton, Engenheiro de Software da BBC News Labs, o software criado pela Arria NLG combina processamento de dados, geração de histórias e aprovação editorial em um processo simples de apenas um clique.

FIGURA 20 - EXEMPLO DE NARRATIVAS DO ARRIA BOOST



Fonte: <<https://www.arria.com/sports/>>

O sistema criado pela Arria NLG é o "Arria Boost", e tem a capacidade de criar conteúdos personalizados. A aplicação funciona com a coleta de dados brutos, gerando automaticamente narrativas jornalísticas com base em modelos projetados por jornalistas.

As três empresas apresentadas acima, embora tenham objetivos diferentes, mantêm em comum o fato de tratarem centenas e até milhares de dados, transformando-os em narrativas jornalísticas facilmente apreendidas pelos leitores. Os softwares, por sua vez, podem ser considerados como "repórteres-robôs", uma vez que centralizam neles todas as funções de um jornalista de redação: recebem a pauta, vão atrás de fontes, estruturam o texto, verificam a veracidade das informações, criam conexões com eventos passados, produzem os textos (falados ou escritos) e, ainda, os publicam automaticamente. Tudo isso, vale ressaltar, em uma velocidade de segundos. São sistemas considerados como abertos, adaptativos e complexos (BERTOCCHI, 2016) à medida em que tratam dados dos mais variados e se retroalimentam, criando rotinas de coletas automatizadas, verificações de acordo com regras pré-estabelecidas e executam funções guiadas por *machine learning*.

Esta pesquisa, que teve início em 2019 e está sendo concluída em 2023, acompanhou de perto a expansão dos sistemas de automação. Em novembro de 2022, acompanhamos o lançamento de um outro sistema, já mencionado nas páginas anteriores, que em poucos meses, revolucionou não apenas o jornalismo, mas diversas áreas e atuações profissionais. Em 30 de novembro de 2022, chegava ao mercado mundial o ChatGPT.

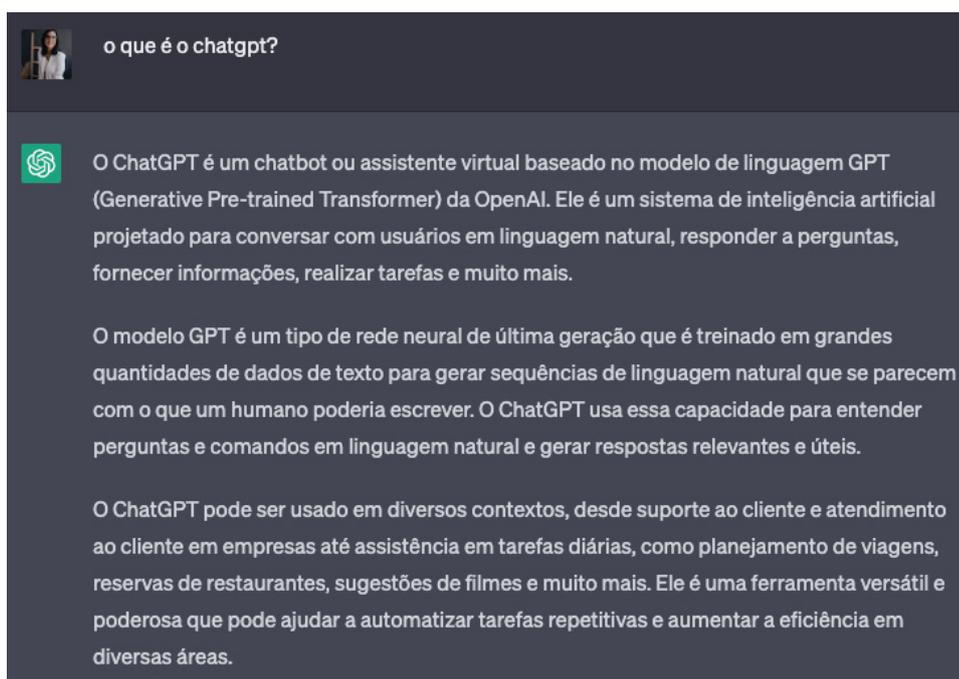
6.3.4. ChatGPT

Embora não seja um sistema de automação jornalístico, o ChatGPT não poderia simplesmente ser ignorado neste estudo. O ChatGPT revolucionou a forma como nos relacionamos com a inteligência artificial, pois temos a impressão de estarmos conversando, de fato, com a ferramenta. As respostas são realizadas em tempo real, embora os resultados nem sempre sejam confiáveis. Além disso, o ChatGPT é gratuito (embora existam versões pagas), o que torna a tecnologia mais acessível, possibilitando a qualquer usuário com acesso à internet fazer uso da plataforma.

Decidimos incluí-lo, mesmo que de forma breve, em nossa pesquisa, porque além de trazer mudanças significativas na forma como nos relacionamos com a inteligência artificial, o sistema utiliza-se de vários conceitos que apresentamos nesta tese, como web semântica e *natural language generation*. O ChatGPT é uma ferramenta de inteligência artificial baseada no modelo de linguagem GPT, que significa *Generative Pre-Trained Transformer*. Esse modelo de linguagem foi criado pela OpenAI, uma organização de pesquisas em inteligência artificial com sede em San Francisco (EUA), e utiliza uma arquitetura de rede neural e *deep learning* para gerar as respostas aos usuários da ferramenta.

A OpenIA foi criada em 2015 por Elon Musk, Sam Altman, Greg Brockman, Ilya Sutskever, John Schulman e Wojciech Zaremba, e tem como objetivo avançar nas pesquisas sobre inteligência artificial. O ChatGPT é uma ferramenta que busca interagir com os usuários por meio de um chat. Para que o sistema funcione de forma adequada, o usuário deve inserir perguntas ou comandos específicos para que a ferramenta retorne uma resposta satisfatória. Para exemplificar o funcionamento da ferramenta, criamos um login na conta e, em seguida, fizemos algumas perguntas. Nosso primeiro comando pedia para que a ferramenta explicasse o que ela é, ou seja, perguntamos para a ferramenta "O que é o ChatGPT?". A resposta foi a de que o sistema é um chatbot ou assistente virtual projetado para conversar com usuários por meio de linguagem natural. Em relação às suas funções, a ferramenta nos indicou que pode "responder perguntas, fornecer informações, realizar tarefas e muito mais", conforme podemos ver a seguir (Figura 21):

FIGURA 21 - INTERAÇÃO COM O CHATGPT

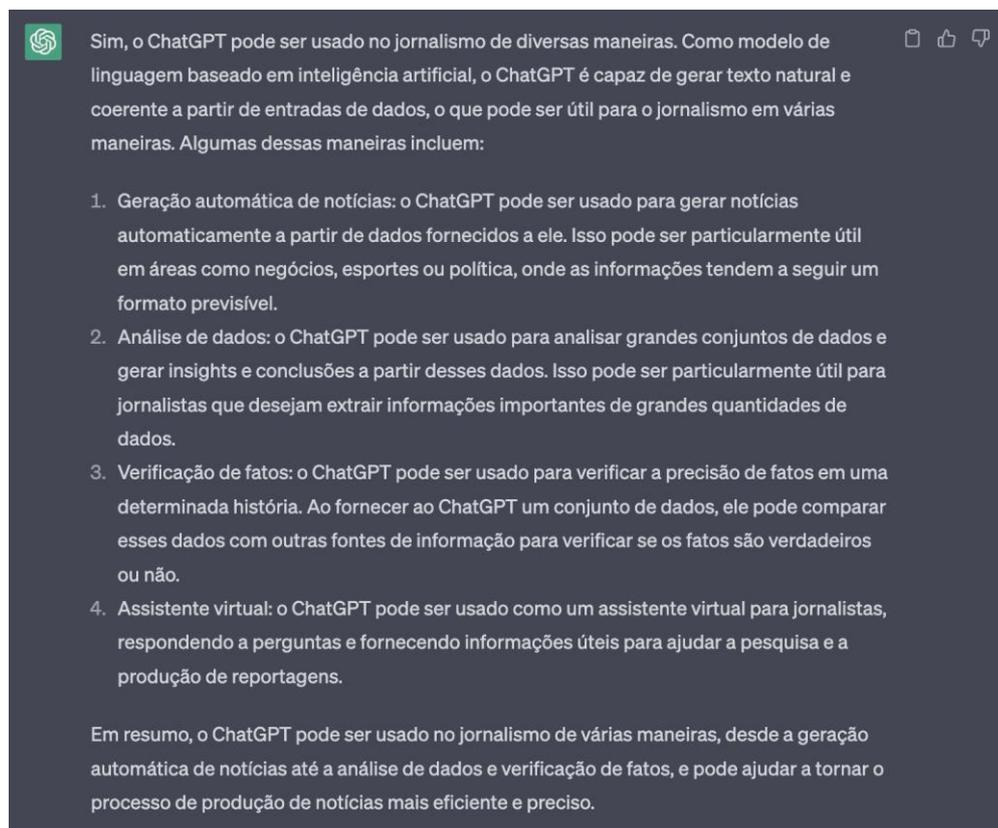


Fonte: <<https://chat.openai.com/>>

Segundo informações do site da OpenIA, o ChatGPT já está em sua quarta atualização. Em abril de 2023, o modelo vigente da plataforma é o GPT-4, único modelo que consegue tratar de eventos em tempo real. Os modelos anteriores tinham algumas limitações em relação aos dados de coleta.

Quando começamos a utilizar o ChatGPT, para entendê-lo melhor e trazer dados empíricos a respeito do uso que os usuários fazem dele, inserimos a pergunta: "O ChatGPT pode ser usado no jornalismo?" A resposta, colocamos abaixo, na íntegra:

FIGURA 22 - COMO O CHATGPT PODE SER USADO NO JORNALISMO



Fonte: <<https://chat.openai.com/>>

Para a ferramenta, os jornalistas poderiam utilizá-la de quatro formas distintas: para a geração automática de notícias, análise de dados, verificação de fatos e como assistente virtual. O que chama a atenção, no entanto, é justamente o primeiro item, no qual a própria ferramenta indica que é capaz de criar notícias, especialmente em algumas "editorias" que têm uma tendência a seguir alguns padrões, como é o caso, por exemplo, de notícias da editoria de esportes, coberturas de jogos/campeonatos, que geralmente obedecem a uma programação predeterminada.

Neste caso, desafiamos a máquina a criar uma notícia sobre um assunto que está nos "assuntos em destaque" do Twitter. Geralmente, estes modelos de aprendizado de máquina conseguem mapear toda a web e identificar os dados relacionados ao assunto em questão. Neste momento, elegemos a seguinte pauta "Presidente Lula visita Primeiro-Ministro Britânico e, em

conversa, Reino Unido anuncia doação de R\$ 500 milhões para o Fundo Amazônia" para desafiar o ChatGPT a criar uma notícia em cima dessas informações.

FIGURA 23 - NOTÍCIA GERADA PELO CHATGPT



Crie uma notícia sobre a pauta: "Presidente Lula visita Primeiro-Ministro Britânico e, em conversa, Reino Unido anuncia doação de R\$ 500 milhões para o Fundo Amazônia"

Notícia:

O ex-presidente brasileiro Luiz Inácio Lula da Silva se encontrou com o Primeiro-Ministro britânico, Boris Johnson, em Londres nesta terça-feira, 5 de maio. Durante a reunião, Johnson anunciou que o Reino Unido fará uma doação de R\$ 500 milhões para o Fundo Amazônia.

A reunião entre Lula e Johnson ocorreu em um momento crítico para a Amazônia, que vem sofrendo com o aumento do desmatamento e das queimadas nos últimos anos. O Fundo Amazônia foi criado em 2008 para financiar projetos de conservação e desenvolvimento sustentável na região.

Em uma entrevista coletiva após a reunião, Lula agradeceu o gesto do Reino Unido e afirmou que a doação ajudará a proteger a Amazônia e a garantir um futuro sustentável para a região. "A Amazônia é um patrimônio da humanidade e precisa ser protegida. A doação do Reino Unido mostra que estamos todos juntos nessa luta", afirmou Lula.

Por sua vez, o Primeiro-Ministro britânico destacou a importância da parceria entre o Brasil e o Reino Unido na luta contra o desmatamento e as mudanças climáticas. "A doação que estamos fazendo hoje é um sinal do nosso compromisso em proteger a Amazônia e ajudar o Brasil a construir um futuro sustentável para a região", disse Johnson.

A doação do Reino Unido para o Fundo Amazônia é vista como um importante gesto de solidariedade e cooperação internacional em um momento em que a região enfrenta desafios ambientais e políticos significativos. A iniciativa deve contribuir para a promoção da conservação ambiental e do desenvolvimento sustentável na Amazônia, ajudando a preservar a biodiversidade, os recursos naturais e o modo de vida das comunidades locais.

Fonte: <<https://chat.openai.com/>>

Constatamos que a ferramenta conseguiu, de fato, criar uma narrativa jornalística, já que, no início do texto, indicou quem estava envolvido, quando aconteceu, o que foi o acontecimento, respondendo algumas das perguntas tradicionais do *lide* jornalístico. Em termos de estrutura, podemos considerar que a ferramenta obteve êxito na resposta. O ponto problemático é que, já na primeira linha do texto, há alguns equívocos e informações erradas. A primeira delas é quando o ChatGPT chama o atual presidente brasileiro, Luiz Inácio Lula da Silva de "ex-presidente". A segunda informação errada é o nome do primeiro-ministro britânico. No texto, o ChatGPT fala em Boris Johnson, que foi, de fato, primeiro-ministro entre

os anos de 2019 até 2022. Atualmente, o político que está no cargo de primeiro-ministro britânico é Rishi Sunak. Ainda, na mesma frase, há outra informação errada: a data.

O encontro entre Lula e Rishi Sunak aconteceu no dia 05 de maio, de fato, como relata a notícia, mas o equívoco é que foi uma sexta-feira, e não uma terça-feira, como indica a ferramenta. Outro ponto que é inconclusivo no texto é que o ChatGPT afirma ter ocorrido uma entrevista coletiva após a reunião entre os dois políticos, colocando, inclusive, algumas falas dos políticos em relação ao assunto. Isso, no entanto, pode ser considerado como um equívoco porque, após a reunião ter sido concluída, o presidente Lula partiu diretamente para outra agenda, uma recepção para convidados que chegaram à Inglaterra para a coroação do rei Charles, no Palácio de Buckingham. Acompanhando a agenda presidencial, pudemos notar que não houve entrevista coletiva após a reunião entre Lula e Sunak.

O que pudemos apreender é que, em termos de estrutura (*lide*, citação de fontes, coerência entre as sentenças, distribuição dos parágrafos, etc.), o texto apresentado pode ser considerado satisfatório. O ChatGPT, de fato, construiu uma notícia. Entretanto, as informações contidas no texto estavam incorretas e algumas delas eram inconclusivas. Isso abre brechas para a discussão a respeito da qualidade das informações geradas pela plataforma, não apenas no exemplo que citamos acima, mas em outras áreas também. Desde seu lançamento até a versão atualizada que estamos utilizando (GPT-4), não faltam exemplos e relatos de usuários que alegam terem usado a ferramenta e ela ter fornecido respostas erradas ou imprecisas⁶³.

Após algumas incursões na ferramenta, e principalmente, após termos desafiado o ChatGPT a escrever uma notícia, podemos concluir que, neste caso, o que chamamos de "repórter-robô" não poderia substituir o repórter humano no processo jornalístico. Isso porque o robô do ChatGPT, mesmo tendo criado uma narrativa jornalística completa, a criou com informações erradas e imprecisas. Neste caso, apenas a título de comparação, reproduzimos (Figura 24) uma notícia publicada no portal G1 (www.g1.globo.com), escrita por Murillo Salviano, correspondente brasileiro em Londres, cuja pauta era a mesma que utilizamos como comando para gerar a notícia pelo ChatGPT.

63 "Os erros do ChatGPT. E as questões quanto à inteligência artificial" em bit.ly/3pfmAu1
"ChatGPT e Inteligência Artificial são incríveis, mas ainda cometem erros bizarros" em bit.ly/3VzUuFI
"Pegadinha, receita de comida, Shakespeare: 4 vezes que o ChatGPT errou" em bit.ly/3LupvpO

FIGURA 23 - NOTÍCIA PUBLICADA EM PORTAL DE NOTÍCIAS

Fonte: <<https://g1.globo.com/politica/noticia/2023/05/05/em-reuniao-com-lula-primeiro-ministro-do-reino-unido-fala-em-doar-r-500-milhoes-para-o-fundo-amazonia.ghtml>>

O texto completo da notícia do G1, reproduzimos na íntegra a seguir:

O primeiro-ministro do Reino Unido, Rishi Sunak, anunciou, em reunião com o presidente Lula em Londres nesta sexta-feira (5), a doação de cerca de R\$ 500 milhões para o Fundo Amazônia. O encontro entre Lula e Sunak durou cerca de 40 minutos. O presidente brasileiro viajou nesta quinta-feira (4) à Inglaterra, onde neste sábado (6) participa da cerimônia de coroação do rei Charles III. O Fundo Amazônia foi criado há 15 anos com o objetivo de captar recursos para o financiamento de ações de redução de emissões provenientes da degradação florestal e do desmatamento. Na gestão do ex-presidente Jair Bolsonaro (PL), o fundo esteve paralisado em razão de um "revogação" de centenas de conselhos federais, que extinguiu o Comitê Orientador (COFA) e o Comitê Técnico (CTFA) da iniciativa. Ao assumir o mandato no início deste ano, Lula reativou o Fundo Amazônia e países anunciaram a retomada de investimentos. Em fevereiro, o petista se reuniu com o presidente norte-americano Joe Biden e o governo dos Estados Unidos anunciou a intenção de fazer aportes ao fundo.

Agora, gostaríamos de resgatar a mesma pergunta que deu início a este capítulo para finalizá-lo com uma reflexão: *can you tell which one of these stories was written by a robot*⁶⁴?

⁶⁴ A frase faz alusão a uma matéria publicada no site *Gizmodo*, e sua tradução é "você consegue dizer qual destas histórias foi escrita por um robô?"

7 A MATEMÁTICA PARA O JORNALISMO: DA FILOSOFIA MATEMÁTICA AOS ESTUDOS SOBRE O TEMA

Neste capítulo, primeiro evidenciamos a obra **Introdução à Filosofia Matemática**, de Bertrand Russel. Escrita em 1919, a obra nos ajuda a compreender o pensamento matemático e, assim, explicar a necessidade de uma teoria matemática para o jornalismo no contexto da inteligência artificial. Após a descrição da obra, apresentamos o estado da arte sobre as pesquisas que relacionam a matemática e o jornalismo, com foco na automação de textos.

7.1 LOGICISTAS, UNI-VOS

Escrito dentro de uma cela de uma prisão em Brixton, no Reino Unido, **Introdução à Filosofia Matemática**, de Bertrand Russell (2007), é um livro que nos trará alguns conceitos importantes para arrematar o percurso teórico que construímos nesta tese. Antes, no entanto, de falarmos sobre os conceitos de lógica e matemática, é necessário realizar uma incursão no perfil de seu autor: Bertrand Russell.

Russell pode ser considerado um dos grandes filósofos do século XX e, com Albert Einstein, formou a dupla de "heróis da juventude", de Augusto J. Franco de Oliveira, professor associado da Universidade de Lisboa, responsável pela tradução para o português, da obra *"Introduction to Mathematical Philosophy"*. Na nota do tradutor, que antecede as páginas traduzidas do livro de Russell, Oliveira conta que Russell é um expoente da chamada corrente logicista na filosofia e que uma das principais características dessa corrente de pensamento é a crença de que a matemática pode ser reduzida à lógica simbólica.

Para Russell, um ponto chave de sua obra é a compreensão de que a lógica está para a matemática assim como a lua está para a noite e o sol para o dia. Na visão do filósofo, lógica e matemática são indivisíveis, pois considera-as como uma coisa só, embora a lógica preceda a matemática. Russell utiliza uma metáfora dizendo que as duas "diferem entre si como rapaz e homem: a lógica é a juventude da matemática e a matemática é a maturidade da lógica" (RUSSELL, 1993, p. 191). Segundo o autor, tanto a lógica quanto a matemática vinham sendo estudadas de formas distintas historicamente, pois a "matemática tem estado relacionada com a ciência e a lógica com o idioma grego. Mas ambas se desenvolveram nos tempos modernos: a lógica tornou-se mais matemática e a matemática tornou-se mais lógica" (idem). Como consequência, afirma o autor, tomando como premissa o fato de que ambas são uma só, seria de fato impossível traçar uma linha divisória entre as duas, ou seja, não é possível delimitar

onde inicia a lógica e onde termina a matemática, pois, segundo seu entendimento, elas constituem uma coisa só. Argumenta que

ao começar com premissas que seriam universalmente admitidas como pertencentes à lógica, e chegar, por dedução, a resultados que de modo igualmente óbvio pertencem à matemática, constatamos não haver um ponto pelo qual possa ser traçada uma linha distinta, a separar a lógica à esquerda e a matemática à direita (RUSSELL, 1993, p. 193)

A lógica, para Russell, antecede a matemática no sentido de que o raciocínio é uma operação que vem antes da conta de mais, menos, dividir ou multiplicar. O autor defende que na lógica clássica, os enunciados levam o leitor a concluir algo por meio de uma representação linguística, como por exemplo: "Todos os homens são mortais, Sócrates é um homem, portanto Sócrates é mortal". Nesse exemplo, o enunciado leva o leitor a deduzir, por meio da lógica, que Sócrates é mortal porque todos os homens são mortais e Sócrates é um homem. Russell, no entanto, coloca esse enunciado em termos de uma equação, demonstrando a possibilidade de reduzir a lógica a um modelo matemático, conforme demonstramos a seguir:

podemos substituir *homens* por α , *mortais* por β e Sócrates por κ , onde α e β são classes quaisquer e κ é um indivíduo qualquer. Chegamos então ao enunciado: «Independentemente dos valores que κ , α e β possam ter, se todos os α forem β e κ for um α , então κ é um β ; por outras palavras, a função proposicional 'se todos os α são β e κ é um α , então κ é um β ' é sempre verdadeira. Temos aqui, finalmente, uma proposição da lógica — aquela que é meramente pelo enunciado tradicional sobre Sócrates e sugerida homens e mortais. (RUSSELL, 1993, p. 193-194).

Russell, portanto, defende que enunciados podem ser descritos dentro de uma linguagem lógica. Se colocarmos em um modelo matemático, o enunciado que Russell traz como exemplo poderia ser ilustrado da seguinte maneira:

Enunciado: "*Todos os homens são mortais, Sócrates é um homem, portanto Sócrates é mortal*"

Onde:

Homem = α

Mortal = β

Sócrates = κ

Logo:

se $\alpha = \beta$ e $\kappa = \alpha$ então, $\kappa = \beta$

Nesse enunciado, é possível também invocar outros conceitos que Russell trata em **Introdução à Filosofia Matemática**, como o conceito de ordem e o conceito de indução. Para o autor, a indução é o que leva a compreendermos a sucessão dentro do sistema matemático. Sendo assim, se tomarmos um número η como sendo $= 0$, o princípio da indução irá nos dizer que o sucessor de η é $(\eta + 1)$. Já o conceito de ordem nos lembra que todo o sistema numérico que utilizamos está disposto de forma ordenada. Para o autor, a ordem é uma característica importante dentro da matemática, mas não a reduz a apenas a ordem dos números progressivos (1, 2, 3, 4, 5 e assim por diante), já que a ordem pode ser conquistada também por meio de outras composições dentro do sistema matemático:

Os números naturais — ou os números indutivos como também lhes chamaremos — aparecem-nos mais frequentemente pela sua ordem de grandeza; mas são capazes de uma infinidade de outros arranjos. Podemos, por exemplo, considerar primeiro todos os números ímpares e depois todos os números pares; ou primeiro 1 e depois todos os números pares, a seguir todos os múltiplos ímpares de 3 e então todos os múltiplos de 5 mas não de 2 ou 3, e depois todos os múltiplos de 7 (mas não os de 2 ou 3 ou 5) e assim por diante, através da progressão dos primos. (RUSSELL, 1993, p. 41-42).

A noção de ordem, segundo Russell (1993), tem uma importância enorme dentro da matemática. Para nosso estudo, inclusive, a noção de ordem nos trará o entendimento de que, dentro de um modelo matemático do jornalismo, algumas ações também são feitas de maneira ordenada. Neste estudo, vimos que os sistemas de automação operam de forma ordenada especialmente os sistemas fechados, e também de forma indutiva, quando pensamos em sistemas abertos, adaptativos e complexos como aponta Bertocchi (2016). Sendo assim, Russell argumenta que a importância da ordem vai além da noção de progressão dos números naturais

Não apenas os inteiros, mas também os números racionais e os números reais possuem uma ordem de grandeza, e isto é essencial à maior parte das suas propriedades matemáticas. A ordem dos pontos numa linha é essencial à geometria; o mesmo se dá no tocante à ordem ligeiramente mais complicada das linhas [rectas] que passam por um ponto num plano ou dos planos que passam por uma linha [recta]. As dimensões são, em geometria, um desenvolvimento do conceito de ordem. O conceito de limite, que alicerça toda a matemática superior, é um conceito serial. Há partes da matemática que não dependem da noção da ordem, mas são pouquíssimas em comparação com as partes em que está envolvida esta noção. (RUSSELL, 1993, p. 41).

Em sua obra, o autor busca sintetizar conceitos importantes para a reflexão sobre a matemática. Determinadas proposições sobre a realidade, de acordo com o autor, são assim, passíveis de uma reprodutibilidade lógica. Se recordarmos, no início deste estudo trouxemos uma frase cuja autoria estima-se que tenha sido de Galileu Galilei, na qual o astrônomo defende que "a matemática é o alfabeto no qual Deus escreveu o Universo". Essa frase orienta-nos a

pensar que as ideias propostas por Russell, de fato, podem ser atribuídas a qualquer fenômeno que se possa observar, levando em consideração as suas conceituações, como a de função proposicional, limites e continuidade, classes, conjuntos, indução, axiomas infinitos, tipos lógicos etc. Em sua obra, Russell busca, de fato, introduzir ao leitor comum conceitos que fazem parte da lógica matemática e que, juntos, podem descrever e sistematizar o mundo.

Por fim, unindo todas as conceituações que trouxemos até aqui, nas próximas páginas iniciamos nossa incursão ao proposto nesta tese: a criação de um modelo matemático para o jornalismo, utilizando como base o funcionamento de sistemas de automação e invocando, neste modelo, toda a estruturação lógica para a produção de uma narrativa jornalística. A seguir, iniciamos a jornada com um olhar exploratório para o estado da arte sobre as relações entre jornalismo e matemática.

8 O ESTADO DA ARTE: ESTUDOS SOBRE JORNALISMO E MATEMÁTICA NO CAMPO DA AUTOMAÇÃO JORNALÍSTICA

A fim de identificarmos o estado da arte em relação ao operador teórico específico deste estudo, investigamos, em diversos repositórios científicos, se há trabalhos que tragam à luz temas relacionados ao jornalismo e à matemática. Aqui, utilizamos também o termo "teoria matemática do jornalismo" como palavra-chave. O termo nos interessa, em particular, neste estudo, pois buscamos iniciar a reflexão sobre essa possibilidade teórica, embora não seja o objetivo central desta pesquisa. O estudo que propomos aqui é uma primeira incursão no que, futuramente, poderá configurar-se como uma nova teoria, com seus pressupostos, métodos e operadores próprios.

Na primeira etapa da pesquisa exploratória, fizemos buscas em repositórios de teses e dissertações, livros, artigos e capítulos de livros. Os bancos de dados escolhidos foram: Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), o banco de teses e dissertações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e o Google Acadêmico, que reúne milhares de livros, artigos, relatórios e documentos, em diversos formatos (pdf, html, páginas da web, etc.). Ainda, com o objetivo de aumentar o foco de nossa pesquisa exploratória, investigamos nos periódicos da área da Comunicação melhor avaliados segundo o sistema Qualis Capes, pois compreendemos que são repositórios com regras e critérios específicos, onde os principais pesquisadores brasileiros costumam publicar suas pesquisas.

Após a pesquisa exploratória, iniciamos uma incursão nos sistemas de automação jornalística para, então, propormos um modelo matemático para o jornalismo, inspirado no funcionamento dos sistemas de automação. A pesquisa a seguir, de caráter exploratório, nos mostrou alguns indícios que corroboram para a necessidade de incentivarmos cada vez mais o surgimento de propostas de investigações acadêmicas com foco na união entre o jornalismo e a matemática.

8.1 BASES DE DADOS: BDTD, CAPES E GOOGLE ACADÊMICO

Iniciamos nossa investigação pela Biblioteca Digital de Teses e Dissertações (BDTD). Segundo dados do site da BDTD⁶⁵, a Biblioteca Digital Brasileira de Teses e Dissertações

⁶⁵ Site da BDTD: <<http://bdttd.ibict.br/vufind>>

(BDTD) tem o objetivo de disseminar os textos completos das teses e dissertações defendidas nas instituições brasileiras de ensino e pesquisa. A Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) foi concebida e é mantida pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) no âmbito do Programa da Biblioteca Digital Brasileira (BDB), com apoio da Financiadora de Estudos e Pesquisas (FINEP).

Como repositório de textos científicos, a BDTD trabalha com os preceitos de interoperabilidade e cooperação para reunir, em um único portal, todos os trabalhos com *open access* defendidos nas instituições brasileiras de Pós-Graduação. O modelo do mecanismo de coleta de dados da BDTD é baseado em padrões de interoperabilidade e, segundo informações da página "Sobre", do site da biblioteca, a BDTD constitui-se como uma rede distribuída de sistemas de informação que gerenciam teses e dissertações, com textos completos, com a existência de dois atores principais:

- Provedor de dados (*data providers*) - administra o depósito e publicação, expondo os metadados para a coleta automática (*harvesting*);
- Provedor de serviços (*service providers*) - fornece serviços de informação com base nos metadados coletados junto aos provedores de dados.

Dessa forma, são as instituições de ensino e pesquisa que oferecem os dados para a biblioteca, enquanto a BDTD opera como uma espécie de ferramenta agregadora dos textos, coletando os metadados das teses e dissertações dos seus provedores (instituições) e disponibilizando-as em sua plataforma. Nossa primeira pesquisa na biblioteca foi realizada pelo termo "teoria matemática do jornalismo", sem nenhum operador de modificação de correspondência de pesquisa, como demonstramos a seguir.

FIGURA 24 - PESQUISA DIRETA SEM OPERADORES DE MODIFICAÇÃO

ACESSO E VISIBILIDADE ÀS TESES E DISSERTAÇÕES BRASILEIRAS			
130	563.529	213.806	777.334
Instituição	Dissertações	Teses	Documentos

Fonte: Captura de Tela página inicial do site BDTD <<https://bdtb.ibict.br/>>

A primeira busca apresentou 40 resultados, com textos de instituições brasileiras que datam de 1992 até 2021. Dessa forma, entre os resultados a BDTD mostrou textos publicados entre 1992 e 1999, que foram excluídos do *corpus* de nossa pesquisa, já que a intenção é observar as pesquisas publicadas entre os anos 2000 e 2020, totalizando 20 anos de coleta. Após a exclusão dos textos, exportamos a tabela de resultados para uma planilha do Google Sheets, onde pudemos observar os títulos, autores, palavras-chave, ano de publicação e instituições relacionadas aos textos resultantes da busca.

De todos os textos publicados, identificamos que apenas dois possuem algum grau de relação direta com o jornalismo, embora nenhum dos dois apresente de forma específica a palavra-chave "teoria matemática do Jornalismo" ou "teoria matemática" ou qualquer outro termo que estabeleça relação entre os operadores "jornalismo" e "matemática":

QUADRO 6 – Textos com relação com o tema jornalismo

Título	Autor	Palavras-chave	Instituição
NOTÍCIAS AUTOMATIZADAS: A evolução que levou o jornalismo a ser feito por não humanos	Krishma Anaísa Coura Carreira	Notícia automatizada; Jornalismo automatizado; Inteligência Artificial e Jornalismo; Algoritmos jornalísticos; narrativa automatizada	METODISTA
A ciência do século XIX na visão do jornalista João Ribeiro em artigos de divulgação científica (1895-1934)	Filipe Silva de Oliveira	Ciências João Ribeiro (jornalista) Visões de ciências Jornalismo científico Ensino de ciências	UFS

Fonte: Elaborado pela autora (2023)

Já os textos que apresentam a palavra-chave "matemática" geralmente colocam o termo em relação com pesquisas do ramo da educação e da pedagogia, por exemplo:

QUADRO 7 – Textos com relação com o tema Matemática (continua)

Título	Autor	Palavras-Chave	Instituição
Ações das IES quanto às Dificuldades Matemáticas dos Alunos Ingressantes: Um Retrato	Wilson De Jesus Masola	Educação Matemática Educação Superior Ensino-Aprendizagem de Matemática Dificuldade de Aprendizagem Matemática Alunos Ingressantes	UNICSUL

QUADRO 7 – Textos com relação com o tema Matemática (conclusão)

Título	Autor	Palavras-Chave	Instituição
Remy Freire e as suas Contribuições para a Matemática e a Educação Matemática no Paraná	Antonio Peixoto de Araujo Neto	Matemática - História - Brasil Sociedade Paranaense de Matemática Ciências Exatas e da Terra Matemática	UEM
Handles - A Trajetória de Desenvolvimento de um Jogo Digital para Ensino de Matemática.	Marcos Henrique de Paula Dias da Silva	Ensino De Matemática Jogo Digital Game Design Teaching Math Digital Game	UNESP
Um Estudo Sobre os Saberes Formativos do Formador de Professores de Matemática do Curso de Licenciatura em Pedagogia	Luciana Miyuki Sado Utsumi	Formação Inicial de Professores Educação Matemática Saberes Formativos	METODIS TA
Sobre o Teorema de Bohnenblurt - Hille	Daniel Núñez Alarcón	Operadores Multilineares Operadores Múltiplo Somantes Polinômios Homogêneos Ciências Exatas e da Terra Matemática	UFPB
A Utilização Da Calculadora Hp-12c No Ensino Da Matemática Financeira Visando A Qualificação Profissional	Júnio César Mendes Simão	Qualificação Profissional Matemática Financeira Calculadora Hp-12c Anúncios Publicitários	UFG
O Projeto São Paulo Faz Escola Para O 1º Ano Do Ensino Médio Sob O Olhar Da Teoria Elementar Dos Números	Francisco De Moura e Silva Júnior	Teoria Elementar Dos Números Análise De Conteúdo Jornal Do Aluno Caderno Do Professor Matemática (Ensino Médio) Teoria Dos Números	PUC_SP

Fonte: Elaborado pela autora (2023)

Sendo assim, identificamos apenas sete trabalhos sobre matemática após a nossa pesquisa pelos termos "teoria matemática do jornalismo" e "jornalismo e matemática". De qualquer forma, apenas um dos trabalhos parece tratar de assuntos afins. É o caso do trabalho da pesquisadora Krishma Anaísa Coura Carreira, que, inclusive, é um dos principais estudos que trazemos nesta pesquisa e o utilizamos no referencial teórico apresentado nas páginas anteriores.

As palavras-chave da pesquisa de Carreira demonstram bastante aderência aos conceitos que utilizamos em nossa trilha teórica e metodológica. Já o trabalho do pesquisador Filipe Silva de Oliveira não apresenta pontos de intersecção relacionados ao tema de nossa pesquisa. Para refinar a busca, com o objetivo de encontrar algum texto que corresponda exatamente ao termo que estamos investigando, utilizamos as aspas para indicar ao mecanismo de busca da BDTD

que buscamos *ipsis litteris* pelo termo "teoria matemática do jornalismo". A segunda pesquisa não retornou resultados, como é possível visualizar a seguir:

FIGURA 25 - RESULTADOS COM O USO DE ASPAS

The screenshot shows the BDTD (Biblioteca Digital Brasileira de Teses e Dissertações) search interface. At the top left is the BDTD logo. Navigation links include 'Página Inicial', 'Sobre a BDTD', 'Rede BDTD', 'Acesso Aberto Brasil', and 'Serviços'. A search bar contains the query '"teoria matemática do jornalismo"'. Below the search bar, it displays 'Search: "teoria matemática do jornalismo"'. The main content area shows 'No Results!' and 'Your search - "teoria matemática do jornalismo" - did not match any resources.' A blue box provides a tip: 'You may be able to get more results by adjusting your search query.' with a bullet point: 'Removing quotes may allow a broader search: [teoria matemática do jornalismo](#).'



Instituto Brasileiro de Informação em Ciência e Tecnologia
SAUS Quadra 05 Lote 06 Bloco H – Asa Sul, Brasília, DF, 70070-912



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÕES



Fonte: Captura de Tela do site da BDTD (<https://btdt.ibict.br/vufind/>)

Após a pesquisa na Biblioteca Digital Brasileira de Teses e Dissertações, consultamos o Catálogo de Teses e Dissertações da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). A CAPES é uma Fundação do Ministério da Educação (MEC), e, segundo dados do site da organização, tem como missão a expansão e consolidação da pós-graduação *stricto sensu* (mestrado e doutorado) no Brasil. A CAPES é uma das principais fundações de avaliação, qualificação, financiamento e promoção da ciência no Brasil e atua em diversas frentes, que são agrupadas por programas/atividades com o objetivo de:

- Avaliar os programas de pós-graduação *stricto sensu* no Brasil;
- Facilitar o acesso e divulgação da produção científica;
- Realizar investimentos na formação de pessoal de alto nível, no País e exterior;
- Promover a cooperação científica internacional;
- Induzir e fomentar a formação inicial e continuada de professores para a educação básica nos formatos presencial e à distância.

O repositório de teses e dissertações da CAPES é uma das atividades vinculadas à Fundação e tem o propósito de reunir as teses e dissertações defendidas nos Programas de Pós-

Graduação brasileiros anualmente. Em abril de 2023, o repositório contava com 1.157.852 arquivos em sua base, sendo 495.163 dissertações de Mestrado e 662.689 teses de Doutorado⁶⁶.

O catálogo fica disponível no banco de dados da CAPES, vinculado ao site <<https://catalogodeteses.capes.gov.br/catalogo-teses>>, onde o usuário pode inserir as palavras-chave no buscador e realizar o primeiro passo para sua pesquisa. Após, o catálogo oferece algumas opções de segmentação e filtros, que auxiliam o usuário a refinar sua busca de acordo com: tipo (tese ou dissertação), ano, autor, orientador, banca, grande área do conhecimento, área de conhecimento, área de avaliação, área de concentração, nome do programa, instituição e biblioteca. Esses filtros segmentam a busca para obter resultados mais direcionados e adequados àquilo que o usuário necessita, além disso, as combinações de filtros podem trazer resultados diferentes a cada nova segmentação. A primeira pesquisa realizada no catálogo de teses e dissertações, assim como na BDTD, foi apenas pelo termo de pesquisa "matemática e jornalismo", sem as aspas e sem operadores booleanos.

FIGURA 26 - BUSCA NO CATÁLOGO DA CAPES

Busca

matemática e jornalismo

1130931 resultados para **matemática e jornalismo**

Exibindo 1-20 de 1130931

Refinar meus resultados

Tipo: 6 opções

Mestrado (Dissertação) 781199

Doutorado (Tese) 266267

1. OLIVEIRA, THALITA RAPHAELA NEVES DE. **JORNALISMO ESPORTIVO E A COBERTURA DA RIVALIDADE GRENAL EM 2016: O TÍTULO DO GRÊMIO E O REBAIXAMENTO DO INTER** 08/08/2018 431 f. Mestrado em JORNALISMO Instituição de Ensino: UNIVERSIDADE FEDERAL DE SANTA CATARINA, Florianópolis Biblioteca Depositária: BIBLIOTECA DA UNIVERSIDADE FEDERAL DE SANTA CATARINA (BU/UFSC) [Detalhes](#)
2. VARASQUIM, MARCELA CRISTINA ZIMOLO. **ESCREVER COM IMAGENS: A NARRATIVA NA PRODUÇÃO CINEMATOGRAFICA DA JORNALISTA ELIANE BRUM** 15/12/2017 173 f. Mestrado em JORNALISMO Instituição de Ensino: UNIVERSIDADE FEDERAL DE SANTA CATARINA, Florianópolis Biblioteca Depositária: Biblioteca Universitária da UFSC - BU

Fonte: Captura de Tela do site da CAPES

⁶⁶ Dados recolhidos do Painel de Informações Quantitativas do Catálogo de Teses e Dissertações da CAPES. Disponível em: <http://visaoanalitica.capes.gov.br/SASVisualAnalyticsViewer/VisualAnalyticsViewer_guest.jsp?reportName=Banco+de+Teses+e+Disserta%C3%A7%C3%B5es+-+Informa%C3%A7%C3%B5es+quantitativas&reportPath=/DTI/Banco_de_teses_e_dissertacoes/Relatorios&reportViewOnly=false&appSwitcherDisabled=true#>>

A primeira pesquisa identificou 1.130.931 resultados para "matemática e jornalismo". Para refinar a busca, utilizamos o filtro de ano, delimitando a análise pelos anos que compreendem 2000 a 2018⁶⁷. A busca reduziu o corpus para 996.090 trabalhos, entre teses e dissertações. O alto número de resultados, no entanto, conflita com a quantidade de trabalhos apresentados no Painel de Informações Quantitativas do Catálogo de Teses e Dissertações da CAPES. Utilizando as aspas como modificador de termo de pesquisa (exemplo: "palavra-chave") com o objetivo de obter uma correspondência exata do termo pesquisado, obtivemos o seguinte resultado:

FIGURA 27 - BUSCA NO CATÁLOGO DA CAPES – ESPECÍFICA



Fonte: Captura de Tela do site da CAPES

Novamente, assim como na BDTD, no repositório de teses e dissertações da CAPES, não encontramos nenhum resultado para a busca exata do termo "jornalismo e matemática". No decorrer deste capítulo, vamos discutir sobre a inexistência de resultados nessas bases de dados e refletir sobre as razões pelas quais o termo não aparece nesses locais quando a pesquisa é feita pela correspondência exata da palavra-chave. Como forma de ampliar a reflexão sobre o assunto, e já vislumbrando possíveis contribuições para o que podemos chamar de uma "teoria matemática do jornalismo", fizemos a pesquisa pelo termo, também utilizando o modificador de correspondência (aspas), e a plataforma também não retornou resultados. Embora não seja o objetivo desta pesquisa, acreditamos que a consulta sobre o termo "teoria matemática do jornalismo" é relevante, principalmente por indicar um caminho que ainda não foi explorado nas teses e dissertações brasileiras.

⁶⁷ Como forma de delimitar a análise exploratória, elencamos um período de 20 anos de análise, que compreende desde os anos 2000 até 2020. Entretanto, o repositório da CAPES contém informações apenas até o ano de 2018. Desta forma, no site da CAPES utilizamos os filtros disponíveis e a análise considera os anos entre 2000 e 2018 por essa razão.

A terceira base de dados consultada nesta análise sistemática foi o Google Acadêmico (*Google Scholar*, em seu nome original na língua inglesa). Trata-se de um repositório online e gratuito de materiais científicos e configura-se como um mecanismo de busca (operando de maneira similar ao próprio mecanismo de busca do Google), mas com o foco especificamente na literatura científica. Nele, é possível consultar livros completos, capítulos de livros, teses, dissertações, artigos e relatórios científicos em diversas línguas. Além dos resultados brutos da pesquisa por palavras-chave, o usuário também tem a possibilidade de utilizar alguns filtros com o intuito de refinar sua busca, como o filtro de período específico, por relevância, por combinação de palavras-chave, por autoria específica e artigos publicados em determinados locais (como revistas científicas, por exemplo):

FIGURA 28 - PESQUISA AVANÇADA GOOGLE ACADÊMICO

Pesquisa avançada

Localizar artigos
 com **todas** as palavras
 com a **expressão exacta**
 com **pelo menos uma** das palavras
sem as palavras

ocorrência das minhas palavras em qualquer parte do artigo
 no título do artigo

Devolver artigos **com autoria** de
 exemplo, "PJ Hayes" ou McCarthy

Devolver artigos **publicados** em
 por exemplo, J Biol Chem ou Nature

Devolver artigos **com datas**
 entre
 por exemplo, 1996

Fonte: Captura de Tela do site do Google Scholar (Acadêmico)

Com o objetivo de ampliar nossa busca pelos trabalhos relacionados aos operadores teóricos que invocamos neste estudo, optamos por consultar o Google Acadêmico justamente pelo fato de ser um dos maiores repositórios de literatura científica gratuita do mundo. Nossa primeira pesquisa foi pelo termo em português, sem o uso de modificador de correspondência de palavra-chave:

FIGURA 29 - BUSCA SIMPLES NO GOOGLE ACADÊMICO

The screenshot shows the Google Scholar interface. At the top, the search bar contains 'matemática e jornalismo' and indicates approximately 49,500 results. On the left, there are filters for 'Artigos', 'Período específico...' (with options for 'Desde 2023', 'Desde 2022', 'Desde 2019'), 'Ordenar por relevância', 'Ordenar por data', 'Em qualquer idioma', 'Pesquisar páginas em Português', 'Qualquer tipo', 'Artigos de revisão', and 'Criar alerta'. The main area displays search results, including a tip to search in Portuguese and several articles with titles like 'O papel da imprensa no Movimento da Matemática Moderna', 'Jornalismo de (im) precisão: o conhecimento matemático e a apuração de números', 'A literacia matemática nas notícias: um problema para o jornalismo?', 'Literacia matemática e prática jornalística: o que o jornal revela sobre o que o jornalista sabe', and 'A matemática na imprensa diária portuguesa'. Each result includes a brief abstract and a link to the full text (PDF).

Fonte: Captura de Tela do site Google Acadêmico

No Google Acadêmico, foram quase 50 mil resultados na busca pelos operadores teóricos "matemática e jornalismo". Após a primeira consulta, utilizamos o filtro de datas para segmentar entre o ano de 2000 e 2020. Nessa segunda pesquisa, o Google mostrou "aproximadamente 15.700 resultados"⁶⁸. Aqui, vale ressaltar que o Google Acadêmico funciona como um grande repositório de conteúdo, e que sua busca retorna resultados por meio da indexação destes conteúdos, ou seja, em determinados resultados, talvez os termos "jornalismo e matemática" sequer apareçam relacionados entre si, mas são palavras-chave que foram, de alguma forma, indexadas pelo Google.

Por exemplo, no resultado a seguir, o Google Acadêmico considerou o livro "Poesia Matemática", de Millôr Fernandes, como um resultado que corresponde à nossa pesquisa. No entanto, se consultarmos de fato o livro, podemos perceber que se trata de uma obra de poesia e que o termo "jornalismo" (correspondência exata) sequer aparece. O que aparece nas páginas do livro é uma variação de "jornalismo": pudemos perceber que a palavra "jornalista" aparece (apenas) duas vezes no livro. A palavra está inscrita na sessão "Sobre o autor", onde há informações sobre a biografia de Millôr Fernandes. Além disso, a "matemática" no livro é

⁶⁸ O Google Acadêmico não reproduz o número exato de resultados, utilizando o termo "aproximadamente" para que o usuário tenha uma noção quantitativa dos resultados de sua busca.

utilizada de forma metafórica, já que o autor usa termos da matemática de forma poética, conforme podemos ver a seguir: “E assim se amaram ao quadrado da velocidade da luz, numa sexta potenciação, traçando, ao sabor do momento e da paixão, retas, curvas, círculos e linhas sinoidais, nos jardins da quarta dimensão.” (FERNANDES, 2014, p. 12).

FIGURA 30 – RESULTADO DE PESQUISA DO GOOGLE ACADÊMICO

[LIVRO] Poesia matemática

M Fernandes - 2014 - books.google.com

... Méier, construiu como **jornalista**, desenhista, humorista, tradutor e dramaturgo uma trajetória
... Órfão de pai com um ano de idade e de mãe aos dez, começou a carreira de **jornalista** aos ...

☆ Salvar  Citar Citado por 7 Artigos relacionados Todas as 8 versões

Fonte: Captura de Tela do site Google Acadêmico

Pudemos perceber, ainda, que o Google Acadêmico mostra resultados que contenham: a) as duas palavras exatas; b) uma palavra exata; c) as duas palavras e/ou suas variações; e d) apenas uma variação da pesquisa. Dessa forma, o Google Acadêmico mostra centenas, até milhares de resultados, pois as variações entre as palavras-chave fazem com que a busca seja mais abrangente. O Google considera esse tipo de pesquisa como "correspondência ampla", ou seja, ele irá mostrar resultados que contenham as palavras-chave pesquisadas ou variações dessas palavras-chave. Por essa razão, acabou mostrando um livro de poesia, um relatório do progresso da COVID-19 no Brasil, uma pesquisa quantitativa do uso do Moodle no ensino do jornalismo e outros tantos conteúdos que não relacionam jornalismo e matemática da forma como estamos fazendo neste estudo. Tendo isso em vista, utilizamos o que o Google chama de "modificador de correspondência de palavra-chave"⁶⁹ para refinar nossa busca e pesquisar pelo termo exato "jornalismo e matemática, utilizando as aspas para delimitar nossa pesquisa, de acordo com o que o Google chama de "correspondência de frase". Nessa pesquisa, o Google Acadêmico retornou seis resultados:

⁶⁹ Suporte do Google, em português, com explicações sobre correspondência de palavra-chave: <https://support.google.com/google-ads/answer/7478529?hl=pt-BR>

QUADRO 8 – Resultados utilizando modificador de correspondência de frase no Google Acadêmico

SANTOS, Antonio Eduardo. É SAL É SOL É SUL DE NOUVEAU... POIS!!!. 2020.
COLFERAI, Sandro Adalberto et al. Um jeito amazônida de ser mundo. A Amazônia como metáfora do ecossistema comunicacional: uma leitura do conceito a partir da região. 2014.
MACHADO, Leticia dos Santos. Dimensão social e educativa da biblioteca e do bibliotecário: estudo de comunidade em Maricá. 2019.
SOARES, Sergei Suarez Dillon. O conhecimento paga bem? Habilidades cognitivas e rendimentos do trabalho no Brasil (e no Chile). 2011.
MARQUES, Daniele Luciano. Entre a escola unitária e a mercadológica: a trajetória para o mundo do trabalho dos egressos das EEEPs do Ceará. 2016.
BRASIL, RENDIMENTOS DO TRABALHO NO. SERGEI SUAREZ DILLON SOARES.

Fonte: Google Acadêmico

Ao analisarmos os seis resultados, observamos que a busca de fato foi insatisfatória. Mesmo utilizando uma estratégia de busca que o próprio Google recomenda para a otimização dos resultados, não conseguimos identificar algum que tenha relação com a nossa pesquisa. Por último, novamente, fizemos uma pesquisa do termo "teoria matemática do jornalismo", sem o modificador de correspondência de palavra-chave:

FIGURA 31 - RESULTADO DE PESQUISA DO GOOGLE ACADÊMICO PARA TEORIA MATEMÁTICA DO JORNALISMO

The screenshot shows the Google Acadêmico search interface. The search bar contains the text "teoria matemática do jornalismo". Below the search bar, there are filters for "Artigos" and "Aproximadamente 16.100 resultados (0,02 s)". On the left side, there are various filters including "A qualquer momento", "Desde 2023", "Desde 2022", "Desde 2019", "Período específico...", "Pesquisar", "Ordenar por relevância", "Ordenar por data", "Em qualquer idioma", "Pesquisar páginas em Português", "Qualquer tipo", "Artigos de revisão", "Incluir patentes", "Incluir citações", and "Criar alerta". The main search results are displayed in a list format, each with a title, author, year, and a link to the full text. The results include:

- Jornalismo e meio ambiente: a contribuição dos meios de comunicação e o conceito de sustentabilidade** (PDF) usp.br. PC Campos - RuMoRes, 2012 - revistas.usp.br. Teoria Cibernética3 que desenvolveu, nos anos 1940, o princípio da circularidade da informação como processo comunicativo, incorporando as bases da Teoria ... que o matemático e ...
- Questões sobre jornalismo e ideologia** (PDF) ufsc.br. A Genro Filho - Estudos em Jornalismo e Mídia, 2004 - periodicos.ufsc.br. ... A aplicação mecânica da Teoria da Informação ao jornalismo é uma delas. Há uma frase ... tidade de informação segundo a visão matemática da Teoria da Informação. Entretanto, é fácil ...
- Teorias e técnicas do jornalismo e da comunicação** (livro) BB Barreira - 2013 - books.google.com. ... o modelo de informação da Teoria Matemática da Comunicação foi o paradigma dominante... Existem três explicações para a continuação da Teoria Matemática da Comunicação como ...
- A literacia matemática nas notícias: um problema para o jornalismo?** (PDF) ua.pt. S Pereira, J Azevedo, A Machiavello - SOPCOM, Associação Portuguesa ..., 2013 - proa.ua.pt. ... matemática nas notícias a dois fatores: a falta de confiança e a falta de conhecimentos lógico-matemáticos dos jornalistas... o nível de literacia matemática dos jornalistas, desde testes ...

Fonte: Captura de tela do site Google Acadêmico

No Google Acadêmico, quando pesquisado o termo "teoria matemática do jornalismo", sem o uso de um modificador de palavra-chave (neste caso, as aspas), o mecanismo de busca trouxe mais de dezesseis mil resultados que correspondem ao termo entre os anos 2000 e 2022. A mesma pesquisa, porém, pela palavra-chave em inglês, trouxe mais de vinte mil resultados. Muitos deles relacionados à Teoria Matemática da *Comunicação e Informação*.

FIGURA 32 - BUSCA PELO TERMO "MATHEMATICAL THEORY OF JOURNALISM"

The screenshot shows the Google Scholar interface. At the top, the search bar contains "Mathematical theory of journalism" and shows approximately 20,600 results in 0.03 seconds. On the left, there are filters for "Artigos" and a date range selector set to "2000" to "2022". Below the filters, two search results are displayed:

- [LIVRO] Lecturing birds on flying: Can mathematical theories destroy the financial markets?** by P Triana - 2009 - books.google.com. The snippet mentions "mathematical theories" and "comprehension of which is way beyond the reach of mere mortals".
- [LIVRO] The mathematical theory of information** by J Kähre - 2002 - books.google.com. The snippet states "The general concept of information is here, for the first time, defined mathematically by adding one single axiom to the probability theory. This **Mathematical Theory** of Information is ...".

Fonte: Captura de tela do site Google Acadêmico

O termo pesquisado de maneira direta, sem o operador⁷⁰ que indica a correspondência exata da busca, trouxe centenas de resultados genéricos, já que o mecanismo de busca expõe resultados que podem, ou não, ter alguma convergência com a área de interesse da busca. Neste caso, para aprimorar a entrega dos resultados, usamos o modificador de termo de pesquisa que indica a correspondência exata do termo, colocando-o entre aspas. Assim, o uso das aspas indica para a ferramenta de busca que não estamos em busca de trabalhos científicos *relacionados* com o tema, mas sim, por trabalhos *sobre* o tema. Essa distinção é importante porque o Google Acadêmico pode mostrar diversos resultados que não correspondem diretamente ao termo que foi pesquisado ou, ainda, resultados que não estão dentro da área de conhecimento pretendida. O Google Acadêmico, como um motor de busca, funciona por meio de comandos e, para

⁷⁰ O uso de aspas indica para o mecanismo de busca que o usuário está pesquisando pelo termo exato. Existem outros operadores que também têm a função de refinar as buscas nos motores de busca como o Google Acadêmico. É o caso do uso de operadores Booleanos, que permitem refinar a busca nos repositórios acadêmicos e bancos de dados científicos com a utilização de termos como *and*, *or* e *not* no momento da pesquisa.

aperfeiçoar nossa busca, também é essencial que saibamos como encontrar os melhores resultados utilizando artifícios permitidos dentro da ferramenta.

A partir disso, refizemos a busca, porém, fazendo o uso do modificador de palavra-chave. Neste caso, quando fazemos a pesquisa direta, sem o uso de nenhum modificador, o Google irá mostrar a maior quantidade de resultados possíveis de acordo com aquele termo, por exemplo: quando a busca é sobre "teoria matemática do jornalismo", como foi o nosso caso, o Google poderá combinar todos os termos, de forma ordenada ou não e apresentar resultados sobre:

- a) Teoria Matemática;
- b) Teoria do Jornalismo;
- c) Matemática e Jornalismo;
- d) Variações que possam ser decorrentes de cada uma das opções acima, como: teoria da informação e teoria da comunicação, que são operadores elementares e que, geralmente, aparecem no conjunto de palavras-chave indexadas juntamente às demais.

Nessa segunda pesquisa, com o objetivo de sermos mais específicas e encontrar resultados com maior aderência ao nosso campo e objeto de estudos, dos mais de dezesseis mil resultados mostrados pelo Google Acadêmico anteriormente, reduzimos a amostra para zero. Buscando exatamente pelo termo "teoria matemática do jornalismo", a plataforma não foi capaz de encontrar artigos ou livros onde essas palavras-chave estejam indexadas. Como sugestão, o Google elenca uma lista de indicações para melhorar a pesquisa realizada:

- Tente pesquisar num período mais abrangente.
- Certifique-se de que todas as palavras estejam escritas corretamente.
- Tente palavras-chave diferentes.
- Tente palavras-chave mais genéricas.
- Tente usar menos palavras-chave.

O que isso indica? Refletindo a respeito dessas pesquisas, podemos considerar que o termo não aparece na correspondência exata nas bases de dados porque:

1. É um termo original e ainda não indexado, portanto, pode ser considerado como um assunto interessante para ser proposto em novas pesquisas;

Ou:

2. Não é um assunto relevante, já que não há registros de outras pesquisas utilizando o mesmo termo.

De qualquer forma, preferimos acreditar que a falta de outros registros é uma oportunidade para criar um novo ambiente de discussões, reflexões e debates dentro do campo

científico, unindo disciplinas distantes com o objetivo de avançar na construção de conhecimento compartilhado e interdisciplinar.

Ao observar as três bases de dados que consultamos, pudemos ainda considerar que:

1. De fato, jornalismo e matemática não são disciplinas pesquisadas de forma conjunta ou colocadas em relação mútua em trabalhos científicos;
2. Há uma abertura teórico-epistemológica para criar discussões e aproximações a respeito das duas áreas de conhecimento;
3. Quando pesquisamos "teoria matemática do jornalismo", frequentemente a teoria de Shannon apareceu listada como termo relacionado, o que pode indicar a grande influência da obra do autor para os estudos acadêmicos;
4. As limitações das ferramentas impõem à ciência alguns desafios, como o de pensar em, além de contribuir para os avanços da ciência, criar obras que sejam legíveis por mecanismos de buscas a fim de indexar os conteúdos;
5. Sendo assim, devemos dar grande importância para os termos de nossas pesquisas, buscando a indexação de nossos trabalhos nas bases de dados e possibilitando maior divulgação da ciência nesses repositórios;

De qualquer forma, embora a pesquisa exploratória nesses repositórios seja importante para observarmos, de forma generalista, o estado da arte em relação aos operadores teóricos de nosso estudo, é necessário dar um passo à frente para que possamos observar mais a fundo o comportamento dessas palavras-chave em outras bases de dados relevantes para a grande área da comunicação, na qual estamos inseridos.

8.2 OS PERIÓDICOS CIENTÍFICOS

Objetivando incrementar ainda mais a busca pelos termos "jornalismo e matemática e, ainda, "teoria matemática do jornalismo", consultamos os sites dos periódicos científicos mais bem avaliados segundo os critérios de qualidade da Capes (o sistema Qualis CAPES avalia os periódicos e a produção intelectual dos programas de pós-graduação brasileiros, concedendo a cada periódico uma classificação entre A1, A2, B1, B2, B3, B4, B5 e C, em que A1 indica o maior e mais relevante fator de impacto científico e C, o menor). Cursos ligados à área da comunicação, como Jornalismo e Relações Públicas, estão agrupados dentro da grande área de avaliação "Comunicação e Informação", segundo a CAPES. Para verificarmos os periódicos mais bem avaliados na área de nosso interesse, é necessário encontrar em qual enquadramento institucional estamos sendo considerados. Nesse caso, na área de avaliação "Comunicação e

Informação", os periódicos mais bem avaliados, ligados a programas de pós-graduação da área da comunicação, são os seguintes: revista E-Compós, Galáxia, revista Intercom, Famecos e Matrizes.

8.2.1 Revista Científica E-Compós

A revista E-Compós é uma publicação científica vinculada à Associação Nacional dos Programas de Pós-Graduação em Comunicação – Compós. Segundo informações do site⁷¹ do periódico, seu formato é apenas eletrônico, com periodicidade regular em modalidade de publicação contínua. Tem como principal missão difundir a produção acadêmica original e inédita em periódicos científicos de pesquisadores da área de comunicação, inseridos em instituições do Brasil e do exterior. São aceitas submissões em português, espanhol ou inglês. A revista E-Compós opera desde 2004 e já publicou 25 volumes, tendo conquistado a importante posição de uma das mais relevantes revistas científicas da área da comunicação no Brasil, com avaliação A2, segundo o sistema Qualis. Utilizamos o sistema de busca dentro do site da E-Compós, entretanto, não obtivemos nenhum resultado pelo buscador interno da aplicação. Realizamos outras buscas pelo mesmo mecanismo e identificamos que este recurso, embora disponível no *header* do site do periódico, é limitado e não consegue identificar pesquisas que incluam mais de uma palavra-chave.

8.2.2 Revista Científica Galáxia

O segundo periódico consultado foi a revista Galáxia, também considerada como A2 na avaliação do sistema Qualis. A Galáxia⁷² é a revista do programa de pós-graduação em Comunicação e Semiótica, da PUC-SP (COS) e, desde sua origem, em 2001, já publicou 48 edições. Utilizando seu sistema interno de busca, não obtivemos resultados para o termo "jornalismo e matemática", embora o termo "jornalismo", quando pesquisado de maneira separada, tenha 173 resultados e "matemática" aparece em 4 artigos. São eles:

⁷¹ Site Revista E-Compós: <<https://www.e-compos.org.br/e-compos>>

⁷² Site Revista Galáxia: <<https://revistas.pucsp.br/galaxia>>

QUADRO 9 – Textos com termo "matemática"

Artigo	Autor	Palavras-chave	Edição
MORIN E FLUSSER A teoria da imagem como aventura antropológica e matemática imaginária	Rose de Melo Rocha	Imagem. Imaginário. Comunicação e antropologia.	n. 25 (2013)
Memórias e percursos semióticos do lado de Greimas	Jean Petitot	Catástrofe. contínuo/descontínuo/discreto. Estética. Teoria gestáltica. Iconicidade. Morfodinâmica	2019: Especial 2 Dossiê A.J.Greimas
Sinéreses cromossômicas: projeto de complementaridade semiótica entre som e cor	Eufrásio Prates	-	n. 4 (2002)
Cartografando comentários e sentimentos no perfil de Jair Bolsonaro no Instagram acerca da Covid-19	Douglas Farias Cordeiro Maiara Raquel Campos Leal Larissa Machado Vieira Núbia Rosa da Silva	Cartografia, Covid-19, Análise de sentimentos, Mineração de dados, Instagram	n. 47 (2022): Galáxia

Fonte: Elaborado pela autora (2023)

Embora a pesquisa na revista Galáxia tenha proporcionado alguns resultados, nenhum deles parece ter alguma relevância à discussão que estamos propondo neste estudo. Basta observar as palavras-chave utilizadas pelos autores para concluir que os artigos encontrados nessa busca, embora tenham alguma relação com a semiótica, não discutem sobre os assuntos relacionados à teoria da informação, teoria matemática da comunicação, jornalismo automatizado ou teorias do jornalismo. Já quando a busca foi feita por teoria da informação, obtivemos 17 resultados de artigos que foram acionados por essa palavra-chave, embora, numa visão geral sobre os títulos e palavras-chave, também identificamos que não há muita aderência ao nosso estudo. Quando pesquisamos sobre "teoria do jornalismo", obtivemos 32 resultados com alguma relação com o termo pesquisado.

8.2.3 Revista Intercom

A Revista Intercom - Revista Brasileira de Ciências da Comunicação (RBCC) é uma publicação da Sociedade Brasileira de Estudos Interdisciplinares da Comunicação (INTERCOM) e também está avaliada como A2 no sistema Qualis. Segundo informações do site⁷³ do periódico, o seu foco de interesse é a Comunicação Social, respeitando a

⁷³ Site Revista Intercom: <<https://revistas.intercom.org.br/index.php/revistaintercom>>

interdisciplinaridade e a abrangência temática características da área do conhecimento. Quando utilizamos o sistema de busca interno do site, o mecanismo não mostrou resultados para "jornalismo e matemática", nem para "teoria matemática do jornalismo", nem para "teoria matemática da comunicação", nem para "teoria da informação", mostrando um total de 33 artigos apenas para a busca por "teoria do jornalismo". Quando fizemos a busca por "jornalismo", apenas, obtivemos 254 artigos como resultado e quando a pesquisa foi feita por "matemática", novamente, não obtivemos nenhum resultado.

8. 2.4 Revista Matrizes

A Revista Matrizes⁷⁴ também é um periódico com a avaliação A2 pela CAPES. É o periódico científico do programa de pós-graduação em Ciências da Comunicação, da Universidade de São Paulo. Lançada no segundo semestre de 2007, é destinada à publicação de estudos que tenham por objeto a comunicação em seus múltiplos aspectos e dimensões. Na busca por "jornalismo e matemática", o buscador mostrou 4 resultados:

QUADRO 10 – Resultados da busca na Revista Matrizes

Artigo	Autor	Palavras-chave	Edição
Acontecimento como singularidade	José Luiz Aidar Prado	Acontecimento, Antagonismo, Regimes de interação	v. 16 n. 1 (2022): MATRIZES
Os Estudos Culturais e a urgência por interdisciplinaridade: cedo, e não tarde, vamos precisar de uma Ciência da Cultura	John Hartley	Estudos Culturais, Comunicação, Cultura, Ciência da Cultura, Interdisciplinaridade	v. 5 n. 1 (2011)
O que é Glocal? Sistematização conceitual e considerações teóricas sobre a mais importante invenção tecnocultural da civilização mediática	Eugênio Rondini Trivinho	Civilização mediática, Fenômeno glocal, Condição glocal	V. 16 N. 2 (2022): MATRIZES
IEML: rumo a uma mudança de paradigma na Inteligência Artificial	Pierre Lévy	Inteligência artificial, Código semântico, Inteligência coletiva, Metalinguagem da Economia da Informação	v. 16 n. 1 (2022): MATRIZES

Fonte: Elaborado pela autora (2023)

Dentre os resultados acima, destacamos um artigo escrito por Pierre Lévy, filósofo cuja literatura intelectual discorre, vastamente, sobre os efeitos da cibercultura na sociedade

⁷⁴ Site Revista Matrizes: <<https://www.revistas.usp.br/matrizes>>

contemporânea. Quando pesquisamos apenas a palavra-chave "jornalismo", obtivemos 98 artigos como resultado, e apenas "matemática", obtivemos 6 resultados, dois a mais do que apresentamos no quadro anterior.

Na busca pelo termo "teoria matemática do jornalismo" obtivemos o mesmo resultado. Para outras pesquisas, como "teorias do jornalismo", o repositório da Matrizes demonstrou 19 artigos como resultado e outros 15 artigos relacionados à palavra-chave "teoria da informação".

8.2.5. Revista Famecos

O último periódico consultado foi a revista Famecos⁷⁵, que é vinculada ao programa de pós-graduação em Comunicação Social, da Pontifícia Universidade Católica do Rio Grande do Sul. Sendo o mais antigo dos periódicos consultados, a revista Famecos foi criada em setembro de 1994 e já publicou 29 volumes (cada um desses, no entanto, costuma contar com no mínimo 3 edições durante o ano). Quando pesquisamos pelo termo "jornalismo e matemática", o buscador não nos mostrou nenhum resultado, enquanto o termo sozinho "jornalismo" aparece em 244 artigos dentro do repositório da revista e o termo "matemática" consta em 2 artigos.

QUADRO 11 – Resultados da busca por "matemática" na Revista Famecos

Artigo	Autor	Palavras-chave	Edição
Imagem e Cognição em Cena: um estudo com a relação Iconicidade/Indexicalidade	Maria Ogécia Drigo	Imagem, representação visual, cognição	v. 17 n. 1 (2010)
Os estudos sobre a hipótese de agendamento	Antonio Hohlfeldt	Comunicação, teorias da comunicação, agendamento	v. 4 n. 7 (1997)

Fonte: Elaborado pela autora (2023)

Desses dois artigos, destacamos o texto escrito pelo pesquisador Antonio Hohlfeldt, sobre a hipótese do agendamento. No artigo, o pesquisador cita a teoria matemática da comunicação, proposta por Shannon, no momento em que menciona outras teorias da comunicação e do jornalismo, embora não trate a teoria de Shannon como um operador teórico de destaque no artigo sobre o agendamento midiático.

⁷⁵ Site Revista Famecos: <<https://revistaseletronicas.pucrs.br/ojs/index.php/revistafamecos/index>>

Na pesquisa por "teoria matemática do jornalismo", não obtivemos resultados. Após, pesquisamos por "teoria do jornalismo" e o sistema mostrou 32 itens em seu banco de dados. Por fim, pesquisamos por "teoria matemática da comunicação" e o sistema retornou, novamente, o artigo do pesquisador Antonio Hohlfeldt, cujas palavras-chave são: comunicação, teorias da comunicação e agendamento.

Passando o olhar sob todos esses repositórios, pudemos observar que a matemática e o jornalismo não aparecem atrelados a praticamente nenhuma obra contemporânea entre os anos de 2000 e 2020. Constatamos, no entanto, que esta pode ser uma oportunidade de inaugurar uma corrente de pensamento que priorize compreender as relações entre essas duas áreas, demonstrando que há, de fato, possibilidades de reflexão, mesmo que nível teórico, sobre essa aproximação.

Agora, após a pesquisa exploratória, gostaríamos de resgatar todo o percurso teórico que nos trouxe até aqui: desde os primeiros conceitos de técnica, a história e evolução da técnica, o desenvolvimento do conceito de tecnologia, passando pelas conceituações de Vieira Pinto (2005) sobre o efeito de maravilhamento da técnica, passando pelos primeiros nomes de destaque, como Ada Lovelace e Charles Babbage com o Analisador Diferencial, Shannon com os sistemas de comunicação e o início da era da informação, passando pelo início da cibernética, até chegar à ascensão dos sistemas de automação.

Nesse momento, também atentamos para o fato de que é necessário costurar todos esses períodos da história da tecnologia com o desenvolvimento do jornalismo, que, desde sua origem, sempre foi permeado pelas tecnologias características de cada época. Sob o ponto de vista desta pesquisa, tecnologia e jornalismo andaram juntos, acompanhando o desenvolvimento da sociedade. O fato é que, neste momento, o jornalismo e a tecnologia inauguraram uma nova era: a era do jornalismo automatizado. Um (novo) jornalismo operado por meio da lógica de sistemas de automação, com *inputs*, *outputs*, dados e circuitos que, se olharmos bem, podem ser reduzidos a operações matemáticas. É por isso que, nas próximas páginas, buscaremos aproximar o jornalismo de automação das noções matemáticas (RUSSEL, 1993).

9 CRIANDO ESTRUTURAS ALGORÍTMICAS: AS BASES PARA UM MODELO MATEMÁTICO DO JORNALISMO POR MEIO DE ITERAÇÕES LÓGICAS

Após a etapa de aproximação teórico-metodológica entre os termos jornalismo e matemática, acima apresentada, partimos para um mergulho mais profundo no entendimento dos sistemas de automação utilizados para criar notícias, para minuciar esses sistemas a fim de compreender seu funcionamento. A partir dessa compreensão, então, buscaremos sistematizar as etapas do sistema dentro de uma lógica matemática, utilizando alguns dos preceitos indicados por Russell (1993) para criar um modelo matemático para o jornalismo, inspirado no funcionamento dos sistemas de automação vigentes.

Aqui, optamos por fazer o caminho inverso: em vez de olhar para a matemática primeiro, olhamos para o modelo jornalístico vigente (o quinto jornalismo, como considera Carreira, 2017), no qual os sistemas de automação operam. Olhando para estes sistemas, buscamos compreender seu funcionamento e destrinchá-lo até chegarmos à sua essência, em como esses softwares operam, quais etapas são necessárias para criar uma notícia. Assim, sistematizando essas operações, será possível criar iterações lógicas, capazes de reproduzir o *modus operandi* jornalístico por meio de teoremas matemáticos. A seguir, iniciamos nosso percurso focando nossas lentes nos sistemas de automação.

9.1 ANATOMIA DO SISTEMA: COMPREENDENDO O FUNCIONAMENTO DE UM SISTEMA DE AUTOMAÇÃO JORNALÍSTICA NA PRÁTICA

Para apresentarmos um teorema que dê conta de ilustrar a proposição de um modelo matemático do jornalismo, tendo como inspiração o funcionamento dos sistemas de automação de notícias, buscamos focar nossa lente de pesquisa na compreensão do funcionamento desses sistemas. Para tanto, buscamos diversas fontes para abastecermos-nos de conhecimento, tendo em vista o fato de que, para compreendermos como um sistema funciona, precisamos familiarizarmos-nos com a área de desenvolvimento de sistemas e programação de softwares. Por isso, passamos para a etapa de desenvolvimento de uma aplicação com foco em automação de notícias em um ambiente de desenvolvimento integrado (IDE, *integrated development environment*), onde apresentamos a estrutura lógica capaz de gerar notícias automáticas por meio de comandos (funções, dentro do sistema).

Assim, buscamos estabelecer aproximações com a ciência da computação para compreender como é possível criar algoritmos capazes de construir notícias automatizadas. A curiosidade em relação a este tema resultou em algumas aproximações teórico-metodológicas em três instâncias distintas, porém complementares:

1. Em primeira instância, a autora realizou um curso introdutório sobre linguagens de programação com foco em *JavaScript*. O curso foi feito de maneira online e remota e contou com módulos teóricos e práticos;
2. Em segunda instância, a autora buscou relacionar-se com profissionais da área de desenvolvimento de sistemas, criando momentos de discussão sobre os objetivos desta tese e coletando referencial prático para dar andamento às proposições;
3. A autora consultou professores da área de desenvolvimento de sistemas, apresentando a pesquisa e buscando sugestões de desenvolvimento metodológico.

Após essas aproximações, que foram de caráter exploratório para este capítulo metodológico, coletamos referências para dar andamento à proposição de nosso modelo matemático inspirado no funcionamento dos sistemas de automação. Para tanto, julgamos necessário compreender como um sistema de automação funciona na prática. Vale ressaltar que nos propusemos a criar essa estrutura algorítmica porque, durante o percurso desta pesquisa, entramos em contato com algumas das empresas citadas (*Automated Insights e Arria NLG*) solicitando uma reunião ou demonstração do funcionamento dos sistemas dessas empresas e recebemos retornos negativos. Após, entramos em contato ainda com uma terceira plataforma, que foi descoberta durante o processo de pesquisa, chamada *ArticleForge*, que cria textos por meio da *natural language generation*, cujo foco não é a criação de textos jornalísticos, mas sim, conteúdos para ranqueamento orgânico e SEO (*Search Engine Optimization*). A *ArticleForge* sequer nos deu retorno. Assim, na impossibilidade de abrir a "caixa preta" dessas empresas, buscamos (re)criar um sistema por conta própria, utilizando como base todos os nossos estudos teóricos e aproximações metodológicas apresentadas nesta tese.

Por isso, optamos pela utilização de um IDE (*integrated development environment*), interface de desenvolvimento de aplicações utilizada por programadores e desenvolvedores de sistemas, com o objetivo de facilitar o desenvolvimento de softwares. Neste estudo, fizemos o uso da *WebStorm*, ambiente de desenvolvimento integrado para *JavaScript* e tecnologias relacionadas. Optamos pelo *WebStorm* porque entre suas funcionalidades existem alguns benefícios como a identificação de *bugs* instantânea durante a criação do algoritmo e *autocomplete* de trechos de código (o próprio sistema do IDE completa as frases durante a

escrita do código, facilitando a construção da aplicação, especialmente para desenvolvedores que não têm experiência no desenvolvimento de softwares, como é o nosso caso).

Optamos por realizar a construção do algoritmo em *JavaScript* por constituir-se como uma linguagem de programação recomendada para a criação de softwares e sistemas de automação. Além disso, para realizar a execução do algoritmo, usamos também a tecnologia de *NodeJs*, responsável por coordenar a execução do processo *JavaScript* no servidor. Segundo o site oficial da Mozilla, que fornece uma extensa documentação sobre tecnologias voltadas para o desenvolvimento web, *JavaScript* é uma linguagem de programação que permite implementar itens complexos em páginas web mostrando conteúdo que se atualiza em um intervalo de tempo, mapas interativos ou gráficos 2D/3D animados, etc." (online). Além disso, utilizando *JavaScript*, temos alguns benefícios como: armazenamento de conteúdo útil dentro de variáveis, operações com trecho de texto (conhecidos como "*strings*") e execução de um código variável em resposta a determinados eventos que ocorrem em uma aplicação⁷⁶.

Iniciamos nossa incursão no mundo do *JavaScript* por meio do resgate da tabela que sistematiza os passos fundamentais de NLG em modo *pipeline* aplicada ao jornalismo, desenvolvida durante o percurso teórico desta pesquisa. O esquema (disponível no capítulo 6), que conta com oito etapas, foi reproduzido dentro do ambiente de desenvolvimento integrado, no qual:

- **Dados** correspondem à etapa da antenarrativa, conforme aponta Bertocchi (2016) e são o subsídio do sistema de geração de notícias automatizadas. Podem ser textos, imagens, gráficos etc.;
- **Entrada do sistema** é representada pela função *main*, cuja finalidade é iniciar o processo de execução dentro do sistema;
- **Rotina de coleta:** dentro da função *main*, criamos a função que busca determinar uma rotina de coleta de dados, representada pelo trecho:

⁷⁶ Dados extraídos do site <https://developer.mozilla.org/>, plataforma open-source e colaborativa sobre CSS, HTML, JavaScript, e Web APIs.

FIGURA 33 - CAPTURA DE TELA DO IDE UTILIZADO PARA CRIAR A FUNÇÃO DE ROTINA DE COLETA DE DADOS

```
function main(){  
    // Create a routine to automatic read data and generated a news  
    setInterval(createNewsRoutine, 1000);  
}  
  
main();
```

Fonte: Elaborado pela autora (2023)

Nesse exemplo, criamos uma rotina de atualização da coleta de dados dentro de um intervalo de tempo (*setInterval(createNewsRoutine, timeout: 1000ms)*) e criamos, na sequência, a instrução para que o sistema colete os dados, por exemplo, diretamente da agenda do presidente da república (dado disponibilizado de forma gratuita no portal do governo federal), cujo primeiro passo é ler a agenda e coletar todos os compromissos agendados para dentro do sistema. Se há programação na agenda do presidente, o sistema irá coletar o dado e passá-lo para a próxima etapa:

FIGURA 34 - CAPTURA DE TELA DO IDE COM O MÉTODO DE COLETA DE DADOS

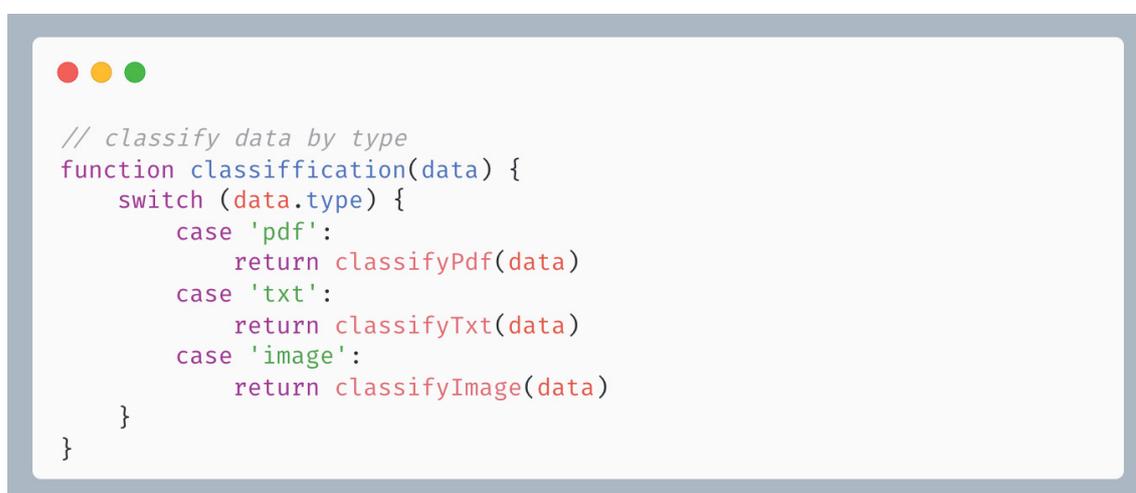
```
// Function to collect data from president schedule  
function collectData() {  
    // read the schedule from the president  
    const data = readPresidentSchedule();  
    // if didn't find any new schedule, return an empty array  
    // and the news generation will be skipped  
    if(data === null) {  
        return [];  
    }  
    // returns the collected data  
    return data;  
}
```

Fonte: Elaborado pela autora (2023)

A próxima etapa, segundo os fundamentos do funcionamento de um sistema de automação de notícias, de acordo com Bertocchi (2016), Devyatkin et.ad. (2019) e Ribeiro (2019) é o da classificação e identificação do formato dos dados coletados. Essa etapa é importante para identificarmos a natureza dos dados e enviá-los para uma leitura e conversão em texto.

- **Classificação do dado:** função que tem o objetivo de extrair o dado de acordo com o seu formato. No modelo algorítmico que criamos, usamos como exemplo os formatos PDF, txt e *image*:

FIGURA 35 - CAPTURA DE TELA DO IDE COM O MÉTODO DE CLASSIFICAÇÃO DE DADOS POR TIPO.



```

// classify data by type
function classification(data) {
  switch (data.type) {
    case 'pdf':
      return classifyPdf(data)
    case 'txt':
      return classifyTxt(data)
    case 'image':
      return classifyImage(data)
  }
}

```

Fonte: Elaborado pela autora (2023)

Após identificar a natureza dos dados, como, por exemplo, uma tabela do excel, no formato .csv com todos os compromissos do presidente da república, passamos para a etapa de validação do dado, na qual a aplicação deverá criar uma etapa de autenticação de todos os dados coletados na etapa 1. Entretanto, para evitar que o sistema deixe passar informações irrelevantes para as demais etapas, criamos, dentro da validação, duas formas de comprovar a veracidade e a relevância daquele dado.

- **Validação do dado:** etapa na qual o sistema deverá validar a informação, com o auxílio de mecanismos de veracidade e relevância, indicados como variáveis, dentro da função, que pode ser descrita da seguinte forma: *function validate (data)* para verificar se o dado é válido, passando pelas linhas de execução *var isValid = verifyVeracity(data)* e *var isRelevant = verifyRelevance (data)*, conforme sintetizado a seguir, no IDE:

FIGURA 36 - CAPTURA DE TELA DO IDE COM FUNÇÃO DE VALIDAÇÃO DO DADO E VARIÁVEIS DE VERACIDADE E RELEVÂNCIA.

A screenshot of an IDE window showing a JavaScript function named 'validate'. The function is designed to check if data is both valid and relevant. It uses two helper functions: 'verifyVeracity' and 'verifyRelevance'. The function returns a boolean value based on the logical AND of these two checks. The code is as follows:

```
// validate the information
function validate(data) {
  // verify if data is valid
  const isValid = verifyVeracity(data);
  // verify if data is relevant
  const isRelevant = verifyRelevance(data);
  // return the validation if the data is valid and relevant
  const isValid && isRelevant;
}
```

Fonte: Elaborado pela autora (2023)

Com os dados coletados, classificados e validados dentro do sistema, passamos para a terceira etapa, cujo principal objetivo é a transcrição dos dados em uma narrativa jornalística (BERTOCCHI, 2016), escrita de acordo com os preceitos levantados por Wolf (1987) e Lage (2021) que caracterizam uma notícia. Assim, nessa etapa do sistema, além de criarmos um comando para que o sistema crie a notícia, ele também precisa levar em consideração algumas regras, como a correspondência ao *lead* jornalístico:

- **Criação da notícia:** o processo de criação da notícia corresponde à etapa em que o dado vai ser transformado em narrativa jornalística, esta é a etapa de produção textual, na qual o sistema deverá responder às perguntas do *lead* jornalístico: *what? who? where? when? why? how?*

Nesta etapa o sistema deverá executar a função *createNews(data)* e retornar as perguntas o que? quem? quando? onde? por que? e como? de acordo com a inspeção realizada anteriormente nos dados imputados dentro do sistema, conforme demonstramos na figura a seguir:

FIGURA 37 - CAPTURA DE TELA DO IDE COM FUNÇÃO CRIAÇÃO TEXTUAL CORRESPONDENDO ÀS REGRAS DO LIDE JORNALÍSTICO

```
// execute the extraction the information from the data
function createNews(data){
  return {
    where: getWhere(data),
    when: getWhen(data),
    who: getWho(data),
    why: getWhy(data),
    what: getWhat(data)
  }
}
```

Fonte: Elaborado pela autora (2023).

Após criar a narrativa jornalística bruta, o sistema deverá operar uma etapa de lexicalização com foco em criar um texto coerente, isto é, nesta etapa, o sistema buscará criar relação entre frases e parágrafos, construindo uma estrutura noticiosa compreensível.

- **Lexicalização:** Momento em que o sistema deverá executar as funções de estruturação do texto entre parágrafos, coerência das frases, lexicalização utilizando as ontologias, a web semântica e *machine learning* para gerar conexões e combinações entre frases, acontecimentos, palavras, etc. É nessa função que a *natural language generation* deve atuar de forma mais aparente, já que suas variáveis utilizam *machine learning* para retornar um texto coerente e contextualizado, levando em consideração o tratamento dos dados efetuado nas etapas anteriores.

FIGURA 38 - CAPTURA DE TELA DO IDE COM FUNÇÃO DE LEXICALIZAÇÃO INTEGRANDO NLG AO SISTEMA.

```
// execute the process of lexicalization
function lexicalization(news) {
  const structured = structuringMachineLearning(news);
  const coherence = coherenceMachineLearning(structured);
  const lexicalized = lexicalizationMachineLearning(coherence)
  return lexicalized;
}
```

Fonte: Elaborado pela autora (2023)

A etapa de lexicalização é de extrema importância para que a notícia atenda a, pelo menos, duas das três etapas elencadas por Lage (2021) para considerarmos um texto em uma narrativa jornalística: para ser uma notícia, o autor sugere que precisa a) relatar um acontecimento, b) obedecer a uma estrutura e c) ser verdadeira. Na etapa da lexicalização, portanto, o sistema irá tratar os dados para atender, ao menos, às características a e b da conceituação de Nilson Lage (2021). Para tanto, incluímos três variáveis distintas: a primeira, com o objetivo de estruturar o texto de acordo com princípios jornalísticos e o manual de estilo do jornalismo; a segunda variável tem o objetivo de conferir coerência às frases e palavras do texto, de acordo com regras gramaticais e probabilidades linguísticas, e a última etapa é a que verifica a lexicalização de acordo com as etapas anteriores, retornando um texto legível. A última variável dentro da etapa de lexicalização é indicada pela seguinte nomenclatura:

FIGURA 39 – ESTRUTURA DE CÓDIGO COM O OBJETIVO DE CRIAR NOTÍCIAS AUTOMATIZADAS.



```

// routine that will run every 1000ms
function createNewsRoutine() {
  const collectedData = collectData();
  // if didn't have any new data, skip the generation process
  if(!collectedData){
    return;
  }
  const classifiedData = classification(collectedData);
  const isValidAndRelevant = validate(classifiedData);
  // if the data is not valid or relevant, skip the generation process
  if(!isValidAndRelevant){
    return;
  }
  const news = createNews(classifiedData);
  const humanFriendlyNews = lexicalization(news);
  publishNews(humanFriendlyNews);
}

```

Fonte: Elaborado pela autora (2023).

Neste caso, utilizamos o termo *humanFriendlyNews* como uma variável dentro do sistema para designar o momento em que a narrativa jornalística é finalizada e, assim, será passível de leitura pelos usuários. Ou seja, após ter passado por todas as etapas dentro do sistema de automação, os dados imputados finalmente tornam-se legíveis para os humanos e o resultado é uma narrativa textual coerente que podemos considerar como uma notícia.

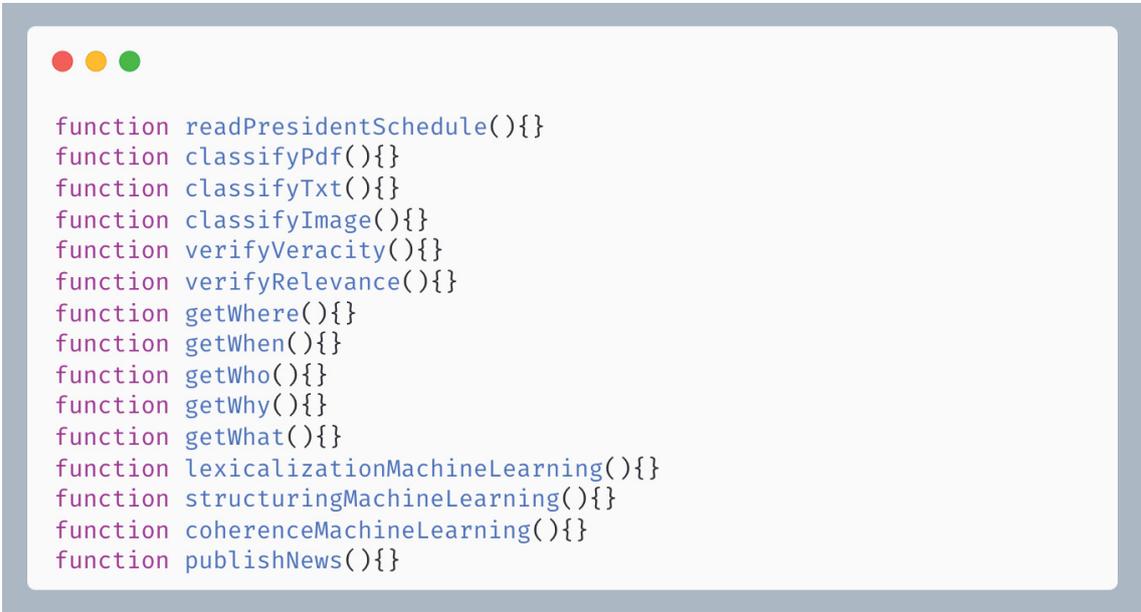
Assim, completando todo o percurso dentro do sistema, o dado se transforma em uma notícia. Nessa etapa, podemos incluir um fechamento automático, indicado a seguir como:

- **Saída:** após completar todas as etapas, o sistema libera a notícia para ser publicada diretamente em seu CMS (sistema de gerenciamento de conteúdo, como por exemplo, o WordPress) e pode, inclusive, liberar o compartilhamento em redes sociais digitais.

Aqui, o sistema é capaz de realizar todas as etapas de forma automática. Se, por exemplo, quiséssemos incluir uma verificação manual, feita por um repórter ou editor humano, na etapa de saída o sistema apenas enviaria o resultado (a notícia) para um ambiente de validação, no qual o repórter teria de abrir a notícia e conferir todas as informações antes de publicá-la.

Para exemplificar a estrutura algorítmica que criamos (com base no esquema dos passos fundamentais de NLG em modo *pipeline* aplicada ao jornalismo) e explicitamos até o momento de forma separada, disponibilizamos os trechos de *JavaScript* na íntegra, com suas funções simplificadas para demonstrar como o sistema deve operar:

FIGURA 40 – ESTRUTURA DE FUNÇÕES DENTRO DO SISTEMA DE CRIAÇÃO DE NOTÍCIAS AUTOMATIZADAS.



```

function readPresidentSchedule(){}
function classifyPdf(){}
function classifyTxt(){}
function classifyImage(){}
function verifyVeracity(){}
function verifyRelevance(){}
function getWhere(){}
function getWhen(){}
function getWho(){}
function getWhy(){}
function getWhat(){}
function lexicalizationMachineLearning(){}
function structuringMachineLearning(){}
function coherenceMachineLearning(){}
function publishNews(){}

```

Fonte: Elaborado pela autora (2023).

Vale ressaltar que, na linguagem de programação, uma função (*function*) é um trecho de código que tem a responsabilidade de executar uma tarefa específica. Frequentemente, uma função faz algum tipo de processamento ou cálculo dentro do sistema e espera-se que ela seja capaz de devolver o resultado desse procedimento (*return*). Dentro de uma função, podemos

utilizar variáveis (*var*) para designar um valor ou uma expressão em relação à função executada (variável local). Ou seja, no caso da função *collectData*, que utilizamos na primeira etapa do nosso sistema, inserimos a variável *var data = readPresidentSchedule()* como um comando para que a função possa acessar a agenda do presidente da república e identificar seus compromissos com o objetivo de gerar notícias relacionadas à agenda presidencial. No caso do trecho de código que escrevemos, ainda fizemos a utilização da condicional *if*, que pode ser considerada como um operador lógico da linguagem de programação, indicando uma condicionante àquela variável.

O exercício de criar uma estrutura de funções com o objetivo de desenvolver um sistema de notícias automatizadas figura como um passo essencial e um grande desafio imposto nesta pesquisa, principalmente pelo fato de ter tomado bastante tempo durante o processo metodológico. O tempo foi um fator importante nesta pesquisa por dois motivos principais:

1. A autora não tinha conhecimento em *JavaScript* e foi em busca de iniciar um processo de aprendizado com o auxílio de um curso online, leituras e conversas com outros programadores para identificar a viabilidade do projeto; e
2. Porque, durante o processo de criação desta estrutura, houve erros que interferiram nas etapas anteriores e posteriores dentro do sistema. A estrutura representada acima passou por três refações para chegar ao modelo que apresentamos neste estudo. Os erros, no entanto, eram facilmente identificados pelo próprio IDE utilizado, que indicava algumas opções de caminhos viáveis. Ainda, a WebStorm auxiliou na identificação de problemas no código e automatizou as possibilidades de código a partir do erro, facilitando o processo com o auxílio de inteligência artificial.

Acima, reproduzimos o funcionamento de um sistema de automação de notícias com base nas contribuições teóricas de diversos autores, como Bertocchi (2016), Devyatkin et.al. (2019) e Ribeiro (2019) porque entendemos que, para criar um *modelo* matemático, somente seria possível compreendendo como opera o sistema. Para justificar nossas escolhas, utilizaremos uma metáfora. Pensemos na tradicional aula de ciências, quando os estudantes, de jaleco branco e óculos de proteção, observam o professor dissecar o corpo de um sapo. Antes da aula prática, o professor os ensinou sobre estruturas anatômicas e explicou-lhes sobre o funcionamento dos órgãos. Embora os estudantes já pudessem aprender apenas com o conteúdo teórico, o professor decide levá-los para um laboratório, onde oferece uma experiência visual, sensorial e prática para que os estudantes observem como o conteúdo teórico acontece na prática.

Só é possível compreender, de fato, a anatomia do sapo observando o sapo. Por isso, o professor leva exemplares de sapos dissecados para que os alunos possam visualizar a anatomia real do anfíbio. A prática de dissecar um animal numa aula de biologia é utilizada justamente para possibilitar aos alunos o desenvolvimento da compreensão sobre como funciona a anatomia do animal, mas também para que os alunos possam colocar em prática todo o conhecimento teórico aprendido em sala de aula. Essa prática de dissecação é o que possibilita aos estudantes o desenvolvimento da observação científica detalhada.

A criação de uma estrutura de sistema de automação de notícias, neste estudo, foi o nosso laboratório. Aqui, não houve dissecação, mas houve, sim, uma análise detalhada da anatomia de um sistema de automação de notícias. Defendemos essa prática porque só podemos modelizar aquilo que é passível de modelização, e compreendemos que, para criar um modelo matemático, precisaríamos reduzir o objeto do "quinto jornalismo" (CARREIRA, 2016), a saber, os sistemas de automação do jornalismo, ao nível anatômico.

Para Russel (1993), o simbolismo lógico é uma resposta às insuficiências da linguagem. O autor vê na matemática uma forma de descrever o mundo como se ela fosse uma linguagem universal, na qual a lógica fosse uma espécie de sintaxe. Segundo ele, "numa tal linguagem poderíamos exprimir todas as proposições da matemática, mesmo se não soubéssemos uma só palavra da linguagem. A linguagem da lógica matemática, caso fosse aperfeiçoada, seria uma linguagem (RUSSELL, 1993, p. 201). Por isso, a partir de agora, após compreendermos, em nível anatômico, como um sistema de automação de notícias funciona, vamos sistematizar cada uma das etapas do sistema em um modelo lógico que, segundo Russell (1993), seria capaz de explicar o fenômeno por meio da matemática.

9.2 TEOREMAS PARA SISTEMAS DE AUTOMAÇÃO DE NOTÍCIAS

Embora ainda identifiquemos espaços para melhorias do sistema de automação proposto acima, constatamos que a estrutura criada nos mostra a anatomia do sistema. Este olhar, que coloca as etapas do sistema de forma ordenada e em relação umas com as outras, é relevante pelo fato de que pudemos observar, de forma sistemática, quais os pressupostos necessários para que o sistema opere. Esses pressupostos nos deram indícios de quais elementos poderíamos tomar como definições lógicas (RUSSELL, 1993) para inserirmos como teoremas que antecedem o modelo matemático que aqui vamos propor. Além disso, criar o sistema nos permitiu colocar à prova todo o arcabouço teórico desta pesquisa em algumas linhas de código.

Pudemos enxergar Shannon (1949), se tomarmos as etapas do sistema com circuitos de comunicação entre um passo e outro do sistema, ou, se observarmos pela entrada do dado e todo o percurso pelo sistema e gerando como resultado uma forma de comunicação. Vimos Bertocchi (2016) no formato da narrativa jornalística digital. Pudemos comprovar os argumentos de Carreira (2017) quando apresentou os pressupostos do "quinto jornalismo", defendendo que os sistemas de automação são a característica primordial desta nova fase de mudanças. Vimos, reproduzidos nas linhas de código, os autores que invocamos durante o percurso teórico desta pesquisa. Vimos Ada Lovelace e Warren Weaver. Vimos Isaacson. Vimos Russell entre funções e variáveis com o objetivo de comprovar que um sistema de notícias pode ser reproduzido por meio de estruturas lógicas.

Assim, o exercício de utilizar uma IDE para (re)criar um sistema de automação de notícias, com base no arcabouço teórico apresentado nesta pesquisa, nos permitiu invocar conceitos e colocá-los à prova, mesmo que de forma breve. Entendemos que o sistema criado acima (adaptativo, complexo e aberto) pode (e deve) ser reformulado e melhorado, entretanto, já serve como uma base para que outros pesquisadores sintam-se encorajados a criar pesquisas multidisciplinares como essa.

Seguimos, portanto, para a apresentação de teoremas que nos darão bases para o modelo lógico do jornalismo de automação. A estrutura lógica proposta a seguir leva em consideração os seguintes princípios do pensamento matemático: o sistema sempre irá trabalhar com uma fonte de informação F , cuja capacidade informativa C pode ser variável:

$$F = C(x, y, z, \dots)$$

A capacidade informativa utiliza o recurso de recorrência de coleta (R) para identificar novos dados a um determinado intervalo temporal (`setInterval(createNewsRoutine, timeout: 1000)`). A função obedece ao comando de `setInterval` utilizando a unidade de medida milissegundos (ms), logo, no caso da estrutura lógica criada acima, a cada um segundo (ou 1000 ms), o sistema deve refazer a coleta de dados na fonte (C variável). Dessa forma, a recorrência (R) de novos dados na fonte informativa (F) pode ser representada por:

$$F = C(x, y, z, \dots) \times R$$

Sendo assim, é possível inferir que F pode ser considerada como uma grandeza de valor infinito, já que a cada recorrência o sistema irá retornar-lhe uma quantidade que pode ser 0 até

um número infinito de resultados (∞). No caso que utilizamos como exemplo, no qual o sistema deve fazer a consulta à agenda presidencial, no entanto, o número de resultados será variável, porém há a possibilidade de ser limitado. Mas tomando como base a estrutura apresentada, da qual a fonte de informação seja, por exemplo, os sites governamentais, coletando dados de diversos setores, a fonte de informação poderá retornar uma quantidade infinita de resultados a cada episódio de recorrência.

Já na etapa de classificação dos dados, o sistema identifica um número x de formatos (PDF, txt. e imagem) e os converte para um símbolo finito e não variável y , que designa o formato textual da narrativa jornalística (neste caso, no sistema apresentado, fizemos a conversão apenas para texto). Esta etapa é necessária para extrair as informações de acordo com a natureza do dado bruto e convertê-las para o formato final da notícia. Em sistemas mais avançados e com estruturas mais rebuscadas, y também pode ser variável, de acordo com os tipos de formatos suportados pela aplicação (vídeo, áudio, gráfico, etc.). Sendo assim, o teorema que representa a etapa de classificação do dado pode ser representado, por simbolismo lógico, da seguinte maneira:

$$3x = y$$

O exercício, neste momento, é o de criar representações simbólicas a respeito das estruturas algorítmicas criadas para ilustrar o funcionamento de um sistema de automação de notícias por meio de deduções lógicas, conforme aponta Russell (1993), e como o fez Shannon (1963) quando propôs a "Teoria matemática da comunicação" por meio da representação gráfica do seu sistema. Consideramos esta etapa de pensamento lógico essencial para realizar a aproximação entre os conceitos do jornalismo (informação, dados, comunicação, transmissão) com os conceitos da matemática (função, variável, condicionais, lógica).

Nesse ponto, buscamos ilustrar, de forma didática, as possibilidades de aproximação teórica entre o jornalismo e a matemática, de modo a criar uma linha de pensamento multidisciplinar (levando também em consideração o fato de que, além dessas duas grandes áreas, a tese que propomos aqui recorre a conceitos de outros campos intelectuais, como a ciência da computação, cujos princípios aparecem em diversos tópicos deste estudo, inclusive nos procedimentos metodológicos). Embora trabalhemos com representações e sistemas imaginados, é importante destacar que toda proposição aqui apresentada carrega preceitos teóricos e conceituais, mas representa também uma aplicação prática, passível de verificação empírica, como é o caso da estrutura algorítmica em *JavaScript*, cujas linhas de código foram

testadas dentro da IDE e verificadas com o auxílio do próprio sistema de inteligência artificial da WebStorm.

Os passos de validação e criação da notícia também podem ser representados com uma equação matemática, utilizando da lógica algorítmica. Na etapa de validação, consideremos que D se refere aos dados que constam dentro do sistema após a coleta e a classificação. Este é o momento em que o sistema deve verificar a veracidade e a relevância daquele dado, assim como um jornalista o faz quando recebe uma informação e faz a posterior checagem da informação. Neste caso, o sistema proposto deverá validar se o dado corresponde a ambas as variáveis (veracidade e relevância) para avançar para a etapa de criação da notícia. Esta é uma etapa que corresponde exatamente ao mesmo posicionamento esperado por jornalistas humanos dentro de uma redação, uma vez que somente partem para a etapa de redação após conferir a veracidade dos dados coletados nas etapas anteriores.

Consideramos essencial fazer esse paralelo, comparando a atividade do sistema de automação às atividades da rotina de um jornalista que trabalha efetivamente dentro de uma redação, conforme aponta Lindén (2018), quando defende que o funcionamento de um software de automação é similar ao trabalho do jornalista, cujos princípios são estritamente baseados em regras. Logo, se esse trabalho é baseado em regras, é passível de simbolismo lógico matemático. A validação das informações, portanto, é uma das regras principais e que rege o funcionamento da prática jornalística, sendo indispensável tanto ao trabalho do jornalista, quanto à atuação de um sistema de automação. Dentro da estrutura algorítmica proposta, podemos sintetizar a lógica da validação como uma soma simples da veracidade com a relevância do dado, conforme sugerimos a seguir:

$$V \equiv v + r$$

Onde V (maiúsculo) corresponde à validação que deve ser equivalente (\equiv) a v (veracidade) somado à r (relevância). O sistema somente irá reconhecer a informação como válida se o dado corresponder às duas variáveis (*var isValid; var isRelevant*), ou seja, se não tiver êxito em uma das variáveis, a validação não acontecerá.

Após a validação das informações, tanto um sistema quanto um jornalista passam para o momento de escrita, ou seja, da efetiva produção da narrativa jornalística: a notícia. Na faculdade de jornalismo, por exemplo, o primeiro princípio jornalístico que é repassado aos alunos, se não na primeira aula, no primeiro contato dos estudantes com as regras do jornalismo, é a estrutura do *lead*. O *lead* corresponde às informações mais importantes relacionadas a

determinado acontecimento e é descrita como cada uma das respostas às cinco perguntas: Quem? O que? Quando? Onde? Por quê?

Geralmente, o *lead* jornalístico aparece no primeiro parágrafo de uma notícia, indicando as informações mais importantes em relação ao fato noticiado. De acordo com essa regra, é possível configurar um sistema de automação para buscar, sempre que possível, responder a essas perguntas dentro do primeiro parágrafo da notícia gerada. Sendo assim, como vimos acima, na estrutura algorítmica apresentada, o sistema reconhece que, quando os dados correspondem às perguntas do *lead*, é possível gerar uma notícia a partir dessas informações. Lindén (2018) argumenta que, com o avanço das pesquisas em web semântica e a utilização de ontologias, sistemas de automação são capazes de entregar resultados ainda mais precisos, com informações que vão além do *lead* jornalístico, trazendo contextualizações, histórico de acontecimentos similares e outras conexões que somente são possíveis com o uso de inteligência artificial e *machine learning*. O *lead*, considerado dentro de um teorema matemático de acordo com sua estrutura algorítmica, pode ser representado da seguinte forma:

$$L = w \times 5$$

Onde L representa o *lead* e w as perguntas *who? where? when? why? what?*. Em relação aos outros teoremas apresentados, este pode parecer um cálculo simplista, embora represente de forma clara o resultado esperado, que é o lide, ao final do processo. A lexicalização corresponde à etapa em que a notícia passa a receber alguns tratamentos para que as frases e parágrafos façam sentido entre si. Um teorema matemático, por dedução lógica, pode ser ilustrado da seguinte forma:

$$N = l (v^2)$$

Sendo v^2 a representação das variáveis da lexicalização: *machine learning* e coerência entre as frases. Essa variável elevada à potência de dois significa que o processo de lexicalização depende dessa operação para completar-se. Aqui, ainda, o sistema conta com *machine learning* e destacamos o ponto porque em relação a essa tecnologia, embora seja possível de ser representada em um teorema matemático, devemos ter o cuidado de olhar para o símbolo utilizado levando em consideração que *machine learning*, por si só, invoca uma série de questões relacionadas à sua própria complexidade. Mesmo os programadores, matemáticos e estatísticos mais experientes do mundo não se atreveram a delimitar o entendimento de

machine learning em poucos símbolos. Por essa razão, a lexicalização não pode ser considerada de forma finita ou estanque, já que, dentro de sua constituição, existe uma estratégia de *machine learning*, cuja principal característica é a complexidade e a mutabilidade constantes.

A última etapa do processo realizado dentro de um sistema de automação é a de saída, já com a notícia escrita, quando o sistema envia diretamente para o CMS (*content management system*) o produto final (narrativa jornalística) para posterior publicação. Aqui, chegamos à etapa que conclui o processo de criação de uma notícia dentro do sistema, e buscamos reproduzi-la, utilizando-nos da linguagem simbólica elencada por Russell (1993), da seguinte forma:

$$S = [N (F + C + V + L + l)]$$

Essa é uma representação simplista de teoremas que busca dar conta de exprimir todas as fases que criamos dentro do sistema de automação de notícias, onde entendemos:

S como a etapa de saída;

F como fonte de informação (onde o dado é coletado e entra no sistema);

C como a classificação dos dados;

V como o processo de validação;

L como *lead*, primeiro passo da criação da notícia;

l como o processo de lexicalização;

S como a saída dos dados do sistema.

Nesse teorema, *N* simboliza a notícia, que só será criada após corresponder a todas as etapas que estão previstas entre parênteses, ou seja, precisamos resolver todas as funções dos parênteses para chegarmos à notícia, que colocamos entre colchetes, sendo o resultado obtido da soma de *F*, *C*, *V*, *L* e *l* e que corresponde à *S*, etapa de saída da notícia de dentro do sistema. Assim, neste ponto, pudemos observar que cada uma das etapas realizadas pelo sistema de automação é passível de uma representação matemática, por meio de teoremas. Sodré (2007) enfatiza que um teorema é uma proposição que pode ser deduzida por meio de uma observação lógica sobre determinado fenômeno. Embora essa seja uma representação matemática *simbólica*, pudemos verificar que Russell (1993) tinha razão ao afirmar que qualquer fenômeno "ordinário", ou seja, do mundo real, pode ser analisado por meio de proposições lógicas, utilizando os preceitos matemáticos como uma linguagem universal de análise. Em sua obra

Introdução à Filosofia Matemática, Russell (1993) argumenta que precisamos "forçar" as palavras para que possamos descrever determinados fenômenos "ordinários" em proposições lógicas. Isso quer dizer que, sob o ponto de vista do autor, apesar de seu uso corrente, a linguagem ainda é insuficiente para descrever a realidade, isto é, segundo ele, a linguagem seria um recurso limitado e gramaticalmente enganoso.

Este é o caso, por exemplo, em relação aos números; 'dez homens' é, gramaticalmente, da mesma forma que 'homem branco' de modo que se pode pensar que 10 seja um adjetivo que está a qualificar homens'. Analogamente, quando estão envolvidas funções proposicionais, e, em especial, no tocante à existência e às descrições. Porque a linguagem é enganosa, e também porque ela é difusa e inexata quando aplicada à lógica (para a qual não era intencionada), o simbolismo lógico é absolutamente necessário para um tratamento exacto ou completo do nosso assunto. (RUSSEL, 1993, p. 201)

Assim, Russell (1993) nos possibilita olhar para os fenômenos ou "assuntos" de interesse de modo a compreender as inúmeras possibilidades que a simbologia lógica nos permite, desprendendo-nos da noção de linguagem atrelada apenas à representação visual das letras e o significado das palavras. O autor considera que precisamos ir além das fronteiras da linguagem corrente, já que, sob sua perspectiva, é possível criar deduções lógicas sobre qualquer assunto, desde que seu enfoque esteja na compreensão matemática deste. E, como citamos no capítulo 7, a lógica e a matemática não existem uma sem a outra. Como o próprio Russell (1993) argumentou, "a lógica é a juventude da matemática e a matemática é a maturidade da lógica." (p. 191).

9.3 UM MODELO MATEMÁTICO PARA O JORNALISMO

Tudo o que apresentamos neste estudo, até o momento, tem um propósito: possibilitar a criação de sinapses. Fizemos um caminho interdisciplinar, recorrendo a estudos sobre tecnologia, jornalismo, ciência da informação e até à filosofia. Isso porque compreendemos que uma tese de doutorado se constrói na intersecção do conhecimento; na assimilação de saberes; nas sinapses cognitivas. Essas sinapses, ou ligações, têm como objetivo evidenciar a aproximação entre o jornalismo e a matemática. Assim, atingimos este tópico, após um percurso teórico-metodológico que buscou atrelar o desenvolvimento da técnica ao desenvolvimento do jornalismo até chegarmos ao período que consideramos como a quinta fase do jornalismo (CARREIRA, 2016).

A principal característica deste período jornalístico é a automação, possibilitada por meio do desenvolvimento de sistemas de automação de notícias. Resgatamos, neste momento,

todo o percurso conceitual desta tese na tentativa de justificar as escolhas teóricas que fizemos até aqui, para que você, interlocutor, seja capaz de criar as sinapses necessárias para compreender as razões deste estudo. O intuito era o de criar uma perspectiva panorâmica acerca da evolução da técnica, demonstrando que o jornalismo também evoluiu, em grande medida, graças aos avanços tecnológicos, em cada uma de suas fases, como aponta Lage (2021). No quinto jornalismo, vemos ascender discussões acerca de um novo fenômeno, que antes parecia fazer parte de um cenário longínquo, um papo demasiadamente tecnófilo: o da automação.

Entendemos, portanto, o jornalismo de automação como um fenômeno técnico e social da contemporaneidade; um fenômeno cujas limitações e possibilidades devem ser pesquisadas por meio de diferentes métodos, dentro do ambiente científico. A lente que usamos neste estudo foi a da matemática. Gehlen (2016) aponta que, desde a origem dos números, filósofos e matemáticos buscam por leis universais para sistematizar os fenômenos do universo, sejam eles de ordem prática ou técnica, por meio da investigação da natureza e de pesquisas experimentais e exploratórias (2016, p. 52). E, na matemática, esses pesquisadores, inquietos, encontraram formas de visualizar o mundo por meio de funções, já que "tudo o que existe passa a ser concebido como função e, ao mesmo tempo, sujeito ao cálculo técnico e matemático." (HEIDEGGER, 1988, p. 106). Outro autor que cita a matemática como uma forma de investigação dos fenômenos da natureza é Francisco Rüdiger, quando relata que "a matemática deixou de ser simples arte liberal para se tornar o veículo da ciência moderna da natureza." (RÜDIGER, 2014, p. 104).

Por isso, defendemos que o fenômeno do jornalismo de automação pode ser observado pelas lentes da matemática, uma vez que buscamos reduzir as etapas da automação a teoremas, por meio da dedução lógica defendida por Russell (1993). Neste momento, portanto, nos aventuramos na proposição de um modelo matemático para o jornalismo, tendo em vista toda a jornada que percorremos até aqui.

Segundo Sodré (2007), um modelo matemático "pode ser apresentado como uma representação de um sistema real, o que significa que um modelo deve representar um sistema e a forma como ocorrem as modificações no mesmo." (p. 03). O autor, em consonância com Russel (1993), afirma que o ato de modelar pode ser aplicado a uma variedade de problemas "ordinários", como, por exemplo, "o estudo da análise ambiental nas proximidades de um rio, a forma da asa de um avião, um sistema econômico, uma cultura agrícola, um estudo populacional, um estudo físico..." (SODRÉ, 2007, p. 03), ou seja, é possível criar modelos matemáticos para compreender diversas situações do mundo e os modelos matemáticos, podem, inclusive, descrever o mundo de forma tão precisa quanto o próprio mundo real (Idem, p. 04).

Conforme Sodré, existem dois tipos de modelos matemáticos mais comuns utilizados no meio acadêmico: o modelo mecanístico e o modelo empírico. Neste estudo, optamos por seguir a linha de pensamento do modelo matemático mecanístico, o qual "pode ser construído pela visão da estrutura do sistema, dividindo-se o sistema em várias componentes e tentando entender o comportamento de todo o sistema através de cada parte e através das interações que ocorrem com as partes." (SODRÉ, 2007, p. 5). O autor relata que os modelos podem ser tanto reais, quanto abstratos, desde que busquem evidenciar a compreensão sobre algum fenômeno da ordem do real. Para tanto, "a formulação de um modelo matemático envolve a escolha de variáveis mais importantes para o fenômeno em estudo e de leis consideradas obedecidas pelas variáveis escolhidas." (p. 10), ou seja, é necessário elencar algumas regras para que o modelo possa ser considerado adequado, mesmo que sua resolução não seja sempre um resultado quantitativo, por exemplo.

O pesquisador Jonei Cerqueira Barboza utiliza o termo "representação de alguma coisa" para conceituar o termo "modelo". Para ele, os modelos são partes substanciais das práticas científicas e possui a denotação das "representações simplificadas, mentais ou não, que os seres humanos fazem sobre a realidade (ou a suposta realidade)" (2009, p. 70). Sendo assim, nas linhas que se seguem, buscamos apresentar um modelo matemático para o jornalismo, no intuito de representar a realidade que vivemos na fase do quinto jornalismo: a de um jornalismo permeado, marcado e caracterizado pelos sistemas de automação de notícias. Para a criação do modelo, elencamos alguns pressupostos relevantes:

1. Escolha do fenômeno;
2. Escolha das variáveis;
3. Desenvolvimento de equação diferencial;
4. Simulação simbólica

O fenômeno escolhido, como já apresentado, é o da automação no jornalismo. Suas variáveis são os procedimentos, ou etapas, que um sistema de automação realiza para criar notícias de forma automatizada (vide capítulo 9, subcapítulo 9.1, sobre a anatomia dos sistemas). Sendo assim, o modelo matemático para o jornalismo de automação pode ser representado da seguinte forma.

Entendendo o fenômeno da automação das notícias e as variáveis como as etapas dentro do sistema de automação, podemos inferir que:

- Entrada dos dados (E): representa a quantidade de dados disponíveis para o sistema de automação;

- Classificação dos dados (C): indica a eficiência do sistema em classificar e organizar os dados;
- Validação dos dados (V): mede a precisão do sistema em verificar a veracidade e a confiabilidade dos dados;
- Criação da notícia (N): representa a estruturação da notícia gerada pelo sistema de automação;
- Publicação (P): indica capacidade de publicação das notícias geradas.

Desse modo, as equações diferenciais podem ser apresentadas da seguinte forma:

Equações:

- $C = x(E)$: A classificação dos dados depende da quantidade de dados de entrada;
- $V = x(C)$: A validação dos dados depende da classificação realizada pelo sistema;
- $N = x(V)$: A produção da notícia depende da validação dos dados;
- $P = x(N)$: A publicação do conteúdo depende da criação da notícia, na etapa N.

Vale ressaltar que, nesta simulação, o modelo não opera com uma entrada de dados infinita, sendo $E = 0$. Se observássemos o modelo com uma entrada de dados finitos, como 5 dados, poderíamos representar a quantidade de dados como

$$E = n + 5$$

Neste modelo, levamos em consideração o *modus operandi* de um sistema de automação de notícias, objeto característico da quinta fase do jornalismo, segundo Carreira (2016). Essa simulação demonstra que cada etapa do sistema de automação (variável) poderá ser representada por uma equação matemática e o comportamento da equação seguinte será condicionado pelo êxito da função anterior, num processo que Russell (1993) considera como o princípio da progressão matemática.

O exercício que fizemos aqui, foi o de olhar para o sistema de automação de notícias como se fosse um sapo: dissecando-o em suas partes para podermos observá-lo a nível anatômico. Somente assim, olhando o sistema detalhadamente e compreendendo seu funcionamento, é que foi possível chegar a esses modelos matemáticos. Vale ressaltar que trouxemos representações simbólicas que utilizam da dedução lógica proposta por Russell (1993) para representar o fenômeno da automação do jornalismo de acordo com a matemática.

Destacamos a palavra *simbólica*, porque o próprio Russell (1993) defende o uso da "simbologia logicista" para descrever o que considera como os "fenômenos ordinários" do cotidiano.

CONSIDERAÇÕES FINAIS

Um teorema traz-nos uma dedução lógica de algo. Um modelo traz-nos uma representação da realidade. O jornalismo? Busca a mesma coisa: representar a realidade. Embora tenham andado separados por séculos, o jornalismo e a matemática reencontraram-se nas páginas desta tese. Fornecemos ao leitor uma espécie de encontro inusitado: colocando o jornalismo a operar com números, e os números a operarem com as letras do alfabeto. Há quem possa dizer que esta é uma visão romântica, até poética da ciência. Seria um elogio, tendo em vista que Ada Lovelace também foi acusada de cometer esse delito acadêmico.

A precursora da "ciência poética", que inaugurou uma nova fase nos estudos sobre tecnologia e é considerada a primeira programadora da história, não por acaso, era também uma estudiosa da matemática e trouxe-nos contribuições valiosas para pensarmos no modelo matemático para o jornalismo. Segundo Isaacson (2014), Ada "ajudou a plantar as sementes de uma era digital que floresceria cem anos mais tarde" (p. 45). Na primavera em que vivemos, as sementes de Ada parecem ter crescido e, de fato, germinado. O jornalismo de automação pode ser considerado um de seus frutos, que modificou o *modus operandi* do jornalismo, porém não foi capaz de interferir no *ethos* da profissão. A busca por relatar a realidade continua, só que, agora, é feita por repórteres-robôs.

Vimos, no capítulo 6, que existem diversos sistemas de automação de notícias no mercado. Alguns já estão estabelecidos no mercado profissional e são utilizados por grandes corporações, como a *Associated Press* e a *BBC News*. Outros, como ChatGPT, embora não sejam ferramentas criadas com o objetivo de automatizar notícias, são capazes de criar narrativas jornalísticas (mesmo que com alguns problemas de precisão e veracidade das informações). Os avanços da automação nas redações jornalísticas já não é mais uma previsão para o futuro ou um delírio tecnófilo, e sim uma realidade que centenas de jornalistas já enfrentam pelo mundo afora. Essa realidade, no entanto, traz consigo novos desafios e aberturas teórico-epistemológicas. É a ciência a responsável por compreender os movimentos decorrentes dessa nova realidade. Do mesmo modo, é papel também da ciência observar os impactos da automação no jornalismo não apenas do ponto de vista técnico, mas também social, econômico, político, educacional e, quem diria: humano.

O jornalismo de automação, assim, é um objeto de pesquisa que demanda uma investigação profunda e interdisciplinar. Não são as lentes do jornalismo as únicas capazes de observar e explicar tal fenômeno, tampouco apenas as lentes da matemática ou da ciência da informação. Por isso, na trilha desta pesquisa, invocamos um quadro conceitual desenhado por

pesquisadores de diversas áreas: jornalistas, filósofos, matemáticos, programadores, professores, etc., que trouxeram, cada um à sua medida, alguma contribuição para compreender melhor o fenômeno técnico do jornalismo de automação.

Neste estudo, deixamos claro que a lente de observação, embora multidisciplinar, focou o fenômeno técnico, atrelando o jornalismo de automação à matemática. No entanto, não podemos fechar os olhos para as tantas outras faces do fenômeno, por isso, registramos, durante o percurso da pesquisa, e reiteramos nestas considerações finais, a urgente necessidade de investigar-se a automação jornalística por diversas lentes, buscando compreender limitações e possibilidades deste objeto de pesquisa que é tão característico do que convencionamos chamar de "quinto jornalismo" (CARREIRA, 2016).

Desse modo, na impossibilidade de esgotar as análises sobre o tema, buscamos focar apenas um direcionamento, compreendendo que existem outros assuntos relacionados que também precisam ser refletidos cientificamente. Ainda assim, mesmo que não tenha sido um objetivo explícito desta pesquisa, buscamos evidenciar, de forma breve, os aspectos éticos, sociais e econômicos que estão relacionados ao tema (especialmente no capítulo 5, no qual discorreremos de forma mais aprofundada sobre o conceito de jornalismo e suas fases). Deixamos, assim, registradas as nossas observações sobre a necessidade de pesquisarmos de maneira científica o fenômeno da automação do jornalismo, colocando o objeto em análise não apenas pelo microscópio das ciências sociais aplicadas, mas trazendo outras áreas do conhecimento para contribuir no desenvolvimento da ciência e na construção de novos saberes acerca do assunto.

Optamos por olhar para o fenômeno sob uma perspectiva diferente. Diferente porque vai de encontro a toda a formação da autora desta tese. Diferente porque vimos, no estado da arte (capítulo 8), que foram poucas ou praticamente inexistentes as tentativas de aproximar o jornalismo da matemática. Como aponta Gehlen (2016), o jornalismo e a matemática parecem ter se distanciado ao longo das décadas, embora muitas reportagens e esforços investigativos predominantes no trabalho do jornalista necessitem de um conhecimento (ou, ao menos, noções) da matemática, da estatística e/ou da lógica.

Trouxemos, nas primeiras páginas deste trabalho, lá na introdução, uma frase cuja autoria não encontramos em nossas pesquisas. Alguns relatos atribuem a autoria da frase a Galileu Galilei, outras a Johann Kepler e Pierre-Simon Lapla. Buscamos identificar, na internet, a origem da frase, porém não obtivemos êxito. Embora fosse importante sabermos de fato o autor da frase, o mais relevante, para esta tese, é o significado que extraímos dela: "a matemática é a linguagem pela qual Deus escreveu o mundo.". Não nos cabe, aqui, uma discussão sobre a

existência ou não de um Deus, criador de todas as coisas. Mas, a título de metáfora, podemos captar a essência de tudo o que essa frase é capaz de carregar: ela nos indica que o mundo foi escrito pela matemática. Não à toa Russell (1993) também nos apresenta o entendimento da matemática enquanto linguagem. Uma linguagem capaz de descrever o mundo por meio da simbologia lógica.

Essas noções são vitais para compreender-se a forma como conduzimos esta pesquisa. Isso porque vislumbramos uma maneira de sistematizar o jornalismo de automação, enquanto fenômeno da realidade, em um modelo matemático. Aqui também entendemos a matemática enquanto linguagem, enquanto um alicerce capaz de descrever situações da ordem do real. Tomando, especialmente, as contribuições de Shannon (1963) e Russell (1993) como base, pudemos observar o sistema de automação de notícias como um produto da matemática.

Buscamos assimilar os conceitos teóricos a fim de justificar nossas escolhas metodológicas (e nada ortodoxas) em relação à proposição de um modelo matemático para o jornalismo. Para chegarmos a este modelo, no entanto, buscamos dissecar a forma como os sistemas de automação operam, examinando cada uma das etapas do trabalho do "repórter-robô" até chegar à publicação da notícia. Conforme apontou Ribeiro (2019), um sistema de automação opera de forma bastante similar ao trabalho humano dentro de uma redação: ele coleta um dado, classifica, valida, escreve a notícia com base em determinadas regras de redação, revisa e publica. Essas etapas, já tão disseminadas na práxis jornalística, passaram a ser reproduzidas pelos sistemas de automação na contemporaneidade. Podemos afirmar, portanto, que, no quinto jornalismo, o *modus operandi* do "fazer jornalístico" transmuta-se, enquanto o *ethos*, isto é, o conjunto de postulados que caracteriza o jornalismo, não se altera.

Após dissecar o objeto de pesquisa, a saber, o sistema de automação de notícias, buscamos olhar para cada uma das etapas a fim de criar teoremas, por meio da dedução lógica proposta por Russell (1993). Os teoremas foram criados com o objetivo de sistematizar cada uma das etapas do processo de tratamento do dado até a criação e publicação da notícia. Esses teoremas, apresentados no capítulo 9, foram a primeira aproximação entre o jornalismo e a matemática na arena prática desta pesquisa. Já o nosso modelo matemático, utilizou uma série de variáveis, interligadas entre si, para criarmos uma representação simbólica capaz de dar conta do fenômeno do jornalismo de automação de maneira técnica. A construção do modelo sinalizou-nos que é possível reduzir o *modus operandi* do jornalismo a um modelo matemático, entretanto, dentro desse modelo, existem limitações que extrapolam a simples representação simbólica. Explicaremos, a seguir, quais são algumas dessas limitações percebidas.

No modelo, não mostramos as demissões em massa de jornalistas humanos em “benefício” de um repórter-robô (embora tenhamos utilizado uma nota de rodapé para destacar esse movimento, conhecido entre a comunidade jornalística como *passaralho*). No modelo matemático, não contamos com a entropia que pode ser causada pelo excesso de informações falsas que circulam no ambiente digital, por exemplo. No modelo, embora possa, de fato, sistematizar o trabalho do jornalismo, não incluímos a subjetividade característica do jornalista humano. Focamos o ato de estruturar, de forma lógica, o funcionamento de um sistema de automação jornalístico. Então, do ponto de vista técnico, atingimos o objetivo deste estudo.

Do ponto de vista científico, vislumbramos uma série de novas possibilidades que foram criadas a partir deste estudo. A principal delas, mais ousada, é a proposição de uma nova teoria do jornalismo baseada nas observações levantadas nesta tese. Isso só o futuro nos dirá. Mas já podemos afirmar que a automação não apenas inaugura a quinta fase do jornalismo, mas que também desestabiliza tudo o que se sabia sobre "como fazer jornalismo" até então. No quinto jornalismo não há espaço para saudosistas. E talvez, de fato, seja o momento ideal para propormos uma nova teoria. Aqui, vislumbramos uma teoria matemática do jornalismo como um caminho que poderá dar conta de explicar a grandiosidade desse fenômeno que nos permeia.

REFERÊNCIAS

- ANDERSON, Chris. BELL, Emily; SHIRKY, Clay. Jornalismo pós-industrial: adaptando aos novos tempos. Revista de Jornalismo ESPM: Revista da Escola Superior de Propaganda e
- ARCE, Tacyana. O Lead Automatizado: Uma Possibilidade de Tratamento da Informação para o Jornalismo Impresso Diário. Revista e-xacta, V. 2, N. 3 (2009)
- BARBOSA, Suzana. Jornalismo digital em base de dados (JDBD) – um paradigma para produtos jornalísticos digitais dinâmicos. 2007. Tese (Doutorado em Comunicação e Cultura Contemporâneas) - Universidade Federal da Bahia, Salvador.
- BARROSO, S.; MARQUES, A. Jornalismo Automatizado: um olhar sob a produção nos periódicos Journalism Studies, Digital Journalism e Journalism. In: XI International Conference on Online Journalism, 2021.
- BERTOCCHI, Daniela. Dos dados aos formatos-Um modelo teórico para o design do sistema narrativo no jornalismo digital. 2013. Tese de Doutorado. Universidade de São Paulo.
- BERTOCCHI, Daniela. Dos dados aos formatos: a construção de narrativas no jornalismo digital. Appris Editora e Livraria Eireli-ME, 2016.
- BROOKES, B.C. The foundations of Information Science: Part I. Philosophical aspect. Journal of information Science, v.2, n.3-4, p.125-133, Jun.1980.
- CARREIRA, K. A. C. Notícias Automatizadas: A evolução que levou o jornalismo a ser feito por não humanos. 2017.
- CARREIRA, Krishma; SQUIRRA, Sebastião. JORNALISMO AUTOMATIZADO, GERAÇÃO DE LINGUAGEM NATURAL E A LÓGICA DO BOM SUFICIENTE. Revista Observatorio, v. 3, n. 3, p. 60-84, 2017.
- CASTELLS, Manuel. A era da informação: a sociedade em rede. Tradução de Alexandra
- COSTA, C. R.. O futuro do trabalho do Jornalista é o digital. Líbero, v. 43, p. 43-54, 2019.
- COSTELLA, Antonio F. Comunicação: do grito ao satélite. Campos do Jordão: Mantiqueira, 2001.
- COUTINHO, Iluska; SILVEIRA JR, Silveira Júnior. Comunicação: tecnologia e identidade. Mauad Editora Ltda, 2007
- CUPANI, Alberto. Filosofia da tecnologia: um convite. Editora da UFSC, 2016
- DAVENPORT, Thomas H. Ecologia da informação: por que só a tecnologia não basta para
- DEVYATKIN, Dmitry; ISAKOV, Vadim; SHVETS, Alexander. Genetic algorithm based sentence packaging in natural language text generation. In: IOP Conf. Ser.: Mater. Sci. Eng., 2019

Diakopoulos, N. (2019). Automating the news: How algorithms are rewriting the media. Harvard University Press.

DOS SANTOS, Márcio Carneiro. A datificação de um campo de conhecimento: como algoritmos, números e abordagens quantitativas estão mudando a comunicação. *Organicom*, v. 16, n. 31, p. 145-157, 2019.

DUSEK, Val. Filosofia da tecnologia. Tradução de Luis Carlos Borges. São Paulo: Loyola, 2009.

ESCOSTEGUY, Ana Carolina D.; GUTFREIND, Cristiane Freitas. Leituras em comunicação, cultura e tecnologia. EdiPUCRS, 2007.

EXAME. Tecnologia de geração de linguagem natural da Arria expande cobertura eleitoral da BBC no Reino Unido. *Revista Exame*. 17 de dezembro de 2019. Disponível em: <<https://exame.abril.com.br/negocios/releases/tecnologia-de-geracao-de-linguagem-natural-da-arria-expande-cobertura-eleitoral-da-bbc-no-reino-unido/>>

FEENBERG, Andrew. O que é a filosofia da tecnologia. In: Conferência pronunciada para os estudantes universitários de Komaba. 2003.

FRANÇA, Vera; ALDÉ, Alessandra; RAMOS, Murilo C. (Orgs.) Teorias da Comunicação no Brasil: reflexões contemporâneas. Salvador: Edufba, 2014. p.39-62.

GALIMBERTI, Umberto. Psiche e techne: o homem na idade da técnica. Paulus, 2006.

GALLOWAY, Alexander R. The cybernetic hypothesis. *differences*, v. 25, n. 1, p. 107-131, 2014.

GLEICK, James. A informação: uma história, uma teoria, uma enxurrada. Tradução de

GUEDES, W. A Teoria Matemática da Comunicação e a Ciência da Informação. <http://repositorios.questoesemrede.uff.br/repositorios/bitstream/handle/123456789/1130/A%20teoria%20-%20Guedes.pdf?sequence=1>

KIM, Joon Ho. Cibernetica, ciborgues e ciberespaço: notas sobre as origens da cibernetica e sua reinvenção cultural. *Horizontes antropológicos*, v. 10, n. 21, p. 199-219, 2004.

KUNCZIK, Michael. Conceito de jornalismo: norte e sul. Tradução de Rafael Varela Jr. 2.ed. 1.reimp. São Paulo: Editora da Universidade de São Paulo, 2002.

LAGE, Nilson. Estrutura da notícia. São Paulo: Ática, 1985a.

_____. Linguagem jornalística. São Paulo: Ática, 1985b.

LEMOS, A. Cibercultura: tecnologia e vida social na cultura contemporânea. Porto Alegre: Sulina, 2002.

LÉVY, P. Cibercultura. Tradução de Carlos Irineu da Costa. São Paulo: Editora 34, 1999.

LIMA JÚNIOR, Walter Teixeira de. Big Data, Jornalismo Computacional e Data Journalism: estrutura, pensamento e prática profissional na Web de dados. In: Estudos em Comunicação. Dezembro de 2012. V. 12, p.207-222.

LINDÉN, C.-G. (2017). Decades of Automation in the Newsroom. *Digital Journalism*, 5(2), 123–140. <https://doi.org/10.1080/21670811.2016.1160791>

LINDÉN, Carl-Gustav. Algoritmos para Jornalismo: o futuro da produção de notícias. *Líbero*, v. 41, p.05-27, 2018.

MACHADO, Elias. O ciberespaço como fonte para jornalistas. 2002. Disponível em: <http://bocc.unisinos.br/pag/machado-elias-ciberespaco-jornalistas.pdf> >.

MAGALHÃES, Daniel Lima. Precisão, rapidez e robôs: um panorama atual do Jornalismo algorítmico. In: *Revista Temática*. Ano XIII, n. 08 - 2017.

MATTELART, Armand; MATTELART, Michèle. História das teorias da comunicação. Tradução de Luiz Paulo Rouanet. 15.ed. Loyola: São Paulo, 2012.

MEDITSCH, Eduardo. Rádio. In: MARCONDES FILHO, Ciro (Org.). *Dicionário da comunicação*. 2.ed. rev. e aum. São Paulo: Paulus, 2014. p. 394-395.

MENDONÇA, Vinícius de Sousa. Notícias Geradas por Software: O Jornalismo Sem Repórter. 2016

MIELNICZUK, Luciana. Características e implicações no jornalismo na web: o link como elemento paratextual. In: *Anais X Encontro Nacional da COMPÓS (GT Jornalismo)*. Brasília,

MILOSAVLJEVIC, M.; VOBIC, I. Our task is to desmitify fears: analysing newsroom management of automation in journalism. In: *Journalism*, 1:1-19, 2019.

MORTENSEN, C. David. Teoria da comunicação: textos básicos. São Paulo: Mosaico, p. 119-138, 1980
o sucesso na era da informação. São Paulo, Futura, 1998.

PAVLIK, John. *El periodismo y los nuevos medios de comunicación*. Barcelona: Editorial Paidós Comunicación, 2005.

PEISE, Jaclyn. The Rise of the Robot Reporter. *The New York Times*. 05 de fevereiro de 2019. Disponível em:
<<https://www.nytimes.com/2019/02/05/business/media/artificial-intelligence-journalism-robots.html>>

QUADROS, Claudia Irene de. El futuro del periodismo en Internet. In: *IV Jornadas Internacionales de Jóvenes Investigadores en Comunicación*, 1997, Bellaterra -. *IV Jornadas Internacionales de Jóvenes Investigadores en Comunicación*. Bellaterra: Servei de Publicacions de la Universitat Autònoma de Barcelona, 1997. p. 114-122.

QUADROS, Claudia Irene de. Periodistas y diarios electrónicos: las exigencias profesionales en la red. Estudio de los casos El País Digital, El Mundo del Siglo XXI, Netestado y O Globo On. Tese de doutorado. Univ. La Laguna. 1999.

RIBEIRO, Vasco Ferreira. Jornalista-Robot: produção automática de conteúdos de texto como apoio ao jornalismo desportivo. 2019

RÜDIGER, F. As Teorias da Cibercultura: perspectivas, questões e autores. Porto Alegre: Sulina, 2011.

RÜDIGER, F. Breve história do pós-humanismo: elementos de genealogia e criticismo. Revista E-Compós, Brasília, v. 8, 2007.

_____, Francisco. Tendências do jornalismo. 3. ed. Porto Alegre: UFRGS, 2003.

_____, Francisco. As teorias da comunicação. Penso, 2010.

SHANNON, Claude E.; WEAVER, Warren. The Mathematical Theory of Communication. Illini Books Edition, University of Illinois, 1963.

SHANNON, Claude Elwood. A mathematical theory of communication. The Bell system technical journal, v. 27, n. 3, p. 379-423, 1948.

SILVA, Jonathas Luiz Carvalho. A teoria matemática da comunicação na ciência da informação: propondo uma nova relação entre sujeitos da informação. Pesquisa brasileira em Ciência da Informação e Biblioteconomia, v. 11, n. 1, p. 203-217, 2016.

SIRÉN-heikel, S.; LEPPANEN, L.; LINDÉN, C-G; BACK, A. Unboxing news automation: Exploring the imagined affordances of automation in news journalism. In: Nordic Journal of Media Studies. 1:47-66, 2020.

TIQQUN, “L'Hypothèse cybernétique,” Tiqqun 2 (2001): 40-83, 44.

TRAQUINA, Nelson. Teorias do Jornalismo: porque as notícias são como são. 3.ed.rev. Florianópolis: Insular, 2012.

VERKERK, Maarten J. et al. Filosofia da tecnologia: uma introdução. Editora Ultimato, 2021.

VICENTE, P.; FLORES, A.M. Inteligência Artificial e Jornalismo: Temas Emergentes. Universidade Nova de Lisboa iNOVA Media Lab/ICNOVA.

WEAVER, Warren. A matemática da comunicação. In: MORTENSEN, C. David. Teoria da comunicação: textos básicos. 1979.

WEAVER, Warren. Recent Contributions to The Mathematical Theory of Communication, Sep 1949.

WIENER, Norbert. Cibernética e sociedade. O uso humano de seres humanos, v. 2, 1968.

WÖLKER, A., & Powell, T. E. (2018). Algorithms in the newsroom? News readers' perceived credibility and selection of automated journalism. *Journalism*, 1464884918757072. <https://doi.org/10.1177/1464884918757072>

BARBOSA, Jonei Cerqueira. Modelagem e modelos matemáticos na educação científica. **Alexandria: Revista de Educação em Ciência e Tecnologia**, v. 2, n. 2, p. 69-85, 2009.

SODRÉ, Ulysses. Modelos matemáticos. **Londrina: UEL**, 2007.

FERNANDES, Millôr. **Poesia matemática**. Nova Fronteira, 2014.

LAGE, Nilson. **Ideologia e técnica da notícia**. Digitaliza Conteúdo, 2021.

WOLF, Mauro; DE FIGUEIREDO, Maria Jorge Vilar. **Teorias da comunicação**. Presença, 1987.

GEHLEN, Marco Antônio et al. Jornalismo de (im) precisão: o conhecimento matemático e a apuração de números. 2016.

GLEICK, James. **A informação: uma história, uma teoria, uma enxurrada**. Editora Companhia das Letras, 2013.

PENA, Felipe. A Teoria do Jornalismo no Brasil—após 1950—. **São Paulo: Contexto**, 2005.

TRAQUINA, Nelson. **Porque as notícias são como são**. Insular Livros, 2020.

GENRO FILHO, Adelmo. Sobre a necessidade de uma teoria do jornalismo. **Estudos em Jornalismo e Mídia**, v. 1, n. 1, p. 160-163, 2004. <file:///Users/manoellaff/Downloads/2233-Texto%20do%20Artigo-6247-1-10-20080613.pdf>

MARCONDES FILHO, Ciro. Comunicação e jornalismo: a saga dos cães perdidos. 2002 https://edisciplinas.usp.br/pluginfile.php/5213861/mod_resource/content/1/362420686-LIVRO-Jornalismo-a-Saga-Dos-Caes-Perdidos.pdf

MARCONDES FILHO, Ciro. **Ser jornalista: o desafio das tecnologias e o fim das ilusões**. . São Paulo: Paulus. . Acesso em: 03 maio 2023. , 2009

HOLANDA, André et al. Metodologias de pesquisa em jornalismo participativo no Brasil. **Brazilian journalism research**, v. 4, n. 2, p. 57-76, 2008

RODRIGUES, Catarina. **Jornalismo participativo: tecnologia, comunicação e o papel do jornalista**. 2013. Tese de Doutorado. Universidade da Beira Interior (Portugal).

SOUSA, Jorge Pedro. Uma história breve do jornalismo no Ocidente. **Jornalismo: história, teoria e metodologia da pesquisa**. Porto: Edições Universidade Fernando Pessoa, p. 12-93, 2008.

SONI, Jimmy; GOODMAN, Rob. **A mind at play: how Claude Shannon invented the information age**. Simon and Schuster, 2017.

Guilhermano, L. “A Crise não é Do Jornalismo, Mas Do Seu Financiamento: Entrevista Com Jean Charron”. *Intexto*, abril de 2019, p. 5-15, doi:10.19132/1807-858320190.5-15.

LINDEN, Carl-Gustav. Algoritmos para Jornalismo: o futuro da produção de notícias. *Líbero*, n. 41, p. 5-27, 2018.

BARROSO, Suélem; MARQUES, Alberto. Jornalismo Automatizado: um olhar sob a produção nos periódicos Journalism Studies, Digital Journalism e Journalism. In: **Nuevos actores y estrategias en el espacio mediático compartido: XI Congreso Internacional de Ciberperiodismo**. Servicio de Publicaciones, 2020. p. 7-23.

VICENTE, Paulo Nuno; FLORES, Ana Marta M. Inteligência Artificial e Jornalismo. **De que falamos quando dizemos Jornalismo?**, p. 175-194, 2021.

MONTAL, Tal; REICH, Zvi. I, robot. You, journalist. Who is the author? Authorship, bylines and full disclosure in automated journalism. **Digital journalism**, v. 5, n. 7, p. 829-849, 2017.

VICENTE, Paulo Nuno; FLORES, Ana Marta Moreira. INTELIGÊNCIA ARTIFICIAL E JORNALISMO: TEMAS EMERGENTES (2015-2020). **De que falamos quando dizemos jornalismo?: Temas Emergentes de Pesquisa**, p. 175, 2021.

LOZANO TELLO, Adolfo. Ontologías en la Web semántica. 2020.

MACHAADO, Elias. DOS ESTUDOS SOBRE O JORNALISMO ÀS TEORIAS DO JORNALISMO. Universidade Federal da Bahia. <https://www.e-compos.org.br/e-compos/article/view/2/4>

PENA, Felipe. A Teoria do Jornalismo no Brasil – após 19501 – <http://felipepena.com/site/artigos/A%20Teoria%20do%20Jornalismo%20no%20Brasil.pdf>

HEIDEGGER, M. Filosofia e cibernética. Pisa: ETS, 1988.

RÜDIGER, Francisco. Martin Heidegger e a questão da técnica. Porto Alegre: Sulinas, 2014, 247 p.