

UNIVERSIDADE FEDERAL DO PARANÁ

MATEUS RAMBO STREY

HUMAN-DATA IDEATION: UM *FRAMEWORK* CONCEITUAL PARA APOIAR O
ENTENDIMENTO DE PROBLEMAS EM INTERAÇÃO HUMANO-DADOS

CURITIBA PR

2019

MATEUS RAMBO STREY

HUMAN-DATA IDEATION: UM *FRAMEWORK* CONCEITUAL PARA APOIAR O
ENTENDIMENTO DE PROBLEMAS EM INTERAÇÃO HUMANO-DADOS

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

Área de concentração: *Ciência da Computação*.

Orientador: Roberto Pereira.

CURITIBA PR

2019

Catálogo na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

S914h

Strey, Mateus Rambo

Human-Data Ideation: um framework conceitual para apoiar o entendimento de problemas em interação humano-dados [recurso eletrônico] / Mateus Rambo Strey. – Curitiba, 2019.

Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Informática, 2019.

Orientador: Roberto Pereira .

1. Interação homem-máquina. 2. Framework (Programa de computador).
3. Engenharia de requisitos. I. Universidade Federal do Paraná. II. Pereira, Roberto. III. Título.

CDD: 004.019

Bibliotecário: Elias Barbosa da Silva CRB-9/1894

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em INFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **MATEUS RAMBO STREY** intitulada: **Human-Data Ideation - Um framework conceitual para apoiar o entendimento de problemas em Interação Humano-Dados**, sob orientação do Prof. Dr. ROBERTO PEREIRA, que após após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre 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, 29 de Agosto de 2019.



ROBERTO PEREIRA

Presidente da Banca Examinadora (UNIVERSIDADE FEDERAL DO PARANÁ)



MARCOS DIDONET DEL FABRO

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)



LUCIANA CARDOSO DE CASTRO SALGADO

Avaliador Externo (UNIVERSIDADE FEDERAL FLUMINENSE)



*A Robertha, pelo companheirismo e
dedicação com os quais fui agraci-
ado.*

AGRADECIMENTOS

A Roberto, primeiramente, pelo papel exemplar e inspirador que sempre desempenhou, bem como pela paciência. A colegas do grupo de pesquisa, que ajudaram e ajudam com questões acadêmicas e além. A Diego Pasqualin e ao projeto SIMTransparência, fundamentais na execução do meu trabalho de mestrado. Todos os professores do C3SL e DInf, especialmente Castilho, Didonet e Todt, que estiveram presentes em momentos importantes da minha vida acadêmica e profissional. A professoras Laura e Natasha, pela inspiração e brilhantismo de seus trabalhos na comunidade de IHC. A professora Luciana Salgado, sempre muito atenciosa.

Agradeço também imensamente a participantes dos casos de estudo que cederam seu tempo para contribuir com a minha pesquisa. Muito grato a comunidade da IHC do Brasil, que me proporcionou experiências muito significativas. Agradeço a UFPR como um todo, minha *alma mater*, e a CAPES por financiar minha bolsa de estudo e responsável por manter com muita luta a produção científica nacional.

A Zenaide, Leo e Gastão, pelo papel fundamental em minha vida. A Daniel, Nanci e Reni, minha família que eu escolheria mesmo que não fosse. A Rafael, Iago e Cleyton, por estarem comigo desde sempre, de uma forma ou de outra. A Maurício, irmão de longa data. A Valcy, pelo suporte nos momentos difíceis. A Carlos, Gisele e Miriã, sempre presentes e dispostos a ajudar.

RESUMO

A produção, processamento e armazenamento de grandes volumes de dados têm sido uma realidade diária e o surgimento de áreas como *Big Data*, *Internet of Things* e *e-Science* corroboram com a necessidade de interagir com dados em diversos contextos. As influências e consequências dessa interação são estudadas por pesquisadores que atuam na área de Interação Humano-Dados (IHD), preocupados com o fazer sentido, agência e até as consequências técnicas e sociais atreladas a produção e coleta, processamento, visualização e manutenção dos dados, como o direito a privacidade e a propriedade. Considerando a presença cotidiana de contextos compatíveis com a IHD e a diversidade de facetas que a área engloba, se faz necessário que projetistas (analistas, desenvolvedores, engenheiros, arquitetos) de *software* entendam a problemática da interação com dados de maneira sistêmica e integrada ao conceber uma solução computacional que apoie a IHD. Em um Mapeamento Sistemático de Literatura, foram identificados elementos conceituais que merecem atenção em contextos de IHD (e.g. transparência, legibilidade, privacidade, etc.). Entretanto, nenhum método que pudesse apoiar projetistas de *software* para considerar esses elementos desde os estágios iniciais do desenvolvimento de *software* foi encontrado. Nesta pesquisa, foi concebido um *framework* conceitual – chamado *Human-Data Ideation* (HDIdea) – que busca apoiar os envolvidos com o projeto de soluções de *software* em IHD a identificar, entender e especificar elementos que são críticos para o contexto de uso das soluções, de uma perspectiva de diferentes partes interessadas em diferentes estágios do Ciclo de Vida dos Dados. O *framework* apoia os projetistas na identificação de partes interessadas do projeto, suas motivações e ações relacionadas ao elementos conceituais identificados na IHD, para então analisar os problemas em múltiplos níveis de abstração por meio de um artefato adaptado da Semiótica Organizacional. Quatro estudos empíricos foram conduzidos aplicando o *framework* em diferentes situações: os primeiros dois para a sua concepção e os últimos dois para seu uso e análise do uso. O primeiro e quarto experimentos foram conduzidos com estudantes de pós-graduação na área de Interação Humano-Computador (IHC), e o segundo e terceiro experimentos foram conduzidos com graduandos em Ciência da Computação durante uma disciplina de Engenharia de Requisitos. Os graduandos foram escolhidos por representarem o público-alvo do *framework* e os pós-graduandos devido ao seu conhecimento abrangente da área de IHC, relevante para a IHD. Os resultados sugerem que o *framework* é promissor ao apoiar os participantes em atividades de entendimento de problema, aumentando a preocupação com os elementos importantes para a IHD e auxiliando o processo de desenvolvimento de *software* contextualizado na área, aproximando conceitos e processos das áreas de IHC e Engenharia de *Software*. Como principais contribuições da pesquisa: (I) um mapeamento da literatura de IHD; (II) um *framework* conceitual que apoia projetistas de *software* no entendimento de problemas em contextos de IHD; e (III) uma ferramenta que apoia o processo de Mapeamento ou Revisão Sistemática de Literatura e que ajuda a difundir a prática.

Palavras-chave: Interação Humano-Dados. *Framework* Conceitual. Entendimento de problema. Interação Humano-Computador. Requisitos de *Software*.

ABSTRACT

The production, processing and storage of large volumes of data has been a daily reality and the emergence of areas such as Big Text, Internet of Things and e-Science enhance the need to interact with data in different contexts. The influences and consequences of this interaction are studied by researchers working in the area of Human-Data Interaction (HDI), concerned with sensemaking, agency and even the technical and social consequences linked to data production and collection, processing, visualization and disposal, such as rights to privacy and data property. Considering the daily presence of HDI-compliant contexts and the diversity of facets it encompasses, it is necessary that software designers (analysts, developers, engineers, architects) understand the problem of interacting with data in a systemic and integrated manner when creating a computing solution that supports HDI. In a Systematic Mapping of Literature, conceptual elements that deserve attention in HDI contexts were identified (e.g. transparency, legibility, privacy, etc.). However, no method that could support software designers to consider these elements from the early stages of software development was found. In this research, a conceptual framework named Human-Data Ideation was created to support the ones involved with the design of HDI software solutions to identify, understand and specify elements that are critical for the solution's usage context, from the perspective of different stakeholders in different stages of Data Lifecycle. The framework supports designers to identify project stakeholders, their motivations and actions related to the conceptual elements identified in the HDI, to then analyze their problems in multiple levels of abstraction through an artifact adapted from Organizational Semiotics. Four practical studies were conducted applying the framework in different situations: the first two for its conception and the last two for its use and analysis of the use. The first and fourth experiments were carried out with post-graduate students in the area of Human-Computer Interaction (HCI) and the second and third experiments were carried out with graduates in Computer Science during a discipline of Requirements Engineering. The undergraduates were chosen because they represented the target audience of the framework and post-graduate students for their comprehensive knowledge of the HCI area, relevant for HDI. Results suggest that the framework is promising to support the participants in problem understanding activities, raising the concern with important elements for HDI and assisting the process of developing software, bringing closer concepts and processes in the areas of HCI and Software Engineering. As the main contributions from this research: (I) a mapping of the HDI literature; (II) a conceptual framework to support software designers in problem understanding for HDI contexts; and (III) a tool that supports the systematic literature review process and helps spread the practice.

Keywords: Human-Data Interaction. Conceptual Framework. Problem understanding. Human-Computer Interaction. Software requirements.

LISTA DE FIGURAS

2.1	Escada Semiótica, retirada de Liu (2000) e traduzida pelo autor do trabalho. . . .	21
2.2	Diagrama de Partes Interessadas, adaptado de Pereira e Baranauskas (2015).. . .	22
2.3	Interface de seleção de trabalhos na MaRS.	24
2.4	Rede de trabalhos selecionados (em verde) e suas referências (não selecionados em azul e descartados em vermelho) na MaRS.	24
3.1	Esquema geral do HDIdea.	27
3.2	Passo 1 - Diagrama de Partes Interessadas estendido com Tipos de Partes Interessadas.. . . .	29
4.1	Diagrama de Partes Interessadas da MaRS. Os partes interessadas que começam com asterisco (*) foram identificados no piloto do <i>framework</i> conceitual com a ferramenta MaRS.. . . .	36
4.2	Alunos de Engenharia de Requisitos colando suas Histórias de Usuário para o projeto final da disciplina.	39
4.3	Alunos de Engenharia de Requisitos colando suas histórias de usuário do projeto final da disciplina.. . . .	40
4.4	Exemplo pré-preenchido da Escada Semiótica em uma história de usuário genérica para o contexto da ferramenta SIMTransparência.	41
4.5	Diagrama de Partes Interessadas estendido com os Tipos de Partes Interessadas preenchido no dia 17/08/2018.	42

LISTA DE TABELAS

3.1	Valores relacionados ao tema <i>Legibilidade</i> e perguntas expostas no HDIdea.	30
3.2	Valores relacionados ao tema <i>Agência</i> e perguntas expostas no HDIdea.	31
3.3	Valores relacionados ao tema <i>Negociabilidade</i> e perguntas expostas no HDIdea.	32
4.1	Casos de estudo realizados para a fomentar a criação do Human-Data Ideation.	35
4.2	Histórias de usuário identificadas no piloto do <i>framework</i> conceitual com a ferramenta MaRS. Em itálico o que foi pré-preenchido. Em negrito o que foi implementado no protótipo.	38
4.3	Primeira parte das histórias de usuário identificadas durante aplicação do segundo passo do HDIdea com a turma de Engenharia de Requisitos do 2º semestre de 2018.43	43
4.4	Segunda parte das histórias de usuário identificadas durante aplicação do segundo passo do HDIdea com a turma de Engenharia de Requisitos do 2º semestre de 2018.44	44
4.5	Elementos Conceituais úteis no contexto de IHD e SIMTransparência para os participantes do questionário da turma de Engenharia de Requisitos do 2º semestre de 2018.	46
4.6	Elementos Conceituais úteis no contexto de <i>redesign</i> da ferramenta MaRS para os colegas de pós-graduação do autor que participaram da atividade.	50
5.1	Passos do HDIdea relacionados com os Temas Centrais do <i>Design Thinking</i> de Carlgren et al. (2016).	54
C.1	Primeira parte do questionário aplicado com 3 alunos da turma de Engenharia de Requisitos do DINF no 2º semestre de 2018.	89
C.2	Segunda parte do questionário aplicado com 3 alunos da turma de Engenharia de Requisitos do DINF no 2º semestre de 2018.	90
C.3	Terceira parte do questionário aplicado com 3 alunos da turma de Engenharia de Requisitos do DINF no 2º semestre de 2018.	91
C.4	Quarta parte do questionário aplicado com 3 alunos da turma de Engenharia de Requisitos do DINF no 2º semestre de 2018.	92
D.1	Primeira parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.	94
D.2	Segunda parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.	95
D.3	Terceira parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.	96
D.4	Quarta parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.	97
D.5	Quinta parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.	98

D.6	Sexta parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.	99
D.7	Sétima parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.	100
D.8	Oitava parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.	101
D.9	Questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.. .	102
F.1	Histórias de Usuário criadas com colegas de pós-graduação do autor em 14/12/2018.	109
F.2	Primeiro artefato preenchido do Ciclo de Vida dos Dados estendido com a Escada Semiótica, criado com colegas de pós-graduação do autor em 14/12/2018.	110
F.3	Segundo artefato preenchido do Ciclo de Vida dos Dados estendido com a Escada Semiótica, criado com colegas de pós-graduação do autor em 14/12/2018.	111
F.4	Terceiro artefato preenchido do Ciclo de Vida dos Dados estendido com a Escada Semiótica, criado com colegas de pós-graduação do autor em 14/12/2018.	112
F.5	Quarto artefato preenchido do Ciclo de Vida dos Dados estendido com a Escada Semiótica, criado com colegas de pós-graduação do autor em 14/12/2018.	113
F.6	Quinto artefato preenchido do Ciclo de Vida dos Dados estendido com a Escada Semiótica, criado com colegas de pós-graduação do autor em 14/12/2018.	114

LISTA DE ACRÔNIMOS

DINF	Departamento de Informática
PPGINF	Programa de Pós-Graduação em Informática
UFPR	Universidade Federal do Paraná
IHC	Interação Humano-Computador
IHD	Interação Humano-Dados
MaRS	Mapping Review System - Sistema de Mapeamento Sistemático de Literatura
IoT	<i>Internet of Things</i>
HDI	Human-Data Interaction
HDIdea	Human-Data Ideation
DPI	Diagrama de Partes Interessadas
TPI	Tipos de Partes Interessadas
CVD	Ciclo de Vida dos Dados
ES	Escada Semiótica
ONG	Organização Não Governamental
CGU	Controladoria Geral da União
TCU	Tribunal de Contas da União
MEC	Ministério da Educação

SUMÁRIO

1	INTRODUÇÃO	13
2	FUNDAMENTAÇÃO TEÓRICA E TÉCNICA	16
2.1	MAPEAMENTO SISTEMÁTICO DE LITERATURA SOBRE IHD	16
2.2	SEMIÓTICA ORGANIZACIONAL E O CICLO DE VIDA DOS DADOS	20
2.3	MARS - <i>MAPPING REVIEW SYSTEM</i>	23
3	<i>HUMAN-DATA IDEATION</i>	26
3.1	PASSO 1 - IDENTIFICAÇÃO E CLASSIFICAÇÃO DAS PARTES INTERESSADAS	28
3.2	PASSO 2 - IDENTIFICAÇÃO DAS AÇÕES DAS PARTES INTERESSADAS SOB OS ELEMENTOS CONCEITUAIS DE IHD.	29
3.3	PASSO 3 - REFINAMENTO DAS HISTÓRIAS DE USUÁRIO EM DIFERENTES NÍVEIS DE ABSTRAÇÃO	32
3.4	CONSIDERAÇÕES SOBRE O <i>HUMAN-DATA IDEATION</i>	33
4	CASOS DE ESTUDO	35
4.1	CASO DE ESTUDO DE CONCEPÇÃO DO <i>HUMAN DATA IDEATION</i> APLICADO NA FERRAMENTA MARS	35
4.2	CASO DE ESTUDO DE CONCEPÇÃO DO HDIDEA APLICADO COM UMA TURMA DE ENGENHARIA DE REQUISITOS	37
4.3	CASO DE ESTUDO DE USO E ANÁLISE DO HDIDEA APLICADO COM UMA TURMA DE ENGENHARIA DE REQUISITOS	40
4.4	CASO DE ESTUDO DE USO E ANÁLISE DO HDIDEA APLICADO EM UM CONTEXTO DE <i>REDESIGN</i> DA FERRAMENTA MARS COM COLEGAS DE PÓS-GRADUAÇÃO DO AUTOR	47
4.5	CONSIDERAÇÕES SOBRE OS CASOS DE ESTUDO.	51
5	DISCUSSÃO DOS RESULTADOS.	52
6	CONCLUSÕES	55
	REFERÊNCIAS	56
	APÊNDICE A – MAPEAMENTO SISTEMÁTICO DE LITERATURA SOBRE O ESTADO DA ARTE DA INTERAÇÃO HUMANO-DADOS.	59
	APÊNDICE B – MANUAL DE USO - HUMAN-DATA IDEATION	72
	APÊNDICE C – QUESTIONÁRIO APLICADO COM A TURMA DE ENGENHARIA DE REQUISITOS DO DINF NO 2º SEMESTRE DE 2018	88
	APÊNDICE D – QUESTIONÁRIO APLICADO COM COLEGAS DE PÓS-GRADUAÇÃO DO AUTOR EM 14/12/2018.	93
	APÊNDICE E – ARTEFATOS PREENCHIDOS COM A TURMA DE ENGENHARIA DE REQUISITOS DO DINF NO 2º SEMESTRE DE 2018	103

APÊNDICE F – ARTEFATOS PREENCHIDOS COM COLEGAS DE PÓS-GRADUAÇÃO DO AUTOR EM 14/12/2018.	108
--	------------

1 INTRODUÇÃO

Em um mundo permeado por tecnologias computacionais no qual a sociedade está cada vez mais conectada por dispositivos eletrônicos e produzindo dados sobre toda e qualquer atividade possível, p. ex. a rota que turistas percorrem em uma região ou indícios de aprendizagem em visitantes de uma exposição de museu, há uma demanda crescente por pesquisas e soluções que apoiem e possibilitem a coleta, processamento, uso e manutenção desses dados, considerando diferentes partes interessadas em diferentes contextos, atentando para leis e o direito à privacidade. A área de Interação Humano-Dados (IHD) surge na literatura com Elmqvist (2011) como uma iniciativa de investigar a problemática da interação de humanos com dados em sistemas computacionais, com o foco no entendimento - fazer sentido - dos dados e informações em larga escala. Haddadi et al. (2013) e Mortier et al. (2014) expandem o escopo da IHD, englobando não só a visualização e o fazer sentido do grande volume de dados, mas também as consequências dessa produção, como o consentimento e a privacidade das partes envolvidas ou o modo como as informações são inferidas por algoritmos a partir desses dados.

O objetivo geral da pesquisa é propor uma solução (*framework* conceitual) que considere os diferentes elementos conceituais de IHD na concepção, bem como outras fases, de projetos de *software* contextualizados na área, favorecendo o entendimento do problema e a identificação de requisitos funcionais e não-funcionais. Os objetivos específicos são: (I) Conduzir um Mapeamento Sistemático de Literatura sobre o estado da arte em IHD; (II) Criar o *framework* conceitual, chamado *Human-Data Ideation* (HDIdea), com base na literatura da área; (III) Experimentar o *framework* conceitual em casos de estudo.

Se levarmos em consideração que muitos sistemas existentes e futuros estão alinhados no contexto da IHD, deve-se considerar como o processo de desenvolvimento de sistemas, i.e. *software*, é executado de maneira a abranger as necessidades identificadas na área. O que percebe-se, após a realização do Mapeamento Sistemático de Literatura, é a falta de soluções que reúnam os múltiplos elementos conceituais de IHD e apoiem o entendimento de problema, de modo sistêmico e teoricamente informado, de projetistas (analistas, arquitetos, desenvolvedores, engenheiros) de *software* situados nesse contexto, em especial os leigos.

Segundo Oulasvirta e Hornbæk (2016), a pesquisa em IHC pode ser desenvolvida a partir de uma perspectiva de resolução de problemas que possui 3 tipos de problemas: (a) Empírico, no qual criam-se descrições elaboradas de fenômenos do mundo real relacionados ao uso humano da computação; (b) Conceitual, no qual explicam-se fenômenos previamente não relacionados que ocorrem nas interações, e (c) Construtivo, no qual produz-se entendimento sobre a construção de um artefato interativo para algum propósito no uso humano da computação. Os autores consideram que a literatura de IHC apresenta pares combinados de tipos de problema, e.g. empírico-construtivo ou conceitual-empírico.

Apoiar projetistas de *software* com o entendimento de problema em contextos de IHD visando a identificação de requisitos pode ser classificado como *conceitual-construtivo*, tendo em vista o exemplo dado pelos autores neste caso como "novas maneiras de projetar fenômenos previstos teoricamente na interação". O mapeamento de literatura executado identificou que os autores dos trabalhos selecionados começaram a identificar elementos conceituais da IHD e o presente trabalho pretende auxiliar projetistas de *software* a identificar requisitos de diferentes níveis de abstração em quaisquer contextos de IHD, relacionando os elementos conceituais ao entendimento do problema. Para tal, adotou-se uma perspectiva sociotécnica para auxiliar

o reconhecimento dos diversos elementos conceituais relacionados tanto ao contexto técnico quanto ao contexto humano, fazendo uso de artefatos que reconheçam esses dois contextos.

Com a falta de um trabalho que agregue os principais trabalhos e vertentes da área, se fez necessário realizar um Mapeamento Sistemático de Literatura. Com o mapeamento, identificou-se que Elmqvist (2011) inicia e promove as discussões em IHD. Entretanto, embora o termo tenha se popularizado e uma linha de pesquisa tenha se delineado sobre o tópico, ainda não existe um consenso sobre uma definição para IHD. A definição mais abrangente e popular vem de Haddadi et al. (2013)¹: a IHD estuda a interação entre humanos e a análise de grandes conjuntos de dados pessoais, combinando tanto os dados quanto os algoritmos usados para analisá-los. Neste trabalho não limitou-se a IHD aos dados pessoais (dados que são produzidos por pessoas ou que as identifica) e sim aos dados que afetam pessoas, e nem ao volume dos dados, desde que uma solução da área amplie a capacidade humana de interagir com os dados. Foram identificados ainda alguns *frameworks*, boas práticas e recomendações de design, mas apenas Hornung et al. (2015) está alinhado com a problemática que pretendemos abordar - contando com caráter exploratório, sem avaliação - sendo adaptado para compor o *framework* conceitual almejado.

Como não existe uma forma de avaliar o *framework* conceitual por comparação, visto que não foram encontradas propostas similares no mapeamento de literatura, optou-se pela realização de casos de estudo em que os participantes fizeram uso do HDIdea e o avaliaram, e.g. facilidade no uso, utilidade, tendências no uso. Os casos com o público-alvo, projetistas de *software*, foram realizados na ferramenta SIMTransparência², que auxilia na transparência de gastos públicos em educação, onde duas turmas de uma disciplina de Engenharia de Requisitos (de *software*) fizeram uso do *framework* conceitual. Outros dois casos de estudo envolveram pós-graduandos com a aplicação do *framework conceitual* em atividades de ideação e *redesign* de ferramenta de apoio ao Mapeamento Sistemático de Literatura (chamada MaRS), que está inserida no contexto de IHD. O resultado das atividades serviu para o aprimoramento gradual do HDIdea até o estágio atual.

Como resultados, existe uma versão aplicável do HDIdea compilada em um manual para os projetistas (Apêndice B). O mapeamento foi publicado como artigo completo no XVII Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (IHC 2018 - Strey et al. (2018)). Casos de estudo do HDIdea produziram resultados para análise e aprimoramento do mesmo e foram conduzidos em (a) dois momentos na forma de *workshop* de ideação e *redesign* com pós-graduandos de IHC que exploraram a ferramenta MaRS e (b) em duas turmas da disciplina de Engenharia de Requisitos na forma de atividade intraclasse com graduandos de Ciência da Computação, que interagiram com a ferramenta SIMTransparência. Também existe um protótipo funcional da ferramenta MaRS, utilizada na execução do Mapeamento Sistemático de Literatura em IHD.

Finalmente, espera-se que o HDIdea e a ferramenta MaRS tragam contribuições metodológicas e sociais, visto que as duas tem potencial de auxiliar tanto projetistas quanto partes interessadas que se encontrem em contextos de IHD, além aproximar questões de IHC em um processo da Engenharia de Software. A MaRS ainda contribui metodologicamente para os Mapeamentos Sistemáticos de Literatura ao promover a transparência do processo realizado, sendo a transparência um elemento conceitual identificado em IHD. Também espera-se que

¹Haddadi et al. (2013) define a IHD como: "The concern with HDI [Human-Data Interaction] is not interaction between humans and computers generally, but between humans and the analysis of large, rich personal datasets. (...) HDI includes the combination of both data, and the algorithms used to analyse it."

²SIMTransparência: <https://www.c3sl.ufpr.br/transparencia/>

o Mapeamento Sistemático de Literatura realizado promova a disseminação da área e ajude a sintetizar as múltiplas perspectivas abordadas em IHD.

2 FUNDAMENTAÇÃO TEÓRICA E TÉCNICA

Os estudos em IHD surgem em 2011, com Elmqvist (2011), que define o escopo da área como sendo a "manipulação humana, análise e *sensemaking* de dados". O autor ainda cita que os dados trabalhados seriam caracterizados pela larga escala, não estruturação, fontes múltiplas e relações complexas. Com o passar dos anos, outros autores começaram a publicar trabalhos na área sob diferentes perspectivas, as quais chamamos de visões em Strey et al. (2018), classificando-as em quatro categorias: (a) *Embodied*, (b) Ampla, (c) Colaborativa e (d) Visualização.

A visão *Embodied*, ou corporificada, publicada primeiramente por Elmqvist (2011), é a primeira que traz a conceituação da área, conforme apontado no parágrafo anterior. A visão Ampla, que surge com Haddadi et al. (2013), é conceituada como "a interação entre humanos e a análise de grandes conjuntos de dados, focando tanto nos dados quanto nos algoritmos usados para analisá-los". A visão Colaborativa, definida por Kee et al. (2012), considera a IHD, juntamente com a comunicação humano-humano, como envolvidas na Ciência da Interação (estudo relacionado ao desenvolvimento de ferramentas colaborativas). Por fim, na visão Visualização, defendida por Widjojo et al. (2017), IHD seria a interface que liga de forma direta o ser humano a uma representação visual dos dados. Embora tratem de diversas facetas da IHD, nenhuma dessas visões parece ser capaz de contemplar a abrangência total da área e, por conta dessa necessidade, o autor (com base nos estudos realizados através do Mapeamento Sistemático de Literatura e experimentos no decorrer do mestrado), define IHD, para fins desse trabalho, como a área que estuda "a manipulação, análise e *sensemaking* de dados por humanos, sendo que o volume dos dados não é relevante se a interação proporcionada amplie a capacidade humana de realizar tais ações".

Na tentativa de encontrar propostas para apoiar o entendimento do problema por projetistas de *software* em projetos contextualizados em IHD, encontrou-se apenas uma proposta relacionada aos fins dessa dissertação, descrita em Hornung et al. (2015). O trabalho utilizou a Semiótica Organizacional, por meio do artefato Escada Semiótica proposto por Stamper (1993) – que oferece uma perspectiva sociotécnica na abstração do problema – e o que os autores chamaram de Ciclo de Vida dos Dados (*data lifecycle*) – que oferece uma segregação das etapas em que as partes interessadas interagem com os dados. Porém, ao se fazer o uso destes artefatos, é necessário que o projetista já tenha identificado quem são as partes interessadas no projeto e também quais as ações e necessidades destas partes interessadas neste contexto. Assim, o trabalho propõe a utilização de outros artefatos que forneçam esses insumos para a aplicação dos artefatos propostos por Hornung et al. (2015).

2.1 MAPEAMENTO SISTEMÁTICO DE LITERATURA SOBRE IHD

Segundo Petersen et al. (2015), o Mapeamento Sistemático de Literatura (MSL) é um método utilizado para dar uma visão geral de uma área de pesquisa, através da classificação e enumeração das contribuições para a área (em relação às categorias dessa classificação). Diferentemente da Revisão Sistemática de Literatura - que se preocupa em sintetizar evidências - o MSL preocupa-se em estruturar uma área de pesquisa. Considerando que IHD é uma área nova, que não haviam muitas publicações sobre o tema e que não existia nenhuma publicação que sintetizasse a área, optou-se, portanto, em utilizar o Mapeamento Sistemático de Literatura como método para identificar e organizar a literatura no tema (Strey et al. (2018)).

A *string* de busca definida para o mapeamento foi "*Human-Data Interaction*" e "*Human Data Interaction*", visando encontrar qualquer trabalho que citasse explicitamente o termo, e as bases selecionadas para a pesquisa de trabalhos foram: *ACM Digital Library*, *IEEE Xplore*, *Springer Link*, e *Science Direct*. Foram encontrados 57 artigos e selecionados 20 artigos após a aplicação do 1º filtro, aplicando os critérios de inclusão e exclusão considerando o título, *abstract*, palavras-chave e o texto completo do trabalho. O único critério de inclusão dos trabalhos era se os trabalhos consideravam explicitamente o termo e o contexto de IHD, enquanto os critérios de exclusão dos trabalhos estão elencados abaixo:

1. Trabalhos não disponíveis para acesso (e.g. não acessíveis na rede interna da UFPR);
2. Trabalhos que não estejam em Inglês;
3. Trabalhos que são duplicatas de trabalhos já selecionados (apresentavam/discutiam um mesmo projeto ou estratégia), mantendo o trabalho completo ou mais recente;
4. Tabela de conteúdo de simpósio/conferência/etc;
5. Trabalhos não relacionados ao contexto de IHD;
6. Trabalhos que não apresentassem solução prática ou conceitual para o contexto de IHD, ou que referenciavam IHD sem abordá-la de fato (e.g. alguns trabalhos em Big Data colocam IHD como uma preocupação de *Data Quality*, sem explorar a problemática).

Com a leitura dos artigos selecionados, percebeu-se que parte da literatura da área não estava em publicações reconhecidas e indexadas por bases de dados, mas nas páginas institucionais ou pessoais dos autores. Foi utilizada a técnica *snowballing*, que consiste em procurar trabalhos relevantes a pesquisa nas referências dos trabalhos já selecionados, para incluir os trabalhos não indexados. Ao aplicá-la foram encontrados mais 10 artigos e, após a aplicação dos critérios de inclusão e exclusão, 8 artigos foram selecionados. O mapeamento, então, é composto por 28 trabalhos.

Para o mapeamento, foram definidas as oito perguntas de pesquisa listadas abaixo. As mais relevantes para esta dissertação estão destacadas:

- *PP1: O que os trabalhos mapeados consideram como IHD?*
- *PP2: Quais são os países, instituições e autores que publicam sobre IHD?*
- *PP3: Em quais fóruns científicos esses trabalhos têm sido publicados?*
- *PP4: Os trabalhos apresentam alguma proposta?*
- *PP5: Quem são as partes interessadas alvo da proposta?*
- *PP6: Como as propostas são avaliadas?*
- *PP7: Quais bases teóricas fundamentam as propostas mapeadas?*
- *PP8: Quais elementos conceituais básicos para IHD foram apresentados nos trabalhos?*

Ao responder **o que os trabalhos mapeados consideram como IHD (PP1)**, classificou-se 4 formas de entender a área (chamadas de visões): (a) *Embodied*, (b) *Ampla*, (c) *Colaborativa* e (d) *Visualização*. A visão *Ampla* surge com Haddadi et al. (2013) como "a interação entre humanos e a análise de grandes conjuntos de dados, focando tanto nos dados quanto nos

algoritmos usados para analisá-los". A IHD difere da IHC por utilizar cenários mais passivos e com grandes volumes de dados, onde os usuários envolvidos não entendem ou não se importam com o funcionamento da infraestrutura que estão utilizando. Mortier et al. (2014), os mesmos autores do trabalho anterior, propõem três temas centrais da IHD: *Legibilidade* – tornar os dados e algoritmos analíticos transparentes e compreensíveis para as pessoas; *Agência* – pessoas devem ter a capacidade de agir dentro dos sistemas de dados, controlando, informando e corrigindo dados e inferências sobre si e outros, e *Negociabilidade* – as entidades envolvidas nos sistemas precisam ter a capacidade de negociar os usos dos dados, visto que as necessidades, e mesmo a cultura, mudam com o passar do tempo. Outros autores estendem a definição inicial da visão ampla, mas iremos discutir adiante apenas o que for relevante para o presente trabalho. As outras duas visões (*Colaborativa* e *Visualização*) possuem apenas 1 trabalho cada uma e não foram incluídas, mas podem ser verificadas no artigo do mapeamento publicado no IHC 2018, incluído no Apêndice A.

Dos 28 trabalhos mapeados, 20 adotam parcial ou integralmente a visão Ampla e neste trabalho decidiu-se adotá-la também. A visão ampla consegue incluir as outras visões dentro de sua área de atuação (e.g. fazer sentido de um grande volume de dados por diferentes formas de interação) e ainda contempla uma preocupação com o mundo social, com questões sobre privacidade, ética e consentimento na interação humana com os dados.

Limitando os **trabalhos que apresentam alguma proposta (PP4)** aos que possam apoiar o entendimento do problema dos projetistas, foram identificados 9 trabalhos que apresentam alguma boa prática, recomendação de *design* ou um *framework*. Com a visão *Embodied*, mapeou-se 2 trabalhos: Mishra e Cafaro (2018) identificam desafios que os usuários enfrentam ao usar um sistema com instruções mínimas, fornecendo recomendações de *design* para a criação de instalações interativas e envolventes relacionadas a exploração de dados, mesmo sem a intervenção e apoio de intérpretes. Pacheco et al. (2015) propõem um *framework* que aplica Realidade Aumentada para explorar e visualizar dados históricos dentro de um memorial ou museu.

Com a visão Ampla, mapeou-se 7 trabalhos: Mashhadi et al. (2014) propõem um conjunto de diretrizes que os provedores de dispositivo e serviços *Internet of Things* (IoT) podem considerar ao projetar seu modelo de negócios, bem como modelos de propriedade de dados de granularidade variável (*Pay-per-use*, *Data Market*, *Open Data*). Hutton e Henderson (2017) sugerem boas práticas para que coletores de dados garantam que suas atividades de mineração de dados não violem as expectativas (e consentimento) das pessoas a quem os dados se relacionam, com base nos 3 temas centrais da IHD. Hornung et al. (2015) estendem a Escada Semiótica com um Ciclo de Vida de dados (e.g. origem de dados, seleção, limpeza, mapeamento, exibição ou interação) e as partes interessadas envolvidas, além de apontar desafios na área. Cavoukian e Chibba (2016) propõem o uso do *framework Privacy By Design* para lidar com privacidade no contexto de IHD e cidades cognitivas. Wilke e Portmann (2016) propõem a Computação Granular como base teórica, formal e metodológica para a IHD. Leone (2017) identifica os desafios éticos decorrentes do cenário IoT no âmbito da agroalimentação e propõe uma "abordagem ética no *design*" (*ethical in-design approach*) para a *Internet-of-Food*. Koesten et al. (2017) entrevistam pessoas que trabalham com análise de dados e propõem um *framework* para a Interação Humano-"Dados estruturados".

Especificamente no sentido de capacitar projetistas ou outras partes interessadas na criação de soluções de IHD, existem 7 propostas. Mashhadi et al. (2014) focam em modelagem de negócios para IoT/IHD, o que foge do público-alvo do presente trabalho que é o projetista de *software*. Hutton e Henderson (2017) apresentam sugestões para projetar sistemas de IHD ao observar os 3 temas centrais (Legibilidade, Agência, e Negociabilidade). Apesar de ser

algo a se considerar em um *framework* conceitual da área, não é uma solução completa para o entendimento do problema em projetos de *software* – pois foca apenas no consentimento das partes envolvidas em atividades de mineração de dados. Cavoukian e Chibba (2016) sugerem o *framework* holístico *Privacy By Design* para auxiliar o projetista (e demais partes interessadas) a lidar com a privacidade dos dados em contextos de IHD. Também é algo a se considerar, mas o entendimento de problema não se restringe à privacidade e o *framework* conceitual proposto busca considerar os diferentes aspectos identificados na área de forma articulada, não fragmentada por meio de diversas soluções específicas. Wilke e Portmann (2016) propõem o uso do *general granular framework* em nível teórico, o que foge do escopo de conceber um *framework* conceitual para projetistas de *software*. Leone (2017) propõe um *framework* para lidar com a ética em cenários de IoT e *Internet-of-Food*, fugindo também do objetivo de ter uma solução articulada. Koesten et al. (2017) apresentam um *framework* conceitual para projetar sistemas de Interação Humano-"Dados estruturados", já que suas partes interessadas trabalham com dados (e.g. procuram dados em plataformas espaciais ou estatísticas), o que inviabiliza a proposta de apoiar os projetistas independente do domínio de aplicação. Hornung et al. (2015) propõe a extensão do artefato Escada Semiótica com o Ciclo de Vida dos Dados e as partes interessadas envolvidas para identificar questões de *design*.

Os trabalhos encontrados que traziam *frameworks* conceituais tratavam de diversos tipos de partes interessadas envolvidas em diferentes contextos, logo, ficou entendido que um *framework* conceitual que apoie o entendimento de problemas em IHD precisa: **(I) identificar quem são as partes interessadas e (II) o que as partes interessadas fazem em contextos de IHD.** (e.g. Chamberlain e Crabtree (2016) estudam o ato de escutar música em dispositivos eletrônicos com pessoas de diferentes perfis e seu relacionamento com metadados, identificando o *workflow* relacionado a atividade). Ainda está demonstrado que existem vários aspectos a serem considerados ao projetar soluções de IHD, do nível técnico ao social, como: exploração, visualização e entendimento dos dados; propriedade dos dados; consentimento e privacidade das partes interessadas; ciclo de vida dos dados; ou responsabilidade ética ao projetar sistemas de IHD. Desta maneira, um *framework* conceitual para a área também precisa **(III) auxiliar na identificação e entendimento de problemas em diferentes níveis de abstração**, apoiado pelos aspectos mais relevantes dos problemas identificados na literatura. O trabalho de Hornung et al. (2015) serviu de base para a elaboração do HDIdea, pelo seu potencial de análise em amplitude e por permitir entender o problema sob os vários níveis de abstração.

Ainda, em relação a **quais elementos conceituais básicos para IHD foram apresentados nos trabalhos (PP8)**, o autor identificou 25 elementos recorrentes na literatura de IHD, sendo eles: Acessibilidade; Adaptabilidade; Agência; Autonomia; *Awareness*; Colaboração; Compartilhamento; Confiança; Consentimento; Controle; Conversação; Grupos; Identidade; Legibilidade; Negociabilidade; Normas; Portabilidade; Presença; Privacidade; Propriedade; Relacionamentos; Segurança; Transparência; Usabilidade; e Visibilidade. Com exceção de Agência, Legibilidade e Negociabilidade (que são os temas centrais da IHD) e Transparência, todos os outros aspectos são valores já mapeados por Pereira et al. (2013) no contexto de software social.

Por meio do mapeamento sistemático realizado, percebe-se a falta de um consenso sobre a área de IHD, dada sua história recente e a interdisciplinaridade que não favorece um entendimento homogêneo do tema. Os autores identificaram diversos elementos conceituais que podem ser pertinentes a IHD, mas não existem trabalhos que agreguem os múltiplos elementos conceituais de uma forma intuitiva para quem está projetando soluções contextualizadas na área, apesar de Hornung et al. (2015) proporem o uso de um artefato que possibilita o entendimento

de problemas por múltiplas perspectivas em conjunto com artefato que diferencia diferentes momentos que os dados são utilizados no sistema pelas partes interessadas.

Após a realização do MSL, o autor entrou em contato com o trabalho de Barreto et al. (2018), em que os autores identificam e avaliam estratégias de comunicação que viabilizam a IHD da visão Ampla. Por meio do Método de Inspeção Semiótica, foram avaliadas duas ferramentas em relação ao temas centrais de Mortier et al. (2014), e os resultados do trabalho contemplaram a viabilidade do Método de Inspeção Semiótica para avaliar a comunicabilidade dos elementos conceituais da IHD abordados, além de identificar estratégias comunicativas para soluções de IHD.

2.2 SEMIÓTICA ORGANIZACIONAL E O CICLO DE VIDA DOS DADOS

Segundo Liu (2000), a Semiótica surge com Charles S. Peirce (1839-1914) como a "doutrina formal dos signos". Um signo é algo que está (*stands*) para alguém em relação a alguma outra coisa em algum aspecto ou capacidade. Sempre existe uma tríplice significante-significado-interpretante, na qual o significante é a forma material do signo, o significado é o objeto/ação/evento/conceito que o representa, e interpretante é o que faz a significação (faz sentido do signo). De maneira simples, a semiótica estuda como fazemos sentido dos signos e nos comunicamos.

A Semiótica Organizacional é um ramo da Semiótica que entende organizações como um sistema de signos, e a Escada (ou *Framework*) Semiótica é um de seus artefatos que favorece a análise e síntese de problemas. A Escada Semiótica é composta por 6 níveis (divisões): Morris (1938) propôs as 3 divisões tradicionais da Semiótica como sintática (estrutura dos signos), semântica (significado dos signos) e pragmática (uso dos signos); enquanto Stamper (1973) acrescentou 3 novas divisões: física (sinais e marcas), empírica (meios por onde os signos transitam) e mundo social (efeitos do uso dos signos em assuntos humanos). Essas 6 divisões compõem a Escada Semiótica, representada na Figura 2.1, sendo as camadas semântica/pragmática/mundo social responsáveis pelo uso, funcionamento e consequências sociais dos signos na comunicação humana, e os níveis físico/empírico/sintático responsáveis pelo modo como os signos são usados e estruturados em uma linguagem, organizados, armazenados e transmitidos, e quais propriedades físicas eles possuem em uma infraestrutura de técnica. Um analista (ou no caso deste trabalho, um projetista de *software*) pode utilizar a Escada Semiótica para identificar os diferentes problemas que possam surgir no seu objeto de análise, auxiliado pelos diferentes níveis de entendimento de comunicação presentes no artefato.

O Ciclo de Vida dos Dados é um tópico que aparece em áreas ligadas a infraestrutura da computação, e.g. Golfarelli e Rizzi (2009) tratam de *Data Warehouse* e Demchenko et al. (2013) tratam de *Big Data* para infraestruturas de dados científicos. Ball (2012) apresenta alguns Ciclos de Vida de Dados científicos que foram usados como base do que será proposto (mesmo que o escopo do *framework* conceitual seja apoiar qualquer contexto de IHD e não apenas ferramentas científicas). Skoutas e Simitsis (2007) tratam de procedimentos de *Extract, Transform, Load* (ETL) ligados a ontologia, demonstrando técnicas para lidar com dados em diferentes estágios do Ciclo de Vida e oferecendo uma alternativa para classificação e entendimento das relações entre os dados. Uma motivação deste trabalho é não tornar os artefatos complexos sem necessidade para favorecer o uso do *framework*, e para tal adotou-se que o Ciclo de Vida dos Dados possui 4 fases principais: (a) Produção e coleta dos dados, e.g. recepção de dados de sensores, pesquisas ou bases de dados; (b) Transformação e filtragem dos dados, e.g. limpeza e pré-processamento dos dados visando o próximo ciclo; (c) Apresentação e interação com os dados, e.g. apresentações e interações com os dados que possibilitem *insights* nas partes interessadas, agregação dos dados,

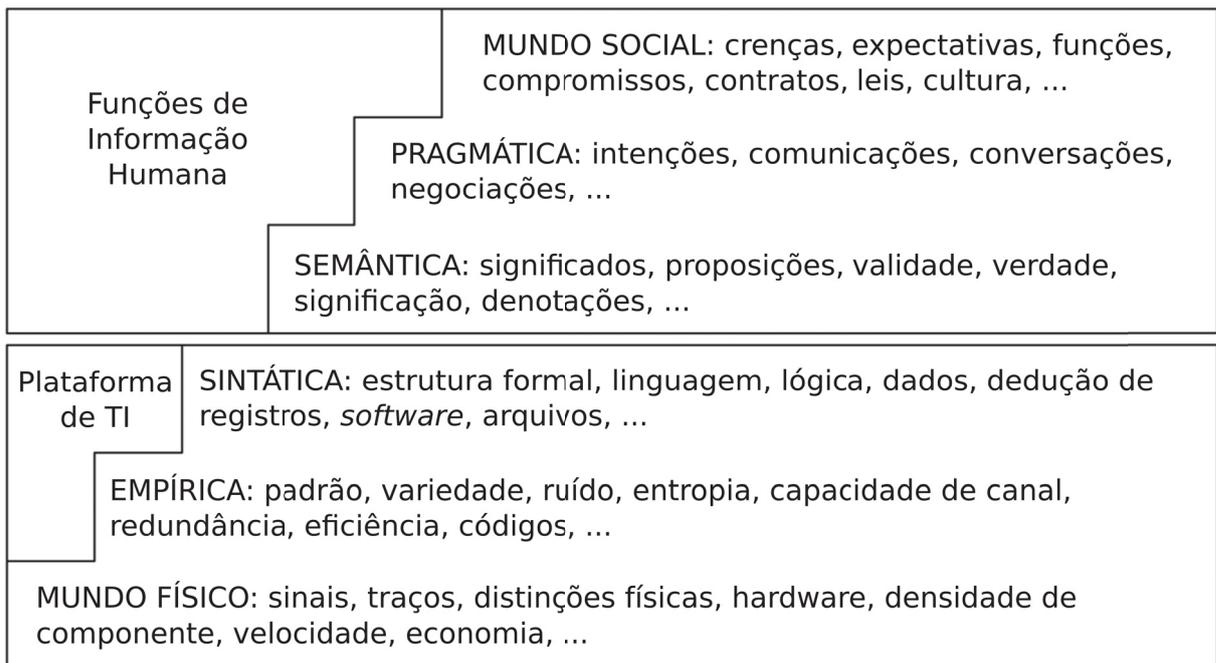


Figura 2.1: Escada Semiótica, retirada de Liu (2000) e traduzida pelo autor do trabalho.

apresentações que permitam novas interações, e (d) Manutenção dos dados, e.g. manter, arquivar e excluir dados armazenados no sistema. Esse modelo finda por ser mais específico para trabalhar com dados e ter várias preocupações diferentes que precisam ser consideradas na concepção do projeto de IHD.

Durante o Mapeamento Sistemático da Literatura, constatou-se que somente Hornung et al. (2015) consideram uma forma de entender os problemas da IHD sem restringi-los a uma faceta da área e, portanto, apresentam a solução que mais se encaixa nas necessidades de projetistas em contextos de IHD num sentido amplo, considerando o problema sob seus vários níveis de abstração. Para tal, os autores indicam justamente que se faça o uso da Escada Semiótica enquanto também se utiliza o Ciclo de Vida dos Dados.

Contudo, para utilizar esses artefatos é importante conhecer previamente o contexto do projeto, i.e., quem são as partes interessadas nesse problema, quais são suas expectativas e ações nesse contexto. A partir desse estudo, surge o questionamento: é possível usar artefatos específicos para reconhecer partes interessadas e suas ações que possam servir no modelo de *framework* proposto por Hornung et al. (2015)?

Seguindo a linha da Semiótica Organizacional, para reconhecer partes interessadas encontrou-se o artefato Diagrama de Partes Interessadas proposto por Stamper e publicado na tese de Kolkman (1993), que segundo Pereira e Baranauskas (2015), permite encontrar partes interessadas baseadas em diferentes níveis de envolvimento, interesses e expectativas. O Diagrama de Partes Interessadas apresentado na Figura 2.2 é ordenado em 5 camadas: a camada de Operação reúne o próprio sistema e quem irá desenvolvê-lo; a camada de Contribuição reúne as partes interessadas que irão ativamente contribuir para a solução do problema; a camada de Fonte reúne as partes interessadas que fornecem e consomem recursos da solução projetada; a camada de Mercado reúne os potenciais parceiros e concorrentes do projeto; e, por fim, a camada de Comunidade reúne terceiros, como cidadãos e legisladores, que afetam e serão afetados pela solução.

Para reconhecer as partes interessadas com base em seu envolvimento, encontrou-se no mapeamento realizado a distinção tripartida dos dados de Locoro (2015) e Cabitza e Locoro

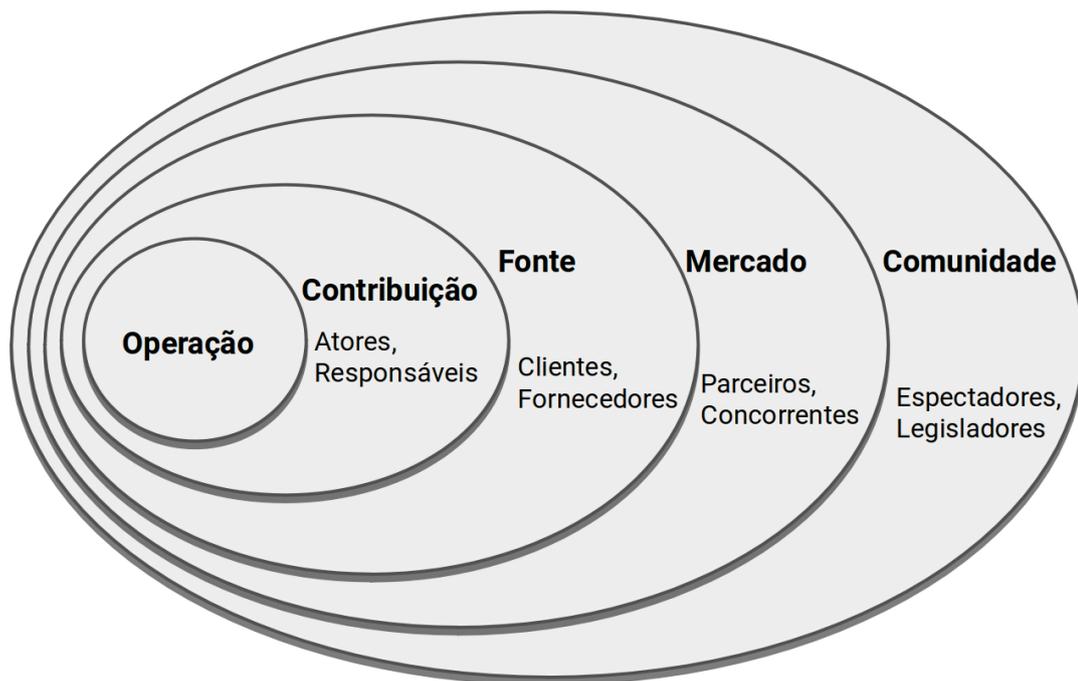


Figura 2.2: Diagrama de Partes Interessadas, adaptado de Pereira e Baranauskas (2015).

(2016), que utiliza o conceito *derivative data* de Floridi (2003), na qual existem três tipos de dados: (a) *dados primários* – produzidos por e para uma prática, portanto, não são facilmente destacáveis de seu contexto de criação, manipulação e uso; (b) *dados secundários* – englobam os processos pelos quais o pessoal administrativo, gerencial e de pesquisa dão sentido aos dados primários, muitas vezes processando-os também para torná-los mais adequados à sua interpretação profissional, e (c) *dados terciários* – englobam atividades que dizem respeito, em grande parte, ao uso de dados de categorias sociais de pessoas, como formuladores de políticas, usuários finais, contribuintes e cidadãos (a maioria desses usuários de dados não compartilha os mesmos propósitos e as mesmas capacidades de outras categorias profissionais em lidar com "seus dados"). Esta diferenciação expõe algumas formas de como as partes interessadas se envolvem com dados e pode servir como complemento ao Diagrama de Partes Interessadas. Neste trabalho, adotou-se o termo Tipos de Partes Interessadas ao invés do termo Tripartição de Dados porque o objetivo do artefato é identificar a parte interessada que manipula o dado (e não apenas o dado sendo manipulado).

Também não foi identificado no mapeamento uma técnica ou artefato para identificar as ações ou demandas das partes interessadas. Para esse propósito, nesta dissertação optou-se por utilizar o artefato Histórias de Usuário, que mapeia a ação de uma parte interessada em um contexto/domínio. Jeffries (2001) e Cohn (2004) descrevem uma História de Usuário como uma mescla de 3 artefatos: cartões (onde são escritas as Histórias de Usuário), conversas (onde o cliente conversa com os programadores sobre o projeto), e confirmações (teste de aceitação para validar a implementação das Histórias de Usuário). Ainda, Cohn (2008) sugere o formato "Eu, enquanto <parte interessada> quero <algum objetivo> para <alguma razão>" para o cartão da HU, o qual foi adotado no presente trabalho.

Dada a natureza genérica das Histórias de Usuário ao identificar as ações das partes interessadas no projeto, se faz necessário orientar sua elaboração para o contexto da IHD. Assim, o projetista deve refletir sobre os elementos conceituais de IHD mapeados na literatura ao executar a atividade. Dado que foram mapeados 25 elementos, com possíveis intersecções no

entendimento do problema, optou-se por selecionar um subconjunto de oito e agrupá-los por afinidade aos temas centrais da IHD: Legibilidade [*Awareness*, e Relacionamentos]; Agência [Autonomia, Consentimento Informado, e Colaboração]; e Negociabilidade [Privacidade, Ética, Propriedade dos dados].

O mapeamento sistemático em IHD demonstrou a necessidade por propostas que visam auxiliar o entendimento de problemas em contextos de IHD, de uma forma abrangente. A proposta de Hornung et al. (2015) é a única que considera o entendimento de problema sem limitar a um ou outro elemento conceitual. Para chegar a esse nível de abstração, faz-se uso do Ciclo de Vida dos Dados em conjunto com a Escada Semiótica. Contudo, projetistas inexperientes podem ter dificuldade em trabalhar com estes artefatos pela complexidade de análise exigida. Uma forma de incrementar essa metodologia, portanto, é fornecer recursos para fomentar o conhecimento do projetista a respeito do domínio abordado, momento oportuno para identificar as partes interessadas e o que elas fazem no projeto.

2.3 MARS - *MAPPING REVIEW SYSTEM*

Realizar um Mapeamento ou Revisão Sistemático de Literatura é uma tarefa complexa e envolve a aquisição, filtragem, análise e produção de uma grande quantidade de dados e informações. Portanto é possível considerá-la uma atividade de IHD, segundo as definições da área discutidas anteriormente. Dada a complexidade da tarefa, é esperado que exista um ferramental técnico que possa suportá-la em completude. Apenas duas ferramentas suportam especificamente a tarefa, sendo a StArt (por Zamboni et al. (2010) e Parsifal (por Freitas (2014)), sendo que os outros *softwares* comumente utilizados em mapeamentos apoiam apenas parte do processo (e.g. gerenciadores de referência ou planilhas eletrônicas). Tanto o StArt quanto o Parsifal apoiam o processo de Mapeamento e/ou Revisão Sistemática de Literatura passando por todas as fases, do planejamento à execução à sumarização dos dados, e possuem suporte limitado a realização do processo de forma colaborativa. O StArt e está restrito ao sistema operacional *Windows*, enquanto o Parsifal é uma ferramenta online *open-source* que pode ser hospedada por qualquer indivíduo ou organização.

Nesse contexto, o autor do presente trabalho percebeu uma oportunidade de criar uma ferramenta de apoio ao Mapeamento e Revisão Sistemática de Literatura com foco na colaboração online, chamada MaRS¹, e aproveitou-a para experimentar empiricamente o processo de desenvolvimento de *software* em um contexto de IHD, fazendo uso do *framework* conceitual desenvolvido neste trabalho no entendimento e *design* do sistema, bem como recebendo *feedbacks* que puderam exemplificar o uso do *framework* conceitual e aprimorá-lo.

Um protótipo funcional do sistema implementado utilizando *Ruby On Rails* e *Ember.js* pode ser conferido nas Figuras 2.3 (apresenta uma lista de trabalhos selecionados no mapeamento) e 2.4 (apresenta um grafo que relaciona os trabalhos mapeados com suas referências). O protótipo funcional foi utilizado pelo autor na execução do Mapeamento Sistemático de Literatura sobre IHD e possui funcionalidades que se mostraram necessárias para o processo. O protótipo também passou por duas atividades de ideação com auxílio do *framework* conceitual (vide Capítulo 4), e encontra-se em fase de desenvolvimento.

O protótipo se baseou no processo de Revisão Sistemática de Literatura apresentado por Petersen et al. (2015) e na ferramenta StArt de Zamboni et al. (2010) e segue três etapas: (1) Planejamento; (2) Execução; e (3) Sumarização. O protótipo permite que um grupo (e.g.

¹Código-fonte disponível em <https://gitlab.com/mappingreviewsystem/mrs-frontend> e <https://gitlab.com/mappingreviewsystem/mrs-backend>

Mapping Review System 1. PLANEJAMENTO 2. EXECUÇÃO 3. SUMARIZAÇÃO INTERAÇÃO HUMANO-DADOS

1. BASES DE DADOS E ARTIGOS 2. SELEÇÃO DE TRABALHOS 3. EXTRAÇÃO DE DADOS

Procurar artigo 28 artigos encontrados
28 extraídos

Extraído x Selecionado x Adicionar filtro

Título	Ano	Tipo	Estado	Tags	Publicador	Referenciado por
Human Data Interaction: Historical Lessons from Social Studies and CSCW	2015	Capítulo de Livro	Extraído	01 04 05 06 07 08 09 10 11 Core Importante	Springer International Publishing	7
Human-Data Interaction: The Human Face of the Data-Driven Society	2014	Artigo de Revista	Extraído	01 04 05 06 07 08 09 10 11 Core Importante	Social Science Electronic Publishing	7
Using embodied allegories to design gesture suites for human-data interaction	2012	Artigo de Evento	Extraído	01 04 05 06 07 08 09 10 11 Core Importante	ACM Press	5
Embodied Human-Data Interaction	2011	Outro	Extraído	01 04 05 06 07 08 09 10 11 Core Importante		3
Human Data Interaction in IoT: The ownership aspect	2014	Artigo de Evento	Extraído	01 04 05 06 07 08 09 10 11 Guidelines Importante	IEEE	3
Human-data.interaction	2013	Outro	Extraído	01 04 05 06 07 08 09 10 11		2

Figura 2.3: Interface de seleção de trabalhos na MaRS.

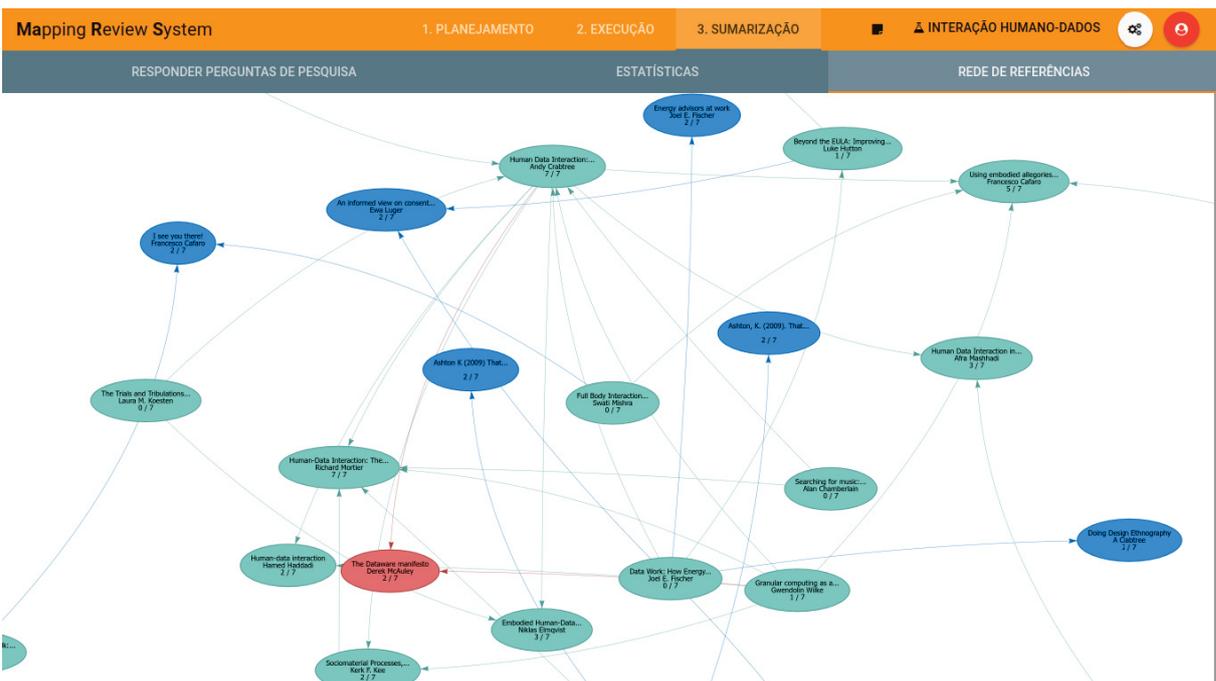


Figura 2.4: Rede de trabalhos selecionados (em verde) e suas referências (não selecionados em azul e descartados em vermelho) na MaRS.

pós-graduando e orientador) realize as seguintes atividades (descritas em Histórias de Usuário na Tabela 4.2):

1. Crie um mapeamento, preenchendo seu título, objetivo, palavras-chave, perguntas de pesquisa, critérios de seleção dos trabalhos, e perguntas para a extração de dados na etapa de planejamento;

2. Importe trabalhos de bases de dados por meio de arquivos *BibTeX* (e.g. da *ACM Digital Library*) ou trabalhos avulsos, importando os metadados automaticamente quando existir DOI. Aplique os critérios de seleção nos trabalhos encontrados, e extraia os dados dos trabalhos incluídos no mapeamento;
3. Responda as perguntas de pesquisa iterando os dados extraídos de cada artigo, e visualize estatísticas e relações entre autores, ano de publicação, publicações, vinculações e bases de dados, e visualize a rede de referências dos trabalhos incluídos no mapeamento.

3 HUMAN-DATA IDEATION

Os estudos focados em IHD são relativamente recentes (os primeiros remetem a 2011). Se por um lado existe muito a ser explorado, por outro a falta de literatura pode complicar seu entendimento a pessoas que queiram estudar ou trabalhar na área. Vejamos: se um novo projetista precisa trabalhar em um projeto que exista manipulação e fazer sentido de dados e não sabe por onde começar, gastará um tempo considerável procurando artigos que tratem sobre assunto e, mesmo assim, vai encontrar produções com propósitos específicos ou que abarquem uma porção limitada do entendimento de IHD. Assim, limita-se a aplicabilidade no contexto do desenvolvimento de *software* ao fragmentar o entendimento sistêmico da área, além de dificultar a divulgação do tema e o reconhecimento de sua importância. A falta de um método de ideação ou concepção de *softwares* que considere boa parte as facetas da IHD, preferencialmente antes do desenvolvimento de uma solução, afeta o potencial dos projetos ao restringir a implementação de um sistema em um conjunto limitado de preocupações. Para abordar esta problemática, neste capítulo será tratado sobre o *framework* conceitual HDIdea, seus passos, execução e motivações. O esquema geral do *framework* pode ser conferido na Figura 3.1 e um manual prático direcionado a apoiar projetistas em sua aplicação encontra-se no Apêndice B.

Considerando a literatura e experiências empíricas na área de IHD, percebe-se que os contextos e aplicações são variados e necessitam, primeiramente, de clarificação para então serem tratados sob a óptica da IHD e suas múltiplas facetas. Trabalhos como Pacheco et al. (2015) ao explorar um museu com Realidade Aumentada, Nunes et al. (2017) estudando a rota dos turistas em uma ilha, ou Roberts e Lyons (2017) mensurando *productive talk* em ambientes de aprendizado exemplificam a diversidade de contextos encontrados na interação com dados. O HDIdea direciona o *design rationale* do *software* ao considerar atividades como a autonomia dos usuários em relação ao sistema e os dados, ou a compreensão por parte dos usuários de como dados são utilizados nos sistemas e o relacionamento entre esses dados, também questões como propriedade dos dados, segurança, privacidade e afins. O *framework* também tem o papel de fomentar o conhecimento da IHD entre projetistas, estes que irão afetar a experiência de muitos indivíduos com seus produtos e disseminar os conceitos da área em contextos variados.

No capítulo anterior foram identificados 3 passos importantes para o processo de projetar soluções em IHD e que, seguindo a ordem apresentada, servem de base para a aplicação do HDIdea:

1. Identificação e classificação das partes interessadas conforme seus papéis e intenções no problema;
2. Identificação das ações dessas partes interessadas no domínio do projeto, observando os principais elementos conceituais tratados em IHD;
3. Refinamento das Histórias de Usuário em diferentes níveis de abstração para definir requisitos de *software*.

Para cada um dos passos expostos anteriormente, existe um artefato conceitual para apoiar seu uso: no primeiro passo há o Diagrama de Partes Interessadas estendido com os Tipos de Partes Interessadas - adaptado da tripartição dos dados; no segundo são elaboradas Histórias de Usuário (HU) estendidas com os elementos conceituais identificados na literatura de IHD e, no terceiro passo é aplicado o Ciclo de Vida dos Dados estendido com a Escada Semiótica. O resultado de cada artefato serve de insumo à execução do próximo artefato, resultando em

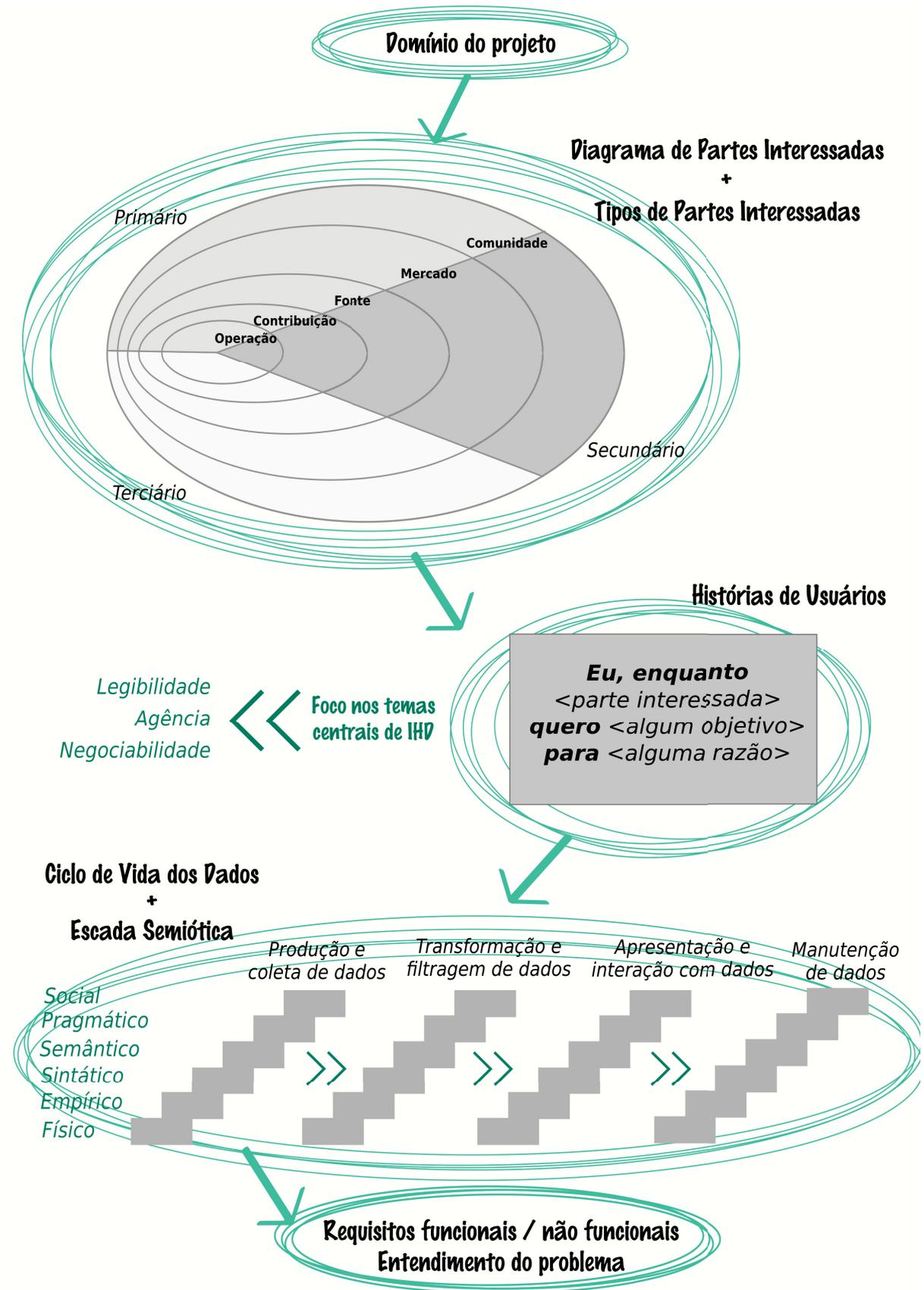


Figura 3.1: Esquema geral do HDIdea.

um entendimento em largura do projeto quando executado em atividades individuais (apenas o projetista) e em um entendimento compartilhado do problema quando executado em atividades colaborativas (com um grupo de partes interessadas no projeto).

3.1 PASSO 1 - IDENTIFICAÇÃO E CLASSIFICAÇÃO DAS PARTES INTERESSADAS

Como *entrada* para a atividade, o projetista deve compilar informações básicas do domínio que irá trabalhar, como: quem está solicitando o projeto, o que é esperado como resultado, prazos, recursos disponíveis, etc. O projetista então deverá ponderar se deseja executar as atividades do HDIdea com um grupo de partes interessadas que auxiliarão o projeto ou se irá executá-las de forma individual. A atividade em grupo permite que o projetista conheça diferentes pontos de vista sobre o problema/domínio por meio de um possível usuário do sistema, bem como permite criar um conhecimento compartilhado sobre o domínio, auxiliando o surgimento de soluções que consideram várias partes interessadas. Sabendo, entretanto, que nem sempre grupos de trabalho podem ser formados (seja por falta de recursos ou distância, por exemplo), o projetista pode utilizar o HDIdea como uma ferramenta para considerar o domínio por diversos ângulos, em largura.

Para exemplificar o uso do HDIdea, considere uma situação hipotética de um órgão federal de Meio Ambiente que precisa se adequar a Lei de Transparência Brasil (2009) e, para isso, licitou uma empresa para criar um *software* para a *web* que agregue os dados produzidos no órgão e seja visualizável pela a sociedade (conforme o estabelecido em lei), mas também que promova a pasta governamental perante a sociedade. Mário, o projetista responsável pelo sistema, dispõe de uma equipe de analistas e desenvolvedores de *software* e pode convidar partes interessadas para auxiliar no entendimento do problema. Como nunca esteve em um contexto com tantas partes interessadas e problemas de naturezas diversas a serem atacados, esse pesquisou o tema e encontrou o HDIdea, decidindo por usá-lo no projeto.

A primeira atividade a ser executada no processo de ideação é o preenchimento do Diagrama de Partes Interessadas, apresentado em Kolkman (1993), identificando as partes interessadas e observando a diferenciação das camadas presentes no artefato: (a) *Operação* – quem irá construir o sistema; (b) *Contribuição* – quem contribui diretamente para o problema ou para a sua solução e/ou são afetados diretamente por ele; (c) *Fonte* – quem fornece os dados/informações ou os consomem; (d) *Mercado* – parceiros ou concorrentes do projeto; (e) *Comunidade* – espectadores externos, sociedade e legisladores. Também deve-se observar a diferenciação dos Tipos de Parte Interessada: (I) *primária* – partes interessadas que produzem dados de suas próprias práticas/atividades/contextos (e.g. pesquisador realizando experimentos); (II) *secundária* – partes interessadas envolvidos em atividades de gerenciamento/administração/pesquisa, que refinam e agregam o que é feito pelos primários, geralmente de forma tabular, para seus próprios interesses; (III) *terciária* – fazem uso do que é produzido pelos primários/secundários para fins sociais (e.g. legisladores, usuários finais, cidadãos, etc). O diagrama do artefato encontra-se na Figura 3.2. Uma mesma parte interessada pode aparecer em mais de uma camada do Diagrama de Partes Interessadas e mais de um tipo do Tipos de Partes Interessadas. A *saída* desse artefato é uma lista de partes interessadas relevantes para o entendimento do problema, clarificando as motivações e relacionamento que estas partes tem com os dados no contexto do projeto.

Exemplo: Para o passo 1, Mário optou por realizar um *workshop* colaborativo entre seus colegas subordinados e partes interessadas que tiveram disponibilidade: um jornalista, um funcionário de uma Organização Não Governamental (ONG) relacionada com Meio Ambiente, alguns cidadãos interessados no projeto, e alguns funcionários do órgão contratante que usarão de forma direta ou indireta o sistema. Durante quarenta e cinco (45) minutos, ele foi o facilitador da

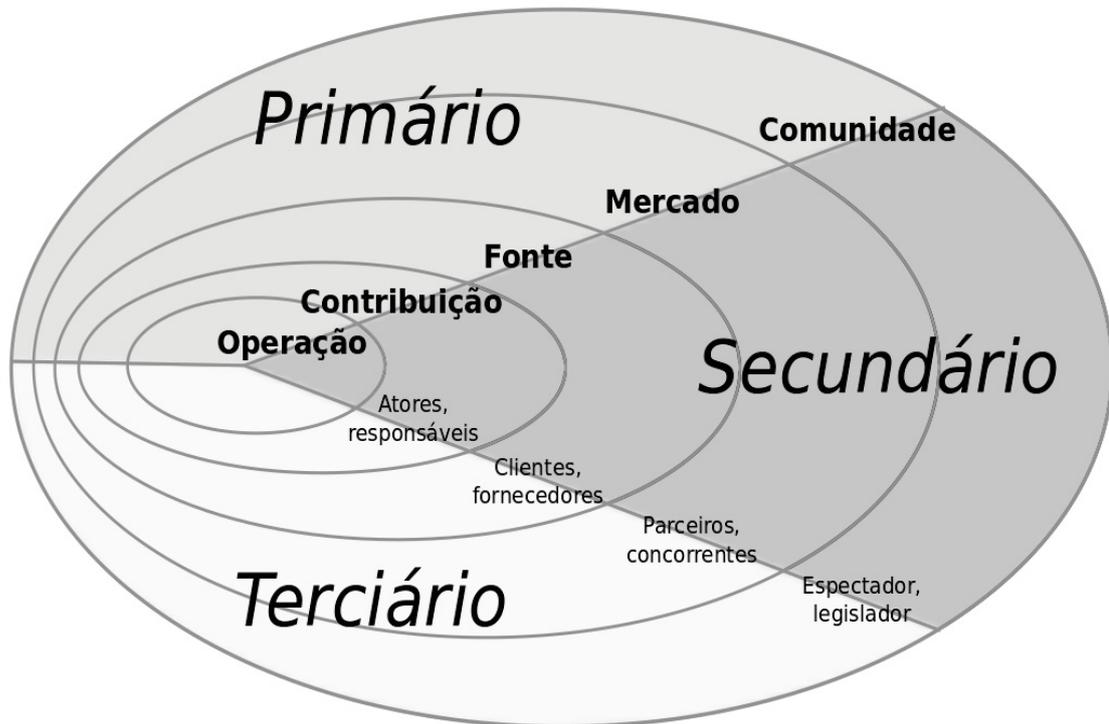


Figura 3.2: Passo 1 - Diagrama de Partes Interessadas estendido com Tipos de Partes Interessadas.

sessão e disponibilizou *post-its* para o preenchimento do artefato Diagrama de Partes Interessadas estendido com Tipos de Partes Interessadas, que foi impresso em um *banner* tamanho A0 (1,19m x 0,84m). Os participantes conversaram sobre as potenciais partes interessadas e seus papéis no sistema ao preencherem o artefato, compartilhando suas visões e entendimentos. As partes interessadas identificadas foram: (a-III) Analistas e desenvolvedores do projeto; (a-III) Empresa licitada; (b/c-I/II) Órgão Federal de Meio Ambiente; (b/c-I/II) Servidores do órgão; (c/d-II/III) ONGs relacionadas ao Meio Ambiente; (c/d-II/III) Jornalistas investigativos; (d-III) Políticos e pessoas com poder político que infringiram alguma lei ambiental; (d-III) Empresas afetadas por regulações ambientais; (e-III) Cidadãos interessados em Meio Ambiente; (d/e-III) Advogados especializados em Meio Ambiente; (d/e-III) Consultorias especializadas em Meio Ambiente; (e-III) Lei da Transparência; (e-III) Leis ambientais. Para aproveitar a reunião do grupo, o projetista decidiu continuar a atividade executando o passo 2.

3.2 PASSO 2 - IDENTIFICAÇÃO DAS AÇÕES DAS PARTES INTERESSADAS SOB OS ELEMENTOS CONCEITUAIS DE IHD

Com a lista de partes interessadas pronta, o projetista (ou o grupo de trabalho) deve escolher ao menos uma parte interessada de cada camada do Diagrama de Partes Interessadas e cada tipo dos Tipos de Partes Interessadas, de forma a diversificar as visões e necessidades que podem surgir no entendimento do problema. As partes interessadas escolhidas serão utilizadas como *entrada* para o preenchimento do artefato do passo 2.

Exemplo: Mário reservou 30 minutos para que os participantes escolhessem as principais partes interessadas, atentando para a diversidade de papéis que elas executam no projeto. As

partes interessadas escolhidas foram: Analistas e desenvolvedores do projeto; Órgão Federal de Meio Ambiente; Servidores do órgão; ONGs relacionadas ao Meio Ambiente; Jornalistas investigativos; Empresas afetadas por regulações ambientais; Cidadãos interessados em Meio Ambiente; Lei da Transparência e, Leis ambientais;

Com as partes interessadas principais selecionadas, pode-se utilizar o artefato de Histórias de Usuário. Uma história de usuário exprime o que as partes interessadas querem e/ou precisam do sistema e deve responder três questões: quem, o quê e por quê. Cohn (2008) sugere o formato "Eu, enquanto <parte interessada> quero <algum objetivo> para <alguma razão>", mas o projetista pode optar pelo formato que estiver familiarizado. O projetista deve criar ao menos uma história de usuário para cada um dos três temas centrais da IHD (Legibilidade, Agência e Negociabilidade), de maneira a considerar as múltiplas facetas da IHD no entendimento do problema, como evidencia Mortier et al. (2014).

Decidiu-se também relacionar os valores¹ de Pereira et al. (2013) identificados por Strey et al. (2018) na literatura de IHD com os temas centrais e expô-los por meio de perguntas ao projetista. Ao final da execução do passo 2, o projetista contará com uma lista de ações das partes interessadas no contexto do projeto. A seguir, serão demonstrados os temas centrais, seus valores relacionados, e um exemplo de uso para cada tema.

O tema *Legibilidade* se relaciona à necessidade de tornar compreensíveis os dados e algoritmos analíticos para pessoas sem domínio da área abordada, além de expor quais e como os dados são utilizados para fazer inferências no sistema projetado. Relacionados a este tema, foram selecionados 2 valores: (a) *Awareness* – o usuário precisa perceber quais dados estão disponíveis no sistema, como os usuários estão utilizando os dados, o que está acontecendo ou aconteceu; e (b) *Relacionamentos* – o usuário precisa ter condições de identificar relações entre dados e usuários do sistema. Na Tabela 3.1 estão relacionados os valores com as perguntas propostas para os projetistas utilizando o HDIdea.

Tabela 3.1: Valores relacionados ao tema *Legibilidade* e perguntas expostas no HDIdea.

Valor Associado	Pergunta para fomentar as HU
Awareness <i>O que está acontecendo no sistema.</i>	As partes interessadas tem como saber o que aconteceu ou está acontecendo com os dados? O processo e o fluxo dos dados é transparente e legível?
Relacionamentos <i>Como os dados e pessoas estão relacionados.</i>	Os dados revelam algum tipo de relação ou conexão entre entidades e pessoas? Existe alguma forma de observar essa conexão?

Exemplo: Mário optou uma atividade com duração de 70 minutos, explicando inicialmente aos participantes o que são Histórias de Usuário e propondo criá-las com base nos três temas centrais da IHD. Tratou então sobre o significado do tema *Legibilidade* e utilizou as perguntas propostas no manual do HDIdea para propiciar a discussão e criação das histórias. Distribuiu *post-its* e canetas para o preenchimento das histórias, fomentando a atividade durante 15 minutos e, posteriormente, utilizou mais 5 minutos para que todos os integrantes conhecessem e entendessem o que havia sido criado. Uma das possíveis Histórias de Usuário criadas pelo grupo foi: <Eu, enquanto> jornalista investigativo, <quero> ver uma rede de conexões entre pessoas físicas e jurídicas que cometem delitos ambientais, <para> investigar grupos criminosos e informar o público.

¹Valores identificados na literatura de IHD: *Awareness*, Relacionamentos, Autonomia, Consentimento Informado, Colaboração, Privacidade, Ética, Propriedade dos dados.

O tema *Agência* se relaciona com a capacidade do usuário de agir sobre os dados, controlando, informando e corrigindo dados ou inferências feitas a partir deles. Relacionados a este tema, foram selecionados 3 valores: (a) *Autonomia* – o usuário precisa poder decidir como agir de forma que atinja seus objetivos no sistema; (b) *Consentimento Informado* – o usuário precisa ficar ciente dos possíveis impactos de suas ações no sistema; e (c) *Colaboração* – o usuário precisa ter a possibilidade de trabalhar com outros usuários em um mesmo objeto. Na Tabela 3.2 estão relacionados os valores com as perguntas propostas para os projetistas utilizando o HDIdea.

Tabela 3.2: Valores relacionados ao tema *Agência* e perguntas expostas no HDIdea.

Valor Associado	Pergunta para fomentar as HU
Autonomia <i>Posso agir em prol dos meus objetivos.</i>	As partes interessadas podem corrigir algum dado sobre elas mesmas? As partes interessadas conseguem fazer o desejam no sistema de forma autônoma?
Consentimento Informado <i>Estou ciente das consequências do que faço.</i>	As partes interessadas estão conscientes sobre o que está sendo feito com os seus dados?
Colaboração <i>Consigo trabalhar com meus pares.</i>	É interessante que as partes interessadas consigam trabalhar em conjunto em determinadas tarefas do sistema?

Exemplo: Mário continuou a atividade tratando sobre o significado do tema *Agência*, e fez uso das perguntas propostas no manual do HDIdea. Durante 15 minutos os participantes discutiram e preencheram novas histórias e, nos 5 minutos restantes, foram apresentadas para que todos os integrantes entendessem a importância sob os diversos pontos de vista. Uma das possíveis Histórias de Usuário criadas pelo grupo foi: *<Eu, enquanto> cidadão interessado no Meio Ambiente, <quero> encontrar facilmente qualquer informação a meu respeito no sistema, <para> descobrir se minha privacidade foi violada ou se estou erroneamente indicado em alguma infração.*

O tema *Negociabilidade* se refere às relações sociais e jurídicas que surgem em torno dos dados e processamento de dados ao longo do tempo. Neste tema foram selecionados 3 valores: (a) *Ética* – o usuário precisa de condições para agir eticamente no sistema, bem como entraves para agir de forma antiética; (b) *Privacidade* – o usuário precisa ter o direito de escolher quais informações pessoais estarão disponíveis no sistema e para quais grupos; e (c) *Propriedade dos dados* – o usuário precisa ter o direito de posse dos dados que lhe pertencem e sobre as ações executadas perante eles. Na Tabela 3.3 estão relacionados os valores com as perguntas propostas para os projetistas utilizando o HDIdea.

Exemplo: Para finalizar a atividade, Mário explicou o significado do tema *Negociabilidade*, fazendo uso das perguntas propostas no manual do HDIdea. Novamente, os participantes discutiram e preencheram novas histórias por 15 minutos e apresentaram o que foi produzido para os outros nos 5 minutos restantes. Uma das possíveis Histórias de Usuário criadas pelo grupo foi: *<Eu, enquanto> Empresa afetada por regulações ambientais, <quero> pressionar o Órgão Federal responsável para retirar menções a meu respeito em crimes ambientais, <para> evitar propagandas negativas a meu respeito.* Assim o projetista se satisfaz com a contribuição do grupo de trabalho e encerrou as atividades do *workshop*.

Tabela 3.3: Valores relacionados ao tema *Negociabilidade* e perguntas expostas no HDIdea.

Valor Associado	Pergunta para fomentar as HU
Privacidade <i>Minha informações pessoais estão sob meu controle.</i>	As partes interessadas podem reivindicar sua privacidade?
Ética <i>Posso agir eticamente e esperar isso dos outros usuários.</i>	As partes interessadas podem ter comportamentos antiéticos com auxílio ou em relação ao sistema?
Propriedade dos dados <i>Os dados que produzo não são manuseados contra a minha vontade.</i>	Qual o aparato legal vigente e como as partes interessadas conseguem ter autonomia na lida com os dados?

3.3 PASSO 3 - REFINAMENTO DAS HISTÓRIAS DE USUÁRIO EM DIFERENTES NÍVEIS DE ABSTRAÇÃO

Neste passo, o projetista (ou o grupo de trabalho) pode conhecer de forma abrangente para quem o sistema será feito e quais as necessidades de uso do mesmo. O profissional, portanto, aumenta o arcabouço de conhecimento situacional, podendo se questionar sobre aspectos mais profundos do projeto. O terceiro passo apresenta o Ciclo de Vida dos Dados aprofundado pela Escada Semiótica, sendo o Ciclo de Vida dos Dados responsável por observar os dados em fases específicas de seu manuseio, enquanto a Escada Semiótica é um artefato vindo da Semiótica Organizacional que analisa um signo – uma informação ou mesmo um dado – por suas propriedades físicas e técnicas e também seu entendimento e uso por outros interpretantes (pessoas que lidam com ou são afetadas pelos dados). Recomenda-se o preenchimento do artefato analisando ao menos uma história de usuário (*entrada*) de cada um dos três temas centrais da IHD, pois o artefato pode levantar questionamentos não aparentes ao se pensar apenas nos anseios e atitudes dos usuários no sistema projetado, como nos passos anteriores. Como *saída* ao término da execução do passo 3, o projetista possuirá questões sobre problemas não necessariamente aparentes no que é esperado das atitudes dos usuários perante ao sistema, além de aprofundar os desafios técnicos que terá ao conceber o projeto.

O Ciclo de Vida dos Dados é composto por quatro fases:

1. Produção e coleta dos dados – quando os dados são produzidos ou recebidos por meio de sensores, pesquisas ou bases;
2. Transformação e filtragem dos dados – quando os dados disponíveis são preparados para o uso;
3. Apresentação e interação com os dados – quando se pode interagir com dados para que façam sentido dentro de um contexto, permitindo a ocorrência de *insights*;
4. Manutenção dos dados – quando são executadas políticas de armazenamento e exclusão dos dados.

Exemplo: Após a aplicação dos passos anteriores, Mário possui vasta noção de quem serão os usuários de seu sistema e quais as vontades e os comportamentos esperados no sistema, é possível refletir de maneira focada sobre um determinado dado. O projetista decide aproveitar a HU sobre o *lobby* de empresas que cometem crimes ambientais para pensar a respeito de como

é possível apresentar esses crimes de uma forma que informe e sensibilize os cidadãos, que poderiam agir de forma organizada para coibir esse tipo de iniciativa.

A Escada Semiótica apresenta uma separação conceitual entre o técnico e o humano. A camada técnica possui 3 níveis: (a) *mundo físico* – qual o suporte físico para o *software*? e.g. o *hardware* necessário; *empírico* – quais são os sistemas que comunicam usando o mundo físico? e.g. o ruído na transmissão dos dados; *sintático* – qual é a estrutura do *software*? e.g. quais dados podem ser cruzados. Já a camada social possui: *semântico* – o que a interface significa para o público alvo? e.g. qual o significado de disponibilizar um tipo de dado; *pragmático* – quais tarefas que a interface apoia? e.g. quais as intenções que as partes interessadas tem com os dados; *mundo social* – quais os efeitos do uso do *software*? e.g. expectativas e compromissos oriundos do uso dos dados.

Exemplo da primeira fase do ciclo: Mário ao utilizar a camada Empírica da Escada Semiótica e a fase "Produção e coleta de dados" do Ciclo de Vida dos Dados no contexto dos analistas e desenvolvedores do projeto se questiona: *O sistema possui uma entrada constante de dados ou existem períodos de tempo em que o tráfego é mais intenso? Usamos uma tecnologia que garante uma maior concorrência ou performance? Se o sistema permitir que cidadãos insiram dados para melhorar a fiscalização e este canal for aproveitado por robôs que postam informações no intuito de derrubar o sistema, como garantir que contribuidores oficiais, p. ex. as secretarias, não sejam afetados?*

Exemplo da segunda fase do ciclo: Mário ao utilizar a camada Sintática da Escada Semiótica e a fase "Transformação e filtragem de dados" do Ciclo de Vida dos Dados no contexto dos jornalistas investigativos se questiona: *Devemos impor uma forma padrão de organização dos dados para as secretarias para evitar possíveis mascaramentos de informação? Será que é possível, dado as diferentes temáticas? Essa forma padrão poderia melhorar a busca e relacionamento dos dados no caso de uma investigação?*

Exemplo da terceira fase do ciclo: Mário ao utilizar a camada Pragmática da Escada Semiótica e a fase "Apresentação e interação com dados" do Ciclo de Vida dos Dados no contexto do lobby das empresas criminosas se questiona: *Essas empresas possuem influência entre vários setores da sociedade e detêm poder sob vários políticos para manter seu lucro, apesar das irregularidades. Com os dados sobre os crimes sendo divulgados de forma que os responsáveis sejam identificados para o público, é possível que a pressão popular faça pressão para que o Poder Judiciário tome medidas em que as empresas sejam obrigadas a reparar os danos causados com grande repercussão, ampliando a visibilidade do projeto e favorecendo a causa ambiental.*

Exemplo da quarta fase do ciclo: Mário ao utilizar a camada Mundo Social da Escada Semiótica e a fase "Manutenção de dados" do Ciclo de Vida dos Dados no contexto das ONGs relacionadas ao Meio Ambiente se questiona: *Os dados preservados podem ser usados pela sociedade civil em diversos fins, ex. analisar gastos públicos com setores diferentes ao longo dos mandatos ou relembrar a atuação política de candidatos a cargos políticos. Nesse sentido, o software auxilia a descoberta de informações que aparecem em períodos grandes de tempo? Se sim, é possível relacionar iniciativas de longo prazo, como políticas e gastos com preservação do Meio Ambiente e cobrar atitudes dos governantes atuais e futuros.*

3.4 CONSIDERAÇÕES SOBRE O HUMAN-DATA IDEATION

Como saída ao final da utilização do HDIdea, o projetista contará com uma relação das diferentes partes interessadas do projeto, Histórias de Usuário das partes interessadas mais importantes orientadas com os principais elementos de IHD, e material com questões e problemas

do projeto que cobrem as diferentes fases do Ciclo de Vida dos Dados detalhados nos vários níveis de significado da Semiótica Organizacional. O resultado da aplicação do *framework* conceitual auxilia o entendimento do problema do projetista, fornecendo uma base apoiada na literatura de IHD e educando-o sobre a área, além de oferecer fomento para a identificação de requisitos funcionais e não funcionais de *software*.

A aplicação do HDIdea na concepção do projeto é recomendada para evitar possíveis mudanças bruscas de modelagem e implementação, mas o autor defende que em qualquer momento do desenvolvimento do *software* é possível utilizar o *framework* para tirar proveito do entendimento proporcionado sobre o domínio e incrementar positivamente o produto. O autor também entende que o trabalho aproxima os desafios e soluções da IHC com processos da Engenharia de Software, permitindo um melhor entendimento do contexto humano que será formalizado em uma solução técnica.

Visto que não foram encontradas soluções com a mesma proposta, o HDIdea foi criado com base nas experiências do autor com o contexto de IHD, seja na realização do mapeamento sistemático e desenvolvimento da ferramenta MaRS, como nos experimentos realizados com graduandos (público-alvo) e pós-graduandos (experiência em entendimento de problema) que serão apresentados no capítulo seguinte.

4 CASOS DE ESTUDO

Neste capítulo estão apresentados os casos de estudo de concepção, uso e análise do HDIdea. Tendo como ferramentas a MaRS e a SIMTransparência, foram realizadas quatro atividades, sendo as duas primeiras focadas em analisar se o uso dos artefatos atingia os resultados esperados para a concepção do HDIdea e as duas últimas focadas em analisar o uso do HDIdea e colher *feedbacks* para a versão de entrega do *framework* conceitual. A MaRS serviu de contexto para as atividades com discentes de pós-graduação, enquanto a SIMTransparência serviu de contexto para graduandos de Ciência da Computação. A Tabela 4.1 sintetiza os casos de estudos realizados.

Tabela 4.1: Casos de estudo realizados para a fomentar a criação do Human-Data Ideation.

	Caso de estudo	Participantes	Objetivos	Resultados
Concepção	maio/2018	6 colegas do grupo de pesquisa	utilizar os artefatos dos passos 1 e 2 e ampliar o entendimento do problema	identificou-se novas partes interessadas e Histórias de Usuário para a MaRS
	1º semestre/2018	31 alunos de Engenharia de Requisitos	utilizar os artefatos dos passos 1 e 2 e ampliar o entendimento do problema	propostas de melhorias para o portal SIMTransparência
Uso e Análise	2º semestre/2018	15 alunos de Engenharia de Requisitos – 3 respondentes do questionário avaliativo ao final das atividades	verificar se o uso do HDIdea auxilia no entendimento do problema	indícios sobre a utilidade do HDIdea para apoiar o seu público-alvo prospectivo
	dezembro/2018	5 colegas do grupo de pesquisa	verificar se o uso do HDIdea auxilia no entendimento do problema	pontos fortes e fracos do HDIdea para apoiar o entendimento de problemas

4.1 CASO DE ESTUDO DE CONCEPÇÃO DO *HUMAN DATA IDEATION* APLICADO NA FERRAMENTA MARS

A primeira utilização de um protótipo do *framework* conceitual se deu com colegas do grupo de pesquisa do autor durante 2 horas no dia 10/05/2018, realizando uma atividade com dois objetivos: 1. elicitare requisitos funcionais para a ferramenta MaRS (que estava em desenvolvimento na época), e 2. entender se o uso dos artefatos alcançava o objetivo do HDIdea (permitindo a futura aplicação do mesmo com o público-alvo, projetistas iniciantes). Esse grupo foi selecionado para a primeira atividade pelo fato dos pós-graduandos estudarem IHC em suas pesquisas de mestrado ou doutorado, o que lhes conferia conhecimento sobre a temática de "entendimento de problemas", e porque também realizavam, no momento, pesquisas no formato de Mapeamento Sistemático de Literatura, tornando-os boas partes interessadas.

Estavam presentes 6 pós-graduandos, que executaram os passos propostos, o autor e o orientador do presente trabalho, que ficaram responsáveis por mediar e registrar a atividade

(atuando como facilitadores). Com base em um Diagrama de Partes Interessadas e Histórias de Usuário pré-preenchidas, os envolvidos executaram apenas os passos 1 (Diagrama de Partes Interessadas estendido com os Tipos de Partes Interessadas) e 2 (Histórias de Usuário) do *framework* conceitual. A execução dos passos ocorreu de maneira escrita e oral. Os materiais disponíveis para a realização desta atividade foram: *post-its*, canetas esferográficas, *banner* A0 com o artefato Diagrama de Partes Interessadas impresso, computador para mostrar a apresentação de *slides* e o *site* da ferramenta MaRS.

Durante os primeiros 15 minutos, os facilitadores apresentaram um panorama geral sobre IHD, Mapeamento Sistemático de Literatura e sobre a ferramenta MaRS (para que servia e como funcionava). Metade do grupo estava utilizando a ferramenta no momento ou já tinha a utilizado (conhecendo o contexto e podendo trazer observações sobre seu uso). Em seguida foi apresentado o Diagrama de Partes Interessadas e a tripartição de dados (que futuramente adaptou-se para Tipos de Partes Interessadas), explicando como funcionavam e quais resultados eram obtidos com sua utilização. Também citaram-se alguns exemplos de partes interessadas (CAPES, C3SL, Desenvolvedores interessados, etc.), buscando facilitar o entendimento e aguçar a percepção sobre o problema. Cada participante recebeu *post-its*, nos quais deveriam escrever as partes interessadas identificadas e, posteriormente, colar no *banner* com o Diagrama de Partes Interessadas. Todo o processo durou, em média, 30 minutos. Na Figura 4.1 observa-se o Diagrama de Partes Interessadas preenchido durante a atividade. As partes interessadas assinaladas com um asterisco foram elencadas pelos participantes da atividade, ajudando o autor da dissertação a ampliar o entendimento do problema.

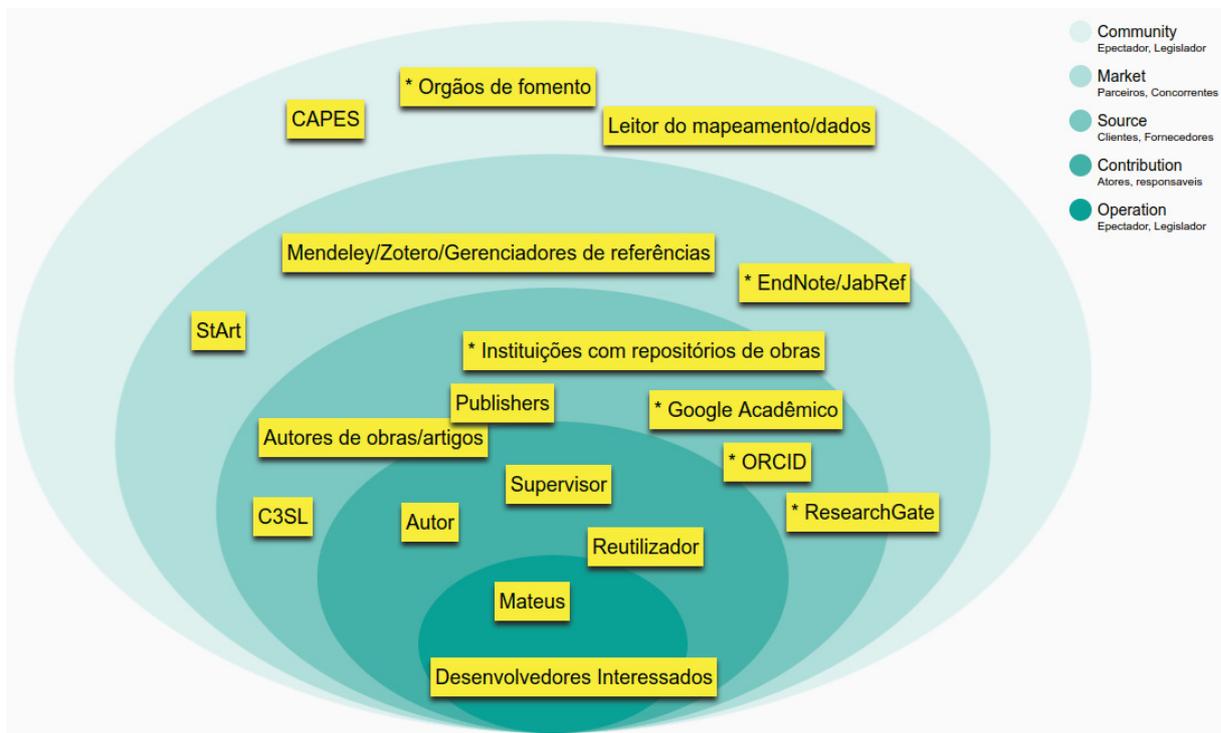


Figura 4.1: Diagrama de Partes Interessadas da MaRS. Os partes interessadas que começam com asterisco (*) foram identificados no piloto do *framework* conceitual com a ferramenta MaRS.

Finalizando essa etapa, o autor leu em voz alta cada parte interessada e pediu que os pós-graduandos selecionassem as mais relevantes para o levantamento de requisitos funcionais da MaRS. Foram selecionadas o Autor do Mapeamento, Supervisor do Mapeamento, Reutilizador do Mapeamento, Leitor do Mapeamento e *Publisher* de artigos. Partindo para as Histórias de

Usuário, o autor explicou o funcionamento da mesma e mostrou algumas já preenchidas para o contexto da MaRS. Entregou-se papel e caneta para cada participante, sendo solicitado que cada um escrevesse uma História de Usuário para cada das cinco partes interessadas selecionadas. Durante essa etapa, que durou 50 minutos, os participantes anotavam as Histórias de Usuário nos papéis e discutiam entre si e com os facilitadores a importância das histórias identificadas segundo seu ponto de vista. O resultado do preenchimento está na Tabela 4.2, sendo que os escritos em itálico foram pré-preenchidas para a atividade e as em negrito estão implementadas no protótipo funcional da MaRS.

Ressalta-se que a atividade inteira foi permeada por diálogo entre os participantes e facilitadores, o que promoveu o compartilhamento do entendimento do problema. Como resultado, foram alcançados requisitos funcionais que fazem sentido para a ferramenta MaRS (sob a forma de Histórias de Usuário) e os participantes pareceram ter aumentado seu entendimento do problema, segundo o que foi relatado pelos mesmos e observado durante a atividade. Outro ponto que considera-se importante é que o autor estava concluindo tanto o Mapeamento Sistemático de Literatura sobre IHD bem como a implementação da MaRS - configurando um cenário de constante contato com aspectos de IHD. A aplicação, embora parcial, do HDIdea ampliou o entendimento do problema do autor após a atividade, indiciando uma constatação embrionária de que o *framework* conceitual serve para seu propósito.

4.2 CASO DE ESTUDO DE CONCEPÇÃO DO HDIDEA APLICADO COM UMA TURMA DE ENGENHARIA DE REQUISITOS

No 1º semestre de 2018, o autor realizou a disciplina de "Prática em Docência de Informática" em uma turma de Engenharia de Requisitos do Departamento de Informática (DINF) do Setor de Ciências Exatas da UFPR. Aproveitando que a ementa de Engenharia de Requisitos trata sobre a temática de requisitos e o projeto da turma ser relacionado a transparência de dados, além de que a disciplina cursada prevê a aplicação de atividades na turma que está sendo acompanhada, optou-se por realizar a aplicação do HDIdea na mesma. A atividade, portanto, foi aplicada no final do semestre em dois momentos distintos utilizando a ferramenta SIMTransparência. Escolheu-se essa ferramenta pois: (a) a temática de projeto final seria em IHD, focando na criação de soluções para a SIMTransparência; (b) a mesma está inserida no contexto de IHD; (c) é desenvolvida como projeto dentro da UFPR, facilitando o acesso às partes interessadas; (d) havia possibilidade de aperfeiçoá-la, e (e) havia possibilidade de estabelecer troca de conhecimento entre os estudantes de Engenharia de Requisitos e os criadores da ferramenta. Com essa turma também foram aplicados apenas os passos 1 e 2 do *framework* conceitual.

A SIMTransparência¹ é uma ferramenta de visualização de gastos públicos do Ministério da Educação (MEC) e foi desenvolvida pelo Centro de Computação Científica e *software* Livre (C3SL) UFPR como uma experimentação dos recursos de um conjunto de *softwares* de armazenamento e visualização de dados, ELK Stack², e encontra-se no contexto de IHD. Como o SIMTransparência é um experimento e seu desenvolvimento ainda é corrente, podem-se explorar e sugerir recursos de IHD (motivo pelo qual foi escolhido para ser trabalhado durante a disciplina de Engenharia de Requisitos).

A atividade com o *framework* conceitual foi executada em 2 dias e teve duração média de 1h30 em cada dia. Em 23/05/2018 o co-criador e até então responsável pelo SIMTransparência, Diego Pasqualin, apresentou a ferramenta para a turma e, em seguida, o autor do presente trabalho

¹SIMTransparência: <http://www.c3sl.ufpr.br/transparencia/>

²ELK Stack: <https://www.elastic.co/elk-stack>

Tabela 4.2: Histórias de usuário identificadas no piloto do *framework* conceitual com a ferramenta MaRS. Em itálico o que foi pré-preenchido. Em negrito o que foi implementado no protótipo.

Como um [ator]	Eu quero/preciso de/devo/gostaria de [ação]	para [funcionalidade]
<i>Autor, Supervisor</i>	<i>criar um novo mapeamento</i>	<i>propor o mapeamento</i>
<i>Autor, Supervisor</i>	<i>importar os artigos das bases de dados</i>	<i>executar o mapeamento</i>
<i>Autor, Supervisor</i>	<i>aplicar os critérios de inclusão/exclusão</i>	<i>selecionar os artigos/obras do mapeamento</i>
<i>Autor, Supervisor</i>	<i>responder as perguntas de extração</i>	<i>extrair os dados para responder as perguntas de pesquisa</i>
<i>Autor, Supervisor</i>	<i>responder as perguntas de pesquisa</i>	<i>completar o mapeamento</i>
<i>Reutilizador</i>	<i>copiar mapeamento</i>	<i>aproveitar/atualizar o mapeamento prévio em um novo mapeamento</i>
<i>Leitor</i>	<i>visualizar os dados</i>	<i>tirar suas próprias conclusões do mapeamento</i>
Autor, Supervisor, Reutilizador	gerar automaticamente a <i>string</i> de busca para as bases de dados	não precisar montar uma <i>string</i> personalizada para cada base de dados
Autor, Supervisor, Reutilizador	formar um grupo de trabalho	trabalhar no mapeamento/revisão em equipe
Autor, Supervisor, Reutilizador	corrigir os dados extraídos facilmente	perder menos tempo arrumando os dados
Autor, Supervisor, Reutilizador	gerar gráficos e tabelas personalizadas a partir do mapeamento	não depender de editores de planilhas e outras ferramentas além da MaRS
Autor, Supervisor, Reutilizador	cruzar perguntas de pesquisa (?)	
Autor, Supervisor, Reutilizador, Leitor	exportar e importar pesquisas	salvar localmente ou em lugares alheios ao servidor que hospeda a MaRS
Autor, Supervisor, Reutilizador, Leitor	ver o h-index/ <i>Google Citations</i> dos autores e JCR das revistas	classificar trabalhos por ordem de relevância
Autor, Supervisor, Reutilizador, Leitor	gerar um relatório com o protocolo da pesquisa	imprimi-lo e ter acesso ao protocolo fora da MaRS
Autor, Supervisor, Reutilizador, Leitor	ver se um artigo permitiu responder a uma pergunta de pesquisa específica	facilitar a releitura ou exploração do artigo
Autor, Supervisor, Reutilizador, Leitor	importar metadados diretamente do PDF	posso não ter o BibTeX e o DOI
Autor, Supervisor, Reutilizador, Leitor	configurar as cores da interface	agradar gostos estéticos
Autor, Supervisor, Reutilizador, Leitor	ver a quantidade de publicações por ano	exportar a estatística para o documento do mapeamento/revisão
Autor, Supervisor, Reutilizador, Leitor	visualizar a <i>tagcloud</i> dos artigos	ter uma noção do que é discutido pelos artigos importados
Autor, Supervisor, Reutilizador, Leitor	legenda do lado dos gráficos (?)	
Leitor	entender o processo do dado até ele ser apresentado (transparência) (?)	
<i>Publisher</i>	apenas usuários autorizados acessem meus artigos	não ter acesso aos artigos sem ganhar por isso

falou brevemente sobre a área de IHD e uma visão geral do HDIdea. Assim o primeiro momento foi focado na ambientação e contextualização da temática. Durante o primeiro dia, os estudantes demonstraram dúvidas em relação à ferramenta (seu funcionamento, as expectativas de uso e potencial) e à IHD (exemplos de aplicabilidade do *framework* conceitual) que foram sanadas tanto por Diego quanto pelo autor. Para a aplicação da atividade foram utilizados computador com apresentação de *slides* e projetor.

Em 25/05/2018, o autor aplicou os passos 1 e 2 do HDIdea com a turma. A atividade contou com 31 alunos e 2 facilitadores (o autor e orientador), iniciando com a retomada das instruções de execução do 1º passo e sua aplicação em conjunto pela turma. Como a turma demonstrou sentir vergonha de colar os *post-its* com suas ideias no *banner* A0 com o Diagrama de Partes Interessadas, os facilitadores iniciaram uma conversa a turma perguntando quais seriam as partes interessadas no contexto da ferramenta e em quais tipos de parte interessada estas inseriam-se. Observou-se resistência da turma em utilizar o Tipo de Partes Interessadas (mostraram-se mais interessados em identificar as partes interessadas do que classificá-las). Mas fazendo perguntas e incentivando a turma a participar, a atividade pode ser continuada.

Em seguida, foi lembrado como o passo 2 deveria ser executado. Individualmente, os estudantes deveriam criar Histórias de Usuários com a necessidade de cobrir pelo menos uma história relacionada com privacidade e uma história relacionada com transparência. As Histórias de Usuário foram escritas em *post-its* e coladas em outro *banner* A0, vide a Figura 4.2. As histórias de usuário identificadas foram então agrupadas por semelhança, como pode ser visto na Figura 4.3, para que os alunos e alunas (em grupos de até 5 integrantes) escolhessem um subconjunto e trabalhassem na criação de uma solução de IHD para o SIMTransparência, gerando um documento de especificação de requisitos, que serviria de avaliação final da disciplina.



Figura 4.2: Alunos de Engenharia de Requisitos colando suas Histórias de Usuário para o projeto final da disciplina.



Figura 4.3: Alunos de Engenharia de Requisitos colando suas histórias de usuário do projeto final da disciplina.

No final do semestre a turma (organizada em grupos) deveria entregar um trabalho final relativo a transparência dos dados, e identificou-se que os alunos propuseram: (a) um grafo de relacionamentos entre as entidades que receberam verba do MEC; (b) uma solução social para a ferramenta que envolveu um fórum, *tags* para pesquisas salvas e enquetes; (c) comparação de gastos entre entidades e uma sessão para que usuários comentem estas comparações; (d) acesso aos dados originais e funcionalidade para reportar dados incorretos; (e) espaço para denunciar possíveis irregularidades (crimes) nos gastos apresentados, salvar filtros aplicados na ferramenta, correções de *User Interface* (UI); (f) módulo de segurança para controlar possíveis ataques à ferramenta; (g) melhorias na interação com as visualizações da ferramenta, e (h) uma funcionalidade que identifica possíveis irregularidades nos gastos com *Machine Learning* e divulga com um *bot* por meio do *Twitter*. Os resultados foram promissores e os alunos cobriram os vários elementos conceituais da IHD, mesmo que com menos tópicos obrigatórios da IHD no 2º passo e sem a aplicação do 3º passo (por falta de tempo). Os alunos autorizaram a divulgação das propostas quando questionados em sala, sem a identificação dos participantes dos grupos, e não foram aplicados questionários para que os alunos avaliassem o (uso do) *framework* conceitual.

Ainda que o 3º passo não tenha sido executado, um artefato Ciclo de Vida dos Dados estendido com a Escada Semiótica foi pré-preenchido com problemas e questões, podendo ser visualizado na Figura 4.4. O *input* utilizado foi "<Eu, como> cidadão <quero> investigar onde as universidades públicas aplicam de forma ruim seu dinheiro, <para> reclamar do Ensino Superior público brasileiro no grupo de família do WhatsApp".

4.3 CASO DE ESTUDO DE USO E ANÁLISE DO HDIDEA APLICADO COM UMA TURMA DE ENGENHARIA DE REQUISITOS

No 2º semestre de 2018, o autor acompanhou a primeira metade do semestre de uma outra turma de Engenharia de Requisitos. Como parte das atividades para a especificação de requisitos de *software*, o autor teve a oportunidade de aplicar o 3 passos do HDIdea. Neste semestre, a

Framework Semiótico	Ciclo de vida dos dados			
	produção e coleta de dados	transformação e filtragem de dados	apresentação e interação com dados	manutenção de dados
Social	- Os cidadãos brasileiros possuem os dados das contas das universidades públicas, mas não conseguem fazer sentido deles.		- A sociedade se beneficia com o acesso aos gastos das universidades e pode cobrar a ferramenta para gastos em outras esferas.	- Os dados preservados podem ser usados pela sociedade civil em diversos fins, e.g. propor o aumento de verba para as universidades públicas.
Pragmático	- Os gestores das universidades públicas estão mais atentos com seus processos internos de movimentação de dinheiro.	- O C3SL depende de stakeholders externos (com seus contextos) para ajudá-lo a entender como os dados adquiridos tornam-se úteis.	- Se o cidadão consegue encontrar os dados que quer, divulga a ferramenta para seus familiares. Se não, tem mais um motivo para reclamar das universidades públicas.	...
Semântico	- Os dados criados pelas universidades expõem como elas são geridas.	- O C3SL precisa processar os dados de forma que a apresentação deles seja possível.	- O cidadão precisa encontrar os dados sobre os gastos fora do padrão em poucos cliques.	...
Sintático	- Os dados criados pelas universidades precisam seguir alguma ordem/sintaxe para seu armazenamento.	- O C3SL precisa transformar o padrão (sintático) de dados do governo em um padrão consumível pelo ElasticSearch.	- Os dados do ElasticSearch precisam ter uma sintaxe que favoreça os diversos usos.	...
Empírico	- O governo federal precisa disponibilizar uma infraestrutura que receba os dados das universidades.	- O C3SL precisa de uma infraestrutura de rede que suporte a carga inicial e as cargas subsequentes de dados dos gastos das universidades brasileiras	- O C3SL precisa de uma infraestrutura de rede que suporte o possível acesso simultâneo de stakeholders do país inteiro.	...
Físico	- O governo federal precisa de um maquinário que suporte a quantidade de dados produzidos pelas universidades.	- O C3SL precisa de um maquinário que suporte a quantidade de dados produzidos pelas universidades e disponibilizados pelo governo federal.	- O C3SL precisa de maquinário que suporte o possível acesso de stakeholders do país inteiro.	- O C3SL precisa manter seu maquinário funcionando, sem perder os dados.
User Story: Eu, como cidadão quero investigar onde as universidades públicas aplicam de forma ruim seu dinheiro para reclamar do Ensino Superior público brasileiro no grupo de família do WhatsApp.				

Figura 4.4: Exemplo pré-preenchido da Escada Semiótica em uma história de usuário genérica para o contexto da ferramenta SIMTransparência.

primeira parte da disciplina utilizou o contexto de IHD com a ferramenta SIMTransparência como uma forma de apresentar os conceitos básicos da matéria, momento em que o autor acompanhou a turma. A atividade com o HDIdea foi executada em 3 dias com 1h30 de duração em cada dia. Ainda, foi possível aplicar um questionário sobre a experiência (que ficou disponível para preenchimento por 3 meses).

Em 17/09/2018 o autor apresentou uma breve explicação sobre a IHD, um panorama geral do HDIdea e também a ferramenta SIMTransparência para a turma, contando com 9 alunos em sala. Após a apresentação inicial, o autor e o seu orientador, atuando como facilitadores, executaram com a turma o primeiro passo do HDIdea, a Identificação das Partes Interessadas - utilizando o artefato Diagrama de Partes Interessadas e Tipos de Parte Interessada. Os facilitadores trouxeram como exemplos de partes interessadas Cidadãos, Fiscais, e C3SL para fomentar o início da atividade. Após explicação detalhada dos artefatos, os alunos e alunas foram reunidos em torno de uma mesa, servidos com *post-its* para anotar as partes interessadas observadas e incentivados a colá-los em um *banner* A0 contendo o Diagrama de Partes Interessadas. Ao final da atividade, os alunos identificaram mais 20 partes interessadas, vide Figura 4.5, mas apenas 3 alunos utilizaram também o artefato Tipos de Partes Interessadas, talvez por não estar explícito no *banner* utilizado na atividade.

Em 22/09/2018, o autor aplicou o segundo passo do HDIdea com a turma. A atividade contou com 15 alunos e 2 facilitadores (o autor e seu orientador) e começou em uma breve retomada do tema de IHD, da ferramenta SIMTransparência, do HDIdea e da sessão anterior com o Diagrama de Partes Interessadas estendido com os Tipos de Partes Interessadas, para situar os alunos que haviam faltado no primeiro dia da atividade. Seguindo para uma explicação sobre as Histórias de Usuário, a definição dos temas centrais³ da IHD de Mortier et al. (2014), valores de Pereira et al. (2013)⁴ e elementos conceituais adicionais identificados por Strey et al. (2018)⁵ na literatura de IHD. Então, os alunos e alunas foram reunidos em torno de uma mesa grande com o artefato do passo anterior preenchido para que fossem escolhidos as partes interessadas mais significativas, resultando em 18 partes interessadas escolhidas das 23 disponíveis. Foram

³Temas Centrais da IHD: Legibilidade, Agência e Negociabilidade.

⁴Valores de Pereira et al. (2013): Autonomia, *Awareness*, Colaboração, Consentimento Informado, Privacidade, Relacionamentos.

⁵Elementos conceituais identificados na literatura de IHD por Strey et al. (2018): Ética, Open Data e Transparência.

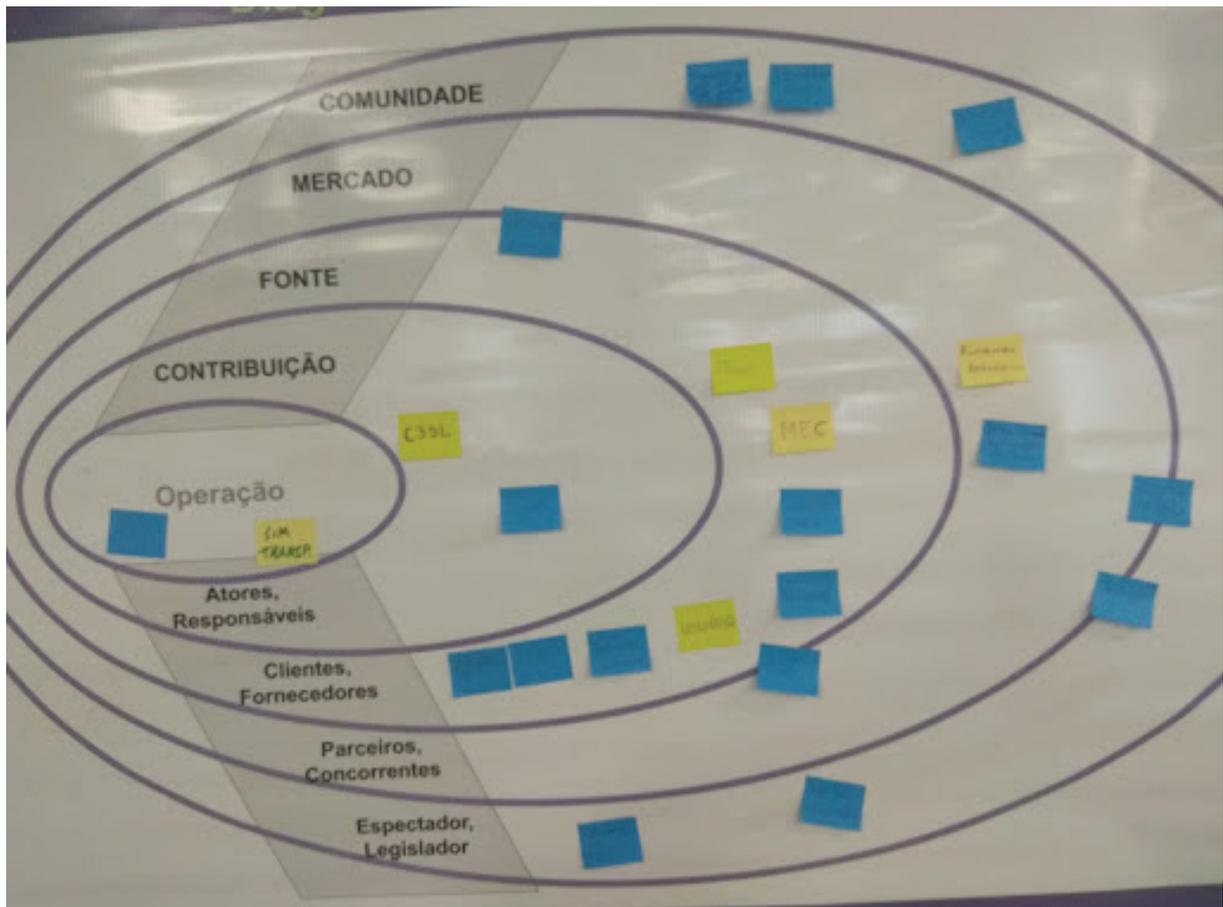


Figura 4.5: Diagrama de Partes Interessadas estendido com os Tipos de Partes Interessadas preenchido no dia 17/08/2018.

distribuídos *post-its* para o preenchimento das Histórias de Usuário, sendo que cada estudante deveria criar 3 Histórias de Usuário referentes a algum dos temas ou elementos conceituais apresentados. O quadro de giz da sala foi usado para colar o resultado da atividade, e os alunos optaram por agrupar as histórias por parte interessada. No final da atividade, os alunos se aproximaram do quadro para ler as outras Histórias de Usuário e discutiram entre si o que elas significavam para o SIMTransparência. O resultado da atividade encontra-se nas Tabelas 4.3 e 4.4.

No dia 24/08/2018, os alunos da matéria tiveram uma aula sobre *Brainwriting* usando como contexto a ferramenta SIMTransparência, sem a participação do autor (aplicado pelo orientador, no caso). No dia 29/08/2018, 12 alunos compareceram e tiveram uma aula com o autor e o orientador sobre o terceiro passo do HDIdea - Ciclo de Vida dos Dados estendido com a Escada Semiótica. O autor apresentou uma retrospectiva dos dois passos anteriores do HDIdea e então explicou os artefatos Ciclo de Vida dos Dados e Escada Semiótica. Os alunos foram orientados a trabalhar no primeiro projeto da matéria, escolhendo uma História de Usuário para embasar o preenchimento dos artefatos do terceiro passo (que seria entregue, portanto, apenas no final do semestre).

Contextualizando, no dia 31/09, a turma teve uma aula sobre Cenários e continuaram trabalhando na primeira entrega do projeto, que deveria conter: (a) Cenário de uso simplificado da solução proposta; (b) Lista de Requisitos Funcionais da solução; (c) Lista de Requisitos Não-Funcionais da solução; (e) Indicar a origem dos requisitos: Histórias de Usuário [HU], *Brainwriting* [BW], Ciclo de Vida dos Dados estendido com Escada Semiótica, e (d) Ciclo de

Tabela 4.3: Primeira parte das histórias de usuário identificadas durante aplicação do segundo passo do HDIdea com a turma de Engenharia de Requisitos do 2º semestre de 2018.

Eu, enquanto [parte interessada]	quero [algum objetivo]	para [alguma razão]
Advogado	acessar dados específicos de empresas ou órgãos quanto aos gastos públicos	poder defender ou atacar com mais e melhores dados. Exemplo: denunciar um gasto exorbitante.
C3SL	desenvolver um sistema de fácil manutenção	que o desenvolvimento não seja prejudicado pela alta rotatividade de bolsistas.
C3SL	saber informações (Browser, OS) de quem acessou o SIMTransparência	melhorar a compatibilidade e legibilidade do site para Browsers específicos.
C3SL	divulgação do C3SL e Conscientização do público	arrecadar fundos com features com mensalidade (SaaS).
Professor de Engenharia de Requisitos	acesso à documentação do projeto	utilizar como exemplo em sala de aula.
CGU	a geração automática e periódica de relatórios	facilitar o controle de gastos de áreas de interesse.
ElasticSearch	relatórios de desenvolvimento	saber como aperfeiçoar a ferramenta que forneço.
Empresa que observa concorrência de produtos nos orçamentos da União	informações sobre o valor e empresa que proveu o orçamento	aumentar negociabilidade com o governo (negociar meus produtos que foram providos por outra empresa).
Fiscais	sistema de criar tabela dinâmica e gráficos (parecido com o Excel). Cubos OLAP.	ajudar no meu trabalho.
Fundação Araúcaria	comparar as prestações de contas dos projetos que eu investi com as informações registradas no portal	conferir a fidelidade dos dados informados na prestação de contas.
Jornalista	analisar despesas realizadas por um departamento que não tenham orçamento previsto	que comprove irregularidades realizadas.
Jornalista	selecionar gastos de um grupo ou instituição específica	reportar fraudes nesse grupo ou instituição.
Ministério Público	ser notificado e ter comprovantes de gastos absurdos de dinheiro público	poder denunciá-los e oprimi-los
MEC	informações sobre gastos	distribuir investimentos com maiores demandas/produktividade.
MEC	alterar dados	poder corrigir possíveis erros nos dados.
Político	formulário de gastos	reportar meus gastos.
Político	possuir relatório de gastos, mais explícito possível	que possa avaliar e criar novos orçamentos de investimentos para a sociedade, assim como onde economizar
Político	relatório de gastos	apresentar melhores propostas e criar projetos de reformas para melhorar o país

Tabela 4.4: Segunda parte das histórias de usuário identificadas durante aplicação do segundo passo do HDIdea com a turma de Engenharia de Requisitos do 2º semestre de 2018.

Eu, enquanto [parte interessada]	quero [algum objetivo]	para [alguma razão]
Público em geral	poder fazer download de gráficos ou outros dados tratados	poder divulgar.
Público em geral	ter acesso a informação não obstante a limitações físicas	ter acesso a informação.
Público em geral	um site estável num servidor que aguento o alto tráfego	que não seja desconectado a cada dez segundos (como no Portal do Aluno, em época de matrícula).
Público em geral	poder acessar esses dados de diferentes formas talvez uma versão impressa com as informações mais buscadas.	
Outras universidades com iniciativas similares	saber mais como é construído o portal	poder construir um semelhante ou participar, junto da UFPR, do portal.
Quem decide o orçamento	poder ver para onde vai o dinheiro	que eu possa decidir mudanças no orçamento.
Servidor público	ter a opção de ser notificado sobre alterações em informações relacionadas a mim	estar ciente das informações pessoais que são públicas.
TCU	saber se o dinheiro está realmente sendo gasto com aquilo que foi orçado	garantir que não estão ocorrendo fraudes.
TCU	visualizar estatísticas comparativas entre vários gastos de mesma categoria	comparar gastos similares.
Usuário do portal	determinar uma data inicial e final	avaliar as despesas realizadas naquele período definido.
Usuário comum	uma interface simples, eficiente e com suporte	diferentes resoluções com um design inteligente. Ou seja, contate um designer.
Usuário comum	ter a possibilidade de dar <i>feedbacks</i> de sistema e possíveis erros de dado	de modo que a experiência do usuário torne-se melhor.
Novo usuário	receber treinamento/capacitação	utilizar as ferramentas do portal e onde buscar informações quando encontrar suspeitas/fazer denúncias.

Vida dos Dados estendido com a Escada Semiótica preenchido. No dia 05/09/2018, data da primeira entrega do projeto, apenas dois grupos entregaram (com três estudantes em cada grupo). Os artefatos do Ciclo de Vida dos Dados estendido com a Escada Semiótica preenchidos podem ser conferido no Apêndice E. Assim, encerrou-se o vínculo do autor com a turma, pois dessa vez a aplicação da atividade ocorreu na primeira parte da disciplina, focada na fundamentação teórica. É importante ressaltar que a baixa adesão da turma na disciplina e a inconstância da turma foram fatores que dificultaram a prática da atividade proposta (poucos alunos matriculados de fato compareciam às aulas e a presença inconstante demandou mais tempo para revisar os tópicos apresentados nas aulas anteriores). Mesmo assim, os resultados foram produtivos e os três questionários respondidos forneceram subsídios essenciais à melhoria da HDIdea.

4.3.0.1 *Questionário aplicado com os alunos para avaliar as atividades*

O autor disponibilizou um questionário a respeito da aplicação do HDIdea para os alunos buscando entender a aplicabilidade do mesmo e colher *feedback* para melhorá-lo. O questionário foi disponibilizado em 19/09/2018, sendo que em três aulas durante o semestre foi lembrado sobre o preenchimento do mesmo, além de um *e-mail* enviado no fim do semestre para todos. Três estudantes autorizaram o uso dos dados para pesquisa, vide Apêndice C. Um participante respondeu o questionário no dia 24/09/2018 (participante A), portanto perto da execução da atividade, enquanto os outros 2 participantes responderam no dia 05/11/2018 (participantes B e C), próximo a entrega final do primeiro projeto da disciplina.

Todos os participantes estavam cursando pela primeira vez a disciplina de Engenharia de Requisitos e tinham idade similar (dois com 22 e um com 25 anos). O participante A não possuía familiaridade com a elicitación de requisitos de *software*, talvez por estar no começo da disciplina, enquanto os outros dois participantes afirmaram possuir, talvez por ter passado por todo o conteúdo programático da disciplina no momento da resposta. O participante A respondeu ser indiferente quanto a ter familiaridade com o desenvolvimento de *software*, enquanto os participantes B e C concordaram com a afirmação. Todos os participantes trabalharam (na Academia ou Indústria) com o desenvolvimento de *software*, o participante A desenvolveu um *plug-in* para um cliente e recebia os requisitos diretamente dele, o participante B afirmou gostar muito de desenvolver *softwares* e de já ter criado um Ambiente de Desenvolvimento Integrado (IDE) para uma linguagem interna da empresa em que trabalha, e finalmente o participante C afirmou trabalhar com desenvolvimento *back-end* e *front-end*, destacando que a visão da comunidade acadêmica em que ele está inserido por vezes despreza os trabalhos mais ligados a experiência do usuário (*front-end*).

Todos os participantes estavam presentes na atividade do primeiro passo do HDIdea, o Diagrama de Partes Interessadas estendido com os Tipos de Partes Interessadas, e concordaram que a identificação de partes interessadas ajuda no entendimento do problema, bem como o artefato utilizado ajuda a identificar partes interessadas relevantes. O participante A foi indiferente a facilidade de usar o artefato do passo 1, enquanto os outros dois acharam fácil o uso. Quanto a utilidade dos Tipos de Partes Interessadas para identificar as partes interessadas relevantes, A e B consideraram útil, enquanto C considerou indiferente. Nenhum dos participantes afirmou conhecer outra forma de identificar partes interessadas. O participante A reforçou que o artefato auxiliou a identificação de partes interessadas que poderiam não ser consideradas no projeto. Já o participante C considerou que o artefato auxiliou na execução do projeto e na escolha de quem são as partes interessadas relevantes.

Quanto a atividade de execução do segundo passo do HDIdea, as Histórias de Usuário inspiradas pelos elementos conceituais de IHD, todos os participantes estiveram em sala no dia, e concordam que as Histórias de Usuário ajudaram a caracterizar as partes interessadas e que a

caracterização das mesmas ajuda no entendimento do problema. Em relação a facilidade de usar o artefato e os elementos conceituais de IHD ajudarem na caracterização das partes interessadas, o participante B concordou, enquanto A e C foram indiferentes. Nenhum dos participantes afirmou conhecer outra forma de caracterizar as partes interessadas, nem quis emitir alguma opinião a respeito do segundo passo do HDIdea.

O terceiro passo do HDIdea foi executado em sala pelos três participantes do questionário, que consideraram que refinar o problema em diferentes níveis de abstração ajuda no entendimento do mesmo, e tanto o Ciclo de Vida dos Dados quanto a Escada Semiótica auxiliavam na visualização em diferentes níveis de abstração do problema. A facilidade de uso da Escada Semiótica só foi atestada pelo participante C, enquanto o A e B discordaram. Nenhum dos participantes afirmou conhecer outra forma de refinar o problema em contextos de Engenharia de Requisitos. O participante A considerou o terceiro passo como o mais próximo do nível técnico e esperava que seus resultados estivessem mais atrelados a atividades práticas no projeto, mas sua experiência o fez aprofundar o entendimento de uma História de Usuário sob várias perspectivas. O participante A também considerou que a dificuldade de preencher o artefato pode estar relacionado com a utilidade do mesmo, visto que o artefato auxilia a perceber funcionalidades e problemas no projeto ou História de Usuário escolhida para o preenchimento do terceiro passo do HDIdea. O participante C acredita que o artefato poderia ser separado em mais passos para facilitar algumas abstrações (talvez relacionadas a entender o funcionamento do mesmo).

Ao final do questionário, os participantes A e C avaliaram que todos os passos foram úteis para entendimento do problema e para elicitar requisitos de *software*, enquanto B considerou apenas o primeiro (Diagrama de Partes Interessadas estendido com os Tipos de Partes Interessadas) e segundo (Histórias de Usuário com elementos conceituais de IHD) como úteis, descartando o terceiro passo (Ciclo de Vida dos Dados estendido com Escada Semiótica). Todos os participantes consideraram que conhecer as perspectivas das partes interessadas no projeto ajudam no entendimento do problema e elicitação de requisitos, bem como que o HDIdea favoreceu o entendimento do problema expondo múltiplas perspectivas das partes interessadas e sendo efetivo na elicitação de requisitos no primeiro projeto da disciplina, contextualizado com a ferramenta SIMTransparência. A Tabela 4.5 expõe quais elementos conceituais foram úteis para os participantes no contexto de IHD e do SIMTransparência. O participante A considerou que é necessário muito esforço mental para utilizar o HDIdea, enquanto o B e C foram indiferentes.

Tabela 4.5: Elementos Conceituais úteis no contexto de IHD e SIMTransparência para os participantes do questionário da turma de Engenharia de Requisitos do 2º semestre de 2018.

<i>Elementos conceituais de IHD</i>	1º Participante	2º Participante	3º Participante
Agência	x		x
Autonomia			
Awareness	x	x	x
Colaboração			x
Consentimento Informado	x		
Ética	x	x	x
Legibilidade	x	x	
Negociabilidade	x		
Open Data			
Privacidade			
Relacionamentos	x		x
Transparência	x	x	

Para finalizar, o autor avaliou as atividades realizadas como positivas e o questionário, mesmo que com pouca adesão, apresentou opiniões sinceras e úteis para melhorar o HDIdea. Os participantes do questionário estão no público-alvo que considerou-se para o *framework* conceitual, que no geral, consideraram a execução do HDIdea como útil no entendimento do problema, inclusive para apoiar as atividades de elicitação de requisitos que se seguiram na disciplina. O projetista A, que respondeu o questionário poucos dias após as atividades realizadas pelo autor com a turma, foi o único que considerou o HDIdea como requerendo muito esforço mental, apesar de considerar que o esforço mental dispendido no terceiro passo justifica sua utilidade. Os projetistas B e C, que responderam o questionário após a conclusão do projeto da disciplina relacionada ao contexto de IHD e ferramenta SIMTransparência, demonstraram ter maior conhecimento da disciplina e consideraram o esforço mental dispendido no HDIdea como indiferente, o que parece apontar para uma ponderação entre o esforço requerido em outras técnicas aprendidas durante o semestre.

O primeiro passo do HDIdea (Diagrama de Partes Interessadas estendido com os Tipos de Partes Interessadas) foi considerado pelos participantes útil tanto na descoberta de possíveis partes interessadas como na sua priorização na execução do projeto. O segundo passo do HDIdea (Histórias de Usuário com elementos conceituais de IHD) foi considerado muito útil para caracterizar as partes interessadas e ajudar no entendimento do problema. Dois participantes foram indiferentes a utilidade dos elementos conceituais de IHD na caracterização das partes interessadas, o que fez o autor considerar que expor muitas categorias para as Histórias de Usuário pode não ter o efeito esperado de diversificação de pontos de vista no contexto de IHD, e sim atrapalhar devido a uma sobrecarga cognitiva no participantes da atividade. O autor decidiu por expor apenas o temas centrais da IHD no segundo passo do HDIdea enquanto o resto dos elementos conceituais de IHD são expostos por meio de perguntas a serem respondidas na criação das Histórias de Usuário. O terceiro passo do HDIdea (Ciclo de Vida dos Dados estendido com Escada Semiótica) foi citado como difícil de preencher pelos participantes A e B (e um deles considerou isso como fazendo parte da utilidade do artefato) enquanto C considerou fácil de preencher, mas também comentou que o artefato poderia ser melhor dividido entre os passos. O autor optou por focar no Ciclo de Vida dos Dados ao apresentar o artefato para dar a ideia de continuidade, direcionando os projetistas para o preenchimento em 4 passos (Ciclo de Vida dos Dados) com 6 especializações (degraus da Escada Semiótica) em cada um, evitando uma possível sobrecarga cognitiva de ser exposto a 24 campos possíveis de especializações do problema de uma só vez.

4.4 CASO DE ESTUDO DE USO E ANÁLISE DO HDIDEA APLICADO EM UM CONTEXTO DE *REDESIGN* DA FERRAMENTA MARS COM COLEGAS DE PÓS-GRADUAÇÃO DO AUTOR

O autor utilizou o contexto de *redesign* da ferramenta MaRS para analisar o uso do HDIdea com seus colegas da pós-graduação em IHC durante uma atividade com a duração de uma tarde. No dia 14/12/2018, às 13h46, o autor e 5 colegas iniciaram a atividade. O autor contextualizou os colegas com o tema da IHD, apresentando exemplos de Koesten et al. (2017), Hutton e Henderson (2017) e Cavoukian e Chibba (2016), bem como a definição adotada por Mortier et al. (2014) e seus temas centrais da IHD, mas o autor optou pela definição de IHD como "(...) (o) estud(o) (d)a manipulação humana, análise e sensemaking de dados. O volume dos dados não é relevante se a interação proporcionada amplia a capacidade humana de realizar a manipulação, análise e sensemaking dos dados". O autor também falou brevemente

sobre Mapeamento Sistemático de Literatura, visto que todos estavam executando ou já haviam executado um, além de explicar sucintamente sobre o HDIdea.

Às 2h, o autor apresentou todas as funcionalidades da ferramenta MaRS para os participantes, respondendo eventuais dúvidas; às 14h29, explicou-se brevemente o primeiro passo do HDIdea (o artefato Diagrama de Partes Interessadas com os Tipos de Parte Interessada) e, em seguida, houve a execução pelos participantes, de forma coletiva, com uso de um diagrama pré-preenchido com os dados do primeiro caso de estudo, vide Seção 4.1, folhas de papel e canetas. Os participantes discutiram a respeito das partes interessadas encontradas anteriormente e não identificaram possíveis partes interessadas significativas além das apresentadas, encerrando a execução do passo ao escolher as partes interessadas principais para o problema.

Às 15h10, o autor explanou sobre Histórias de Usuário, temas centrais da IHD de Mortier et al. (2014) e os elementos conceituais da IHD identificados por Strey et al. (2018). Os participantes então receberam as Histórias de Usuário identificadas no primeiro caso de estudo, vide Seção 4.1, e individualmente criaram Histórias de Usuário referentes a elementos conceituais da IHD (também com papel e caneta), discutindo e compartilhando o que foi feito com os colegas. Às 15h52, os participantes foram introduzidos ao Ciclo de Vida dos Dados e a Escada Semiótica pelo autor, sendo que um colega sabia como usar a Escada Semiótica e os demais não. Cada participante escolheu uma História de Usuário para usar como base no preenchimento do artefato. Após o término da atividade, os participantes entregaram o que foi criado durante o *redesign* para o autor (Apêndice F) e foram convidados a preencher um questionário sobre a atividade.

4.4.0.1 Questionário aplicado com os alunos para avaliar as atividades

As observações a seguir foram sintetizadas a partir do questionário aplicado com os pós-graduandos (Apêndice D). O questionário foi preenchido pelos colegas do autor até o dia 20/12/2018, e todos os participantes autorizaram o uso dos dados para a pesquisa. A amostra era composta por dois mestrados, com 24 anos de idade, e três doutorandos, com idades de 25, 30 e 31 anos. Apenas um dos participantes (mestrando) participou de uma atividade relacionada a IHD ou o HDIdea. Todos os participantes executaram um ou mais Mapeamentos (ou Revisões) Sistemáticas de Literatura. Dois participantes, um mestrando e um doutorando, faziam uso da ferramenta MaRS quando participaram da atividade. Também dois participantes, um mestrando e um doutorando, haviam acessado o sistema sem executar um Mapeamento (ou Revisão) Sistemática de Literatura. Um mestrando não conhecia a ferramenta MaRS, apenas tinha ouvido falar. Quatro participantes, dois mestrados e dois doutorandos, declararam ter familiaridade com o entendimento e identificação de requisitos de *software*, bem como desenvolvimento de *software*, enquanto um doutorando foi indiferente às afirmações. Apenas um mestrando declarou não ter trabalhado com algum projeto de desenvolvimento de *software*. Um doutorando trabalhou como Analista de Sistemas e como *Designer* de *Interfaces* (protótipos), sem relatar sua opinião sobre o desenvolvimento de *software*. Outro doutorando desenvolveu alguns projetos em disciplinas cursadas por ele, e considera que o desenvolvimento de *software* precisa considerar o ponto de vista de usuário na sua concepção, o que exige um processo detalhado. Ainda, outro doutorando relatou que desenvolveu um sistema *web* para gerenciamento de bolsas de Iniciação Científica, de sua concepção e desenvolvimento à implantação do mesmo e considera que o desenvolvimento de *software* tem momentos prazerosos. Por fim, um mestrando participou do desenvolvimento de *softwares* para serviços governamentais relacionados à Previdência Social e considera que o desenvolvimento de *software* exige uma tríade de conhecimento técnico do programador, conhecimento sobre os requisitos de negócio e criatividade no uso dos recursos de tempo e pessoal do projeto.

O primeiro passo do HDIdea (Diagrama de Partes Interessadas e Tipos de Partes Interessadas) foi executado com todos os participantes, apesar de um mestrando não ter participado integralmente. Todos os participantes concordaram que a identificação de partes interessadas ajudam no entendimento do problema e que o Diagrama de Partes Interessadas ajuda a identificar partes interessadas relevantes. Dois doutorandos e dois mestrandos concordaram que o Diagrama de Partes Interessadas era fácil de usar e que os Tipos de Partes Interessadas auxiliam na identificação de partes interessadas, enquanto um doutorando foi indiferente às duas afirmações. Dois mestrandos e um doutorando não conheciam outras formas de identificar partes interessadas. Um doutorando citou conhecer perfis de usuários, etapas de preparação de desenvolvimento de sistemas e Engenharia de *Software*. Outro doutorando disse que conhecia a observação do contexto em que o *software* seria desenvolvido e imersão na cultura do usuário. Em linhas gerais, todos os participantes consideraram importante a atividade de identificação de partes interessadas e seus contextos, bem como reconheceram a utilidade do primeiro passo do HDIdea. Um mestrando ressaltou a importância da retomada do artefato em etapas posteriores do *design* do *software*.

O segundo passo do HDIdea (Histórias de Usuário inspiradas pelo elementos conceituais de IHD) foi executado com todos os participantes. Todos os participantes concordaram que as Histórias de Usuário ajudam a caracterizar as partes interessadas e que a caracterização das partes interessadas ajuda no entendimento do problema. Um doutorando e dois mestrandos consideraram fácil usar as Histórias de Usuário e que os elementos conceituais da IHD ajudam a caracterizar as partes interessadas, enquanto dois doutorandos foram indiferentes as duas afirmações. Um doutorando e um mestrando não conheciam outras formas de caracterizar partes interessadas. Um doutorando conhecia o *template* para conhecimento de perfis de usuários da Engenharia Semiótica. Outro doutorando conhecia a observação e imersão no contexto no qual o *software* será projetado. De modo geral, os participantes consideraram o segundo passo do HDIdea como útil para perceber o contexto das partes interessadas em relação ao *software* projetado. Um mestrando relatou que as Histórias de Usuário aproximam a parte interessada à solução projetada de fato.

O terceiro passo do HDIdea (Ciclo de Vida dos Dados estendido com a Escada Semiótica) foi executado com todos os participantes. Um doutorando e dois mestrandos acreditam que refinar o problema em diferentes níveis de abstração auxilia o entendimento do problema, e dois doutorandos foram indiferentes à afirmação. Dois doutorandos e dois mestrandos concordaram que a Escada Semiótica ajuda a visualizar o problema em diferentes níveis de abstração, enquanto um doutorando discordou da afirmação. Apenas um doutorando e nenhum mestrando atestou ser fácil usar a Escada Semiótica, dois doutorandos e um mestrando discordaram fortemente de que a Escada Semiótica é fácil de usar. Um doutorando e um mestrando não conheciam outras formas de refinar o problema. Um doutorando conhece o artefato Quadro de Avaliação. Outro doutorando conhece a observação e documentação das observações do contexto do *software* que será projetado. Dois doutorandos e um mestrando consideraram o artefato do terceiro passo do HDIdea difícil e confuso de usar, um dos doutorandos apontou a falta de perícia para o uso efetivo do artefato. Um doutorando considerou que o passo 3 é importante para identificar os problemas e pensar em possíveis soluções. Um mestrando ponderou que o artefato pode ser útil, mas analisar os 24 pontos possíveis de análise é uma atividade complexa, que é dificultada sem uma experiência prévia de uso.

Um doutorando e um mestrando consideraram todos os três passos do HDIdea foram úteis para entender o problema e elicitare requisitos de *software*, outros um doutorando e um mestrando apontaram apenas o primeiro e segundo passos como úteis para o entendimento de problema e elicitação de requisitos de *software*, enquanto um doutorando acredita que apenas o

primeiro passo foi útil. Dois doutorandos e dois mestrandos afirmam que conhecer as perspectivas das partes interessadas é importante para o entendimento de problema e elicitação de requisitos, e um doutorando é indiferente à afirmação. Um doutorando e dois mestrandos afirmam que o HDIdea favoreceu o entendimento de problemas por perspectivas das partes interessadas e também foi efetivo ao apoiar o levantamento de requisitos da ferramenta MaRS, enquanto um doutorando discordou das duas afirmações e um mestrando foi indiferente a elas. A Tabela 4.6 demonstra quais elementos conceituais da IHD foram úteis para os participantes no contexto da atividade realizada, e um doutorando considerou também como valores Conversação e Acessibilidade. Dois doutorandos consideraram que o uso do HDIdea requer muito esforço mental, e um doutorando e dois mestrandos foram indiferentes a afirmação. Os participantes avaliaram o que foi útil para complementar o entendimento de problemas e elicitação de requisitos no contexto da ferramenta MaRS, um doutorando atribuiu a conceituação da IHD dado seu desconhecimento da área, outro doutorando acredita que faltou conteúdo explicativo e claro sobre a utilidade e forma de usar o HDIdea, ainda outro doutorando acredita que as discussões com os outros participantes durante a atividade foi interessante, e por fim um mestrando ponderou que conheceu o propósito da ferramenta MaRS e como ela será aperfeiçoada.

Tabela 4.6: Elementos Conceituais úteis no contexto de *redesign* da ferramenta MaRS para os colegas de pós-graduação do autor que participaram da atividade.

<i>Elementos conceituais de IHD</i>	1º Colega	2º Colega	3º Colega	4º Colega	5º Colega
Agência		X			
Autonomia	X	X	X	X	X
<i>Awareness</i>		X		X	X
Colaboração	X	X	X	X	X
Consentimento Informado					X
Ética	X			X	X
Legibilidade					X
Negociabilidade		X			X
<i>Open Data</i>		X			X
Privacidade	X	X		X	X
Relacionamentos	X		X	X	X
Transparência					X

A atividade de *redesign* da ferramenta MaRS foi considerada positiva pelo autor, permitindo uma visão mais crítica sobre o uso do HDIdea. Apesar da falta de conhecimento dos participantes à respeito da IHD, salvo uma exceção, todos possuíam familiaridade com Mapeamentos (ou Revisões) Sistemáticos de Literatura e conheciam a ferramenta MaRS, dois participantes inclusive fazendo uso da mesma, o que permitiu que a atividade pudesse seguir sem tanta contextualização do domínio do projeto. Um ponto que foi notado pelos participantes e autor é que executar todos os passos em uma tarde foi exaustivo e dificultou a execução do terceiro passo do HDIdea (Ciclo de Vida dos Dados estendido com Escada Semiótica) por quem não tinha experiência com ele. O autor notou que os dois participantes com experiência de desenvolvimento de *software* fora de um contexto acadêmico foram os únicos que conheciam outras formas de fazer o que os artefatos do HDIdea fazem e consideraram os três passos do HDIdea importantes no entendimento de problemas e elicitação de requisitos, o que pode indicar que é necessário uma melhor contextualização para os leigos da utilidade e importância do entendimento de problemas em desenvolvimento de *software*.

4.5 CONSIDERAÇÕES SOBRE OS CASOS DE ESTUDO

Os casos de estudo foram divididos em duas fases, a primeira auxiliando a elaboração do HDIdea, e a segunda no uso e análise do HDIdea. A primeira fase conseguiu indicar o potencial dos artefatos escolhidos para a atividade de entendimento do problema, enquanto a segunda fase atestou a capacidade do *framework* conceitual em propiciar o entendimento de problema em contextos de IHD com múltiplas partes interessadas envolvidas. Além disso, em todos os casos de estudo foram elencadas possíveis melhorias para as ferramentas usadas como contexto, contribuindo efetivamente com as partes interessadas do projetos em questão.

Nos casos de estudo de análise e uso do HDIdea, percebe-se uma resistência quanto ao uso do artefato do passo 3 visto a complexidade do mesmo (com seus 24 pontos possíveis de análise de uma só vez). Mas dada as respostas no questionário dos graduandos (pertencentes ao público-alvo do HDIdea), a utilidade do artefato parece compensar sua dificuldade ou, como o participante A indicou, a dificuldade pode ser um produto de sua utilidade. Para a versão do Manual do HDIdea (Apêndice B), optou-se por dividir o artefato nos 4 ciclos do Ciclos de Vida dos Dados, que então são estendidos pelos 6 degraus da Escada Semiótica, tornando o bloco de 24 pontos de análise em 4 blocos de 6 pontos de análise, diminuindo a carga cognitiva infligida ao projetista.

Outro ponto é que os questionários indicaram um reconhecimento da utilidade prática do HDIdea pelos participantes das atividades para fomentar o entendimento do problema. No caso dos graduandos, todos concordaram que o HDIdea foi útil para sua proposta, enquanto no caso dos pós-graduandos, os participantes com experiência em projetos fora do contexto acadêmico também reconheceram sua utilidade em relação a proposta, apesar dos participantes sem essa experiência considerarem o *framework* conceitual complexo para o uso.

5 DISCUSSÃO DOS RESULTADOS

O HDIdea está pautado em apoiar projetistas inexperientes no entendimento de problemas em contextos de IHD e, para tal, sua concepção baseou-se na facilidade de entendimento e uso do *framework*, focando na abrangência de pensamento ao se considerar a IHD, e não na completude ou profundidade de cada elemento conceitual da área. Como resultado do uso, visa-se a educação e sensibilização dos utilizadores do HDIdea com a própria IHD e com o entendimento de problema em projetos de *software*, o que pode ser observado nos casos de estudo realizados com graduandos durante o mestrado do autor.

Visto que os processos de desenvolvimento de *software* tradicionalmente tem seu foco nas camadas mais técnicas da Escada Semiótica, uma iniciativa como o *framework* em questão expõe que o entendimento dos processos e necessidades humanas ampliam a utilidade da solução para as partes interessadas no problema e reduzem o retrabalho durante as fases posteriores do desenvolvimento, aproximando as áreas de IHC e Engenharia de *Software*. Nesse sentido, a partir das atividades com os graduandos do dois estudos percebeu-se que o HDIdea diversificou as visões em relação ao problema e, segundo os relatos e respostas no questionário, o HDIdea ampliou a capacidade de:

1. Perceber a pluralidade das partes interessadas e, dentre estas, quais são importantes para o projeto – Os três participantes concordaram que identificar partes interessadas ajuda no entendimento do problema e que o HDIdea apoia a identificação de partes interessadas, enquanto também são favoráveis ou não discordam de que o artefato é de fácil utilização e é útil na identificação de partes interessadas. Um participante reforçou a utilidade do artefato na identificação de partes interessadas que poderiam não ser consideradas no projeto, enquanto outro participante considerou que o artefato auxiliou na escolha das partes interessadas relevantes para o problema;
2. Caracterizar o que as partes fazem em contextos de IHD – Os três participantes concordaram que a caracterização das partes interessadas ajuda no entendimento do problema e que as Histórias de Usuário ajudaram a caracterizar as partes interessadas. Mas também foi observado que dois dos três alunos foram indiferentes quanto a ajuda dos elementos conceituais na caracterização das partes interessadas e quanto a facilidade de uso do artefato, o que parece indicar uma quantidade excessiva de elementos conceituais demonstradas no segundo caso de estudo, resultando em uma abordagem a partir de perguntas para apresentar os elementos conceituais na versão de entrega do HDIdea;
3. Visualizar o problema em diferentes níveis de abstração – Os três participantes concordaram que refinar o problema em diferentes níveis de abstração auxilia no entendimento do mesmo e que o HDIdea ajuda a visualizar o problema em diferentes níveis de abstração. Apenas um dos participantes considerou a Escada Semiótica fácil de usar, ressaltando que o artefato poderia ser separado em mais passos para facilitar sua abstração. Outro participante citou que o preenchimento do Ciclo de Vida dos Dados estendido com a Escada Semiótica é complexo mas útil para o entendimento do problema. A dificuldade relatada com o uso do artefato é esperada, dada as múltiplas visões oferecidas para um mesmo problema, mas para a versão de entrega do HDIdea optou-se por focar primeiramente nos estágios do Ciclo de Vida dos Dados para só então detalhar os degraus da Escada Semiótica, visando reduzir a carga cognitiva aplicada nos projetistas

e ampliar a percepção em largura dos problemas ao favorecer a análise nos diferentes estágios e degraus do artefato.

Quanto aos pós-graduandos, na primeira experiência encontrou-se uma variedade de partes interessadas e suas caracterizações não identificadas previamente pelo autor da ferramenta utilizada, apontando para o potencial dos artefatos dos passos 1 e 2 quanto a identificação e entendimento em largura dos problemas do projeto em questão. Na segunda experiência o resultado foi considerado misto quanto a utilidade e facilidade de uso pelos participantes do questionário, o que pode ter sido influenciado em partes dado a aplicação extensiva do HDIdea, sem o tempo necessário para a aplicação e entendimento do último passo do *framework* (Ciclo de Vida dos Dados estendidos com Escada Semiótica). Em relação ao HDIdea, os participantes tenderam (com exceção de um dos cinco) a considerá-lo útil para a o entendimento do problema por múltiplas perspectivas, indo de encontro ao objetivo do *framework* de apoiar o projetista em uma visão em largura do projeto. Quanto aos passos individuais:

1. Diagrama de Partes Interessadas estendido com os Tipos de Partes Interessadas – Os cinco participantes concordaram que a identificação de partes interessadas ajudam no entendimento do problema e que o artefato auxilia na identificação de partes interessadas. Os participantes concordaram (ou não discordaram) também quanto a facilidade de uso do artefato. Um dos participantes considerou importante lembrar o resultado do passo 1 em fases posteriores do *design* do software.
2. Histórias de Usuário focadas em elementos conceituais de IHD – Os cinco participantes concordaram que caracterizar as partes interessadas ajuda no entendimento do problema e que o HDIdea auxilia na caracterização das partes interessadas. Os participantes concordaram (ou não discordaram) também em relação a facilidade de uso do artefato. Um participante considerou que as Histórias de Usuário aproximam as parte interessadas da solução projetada.
3. Ciclo de Vida dos Dados estendidos com Escada Semiótica – Os cinco participantes concordaram (ou não discordaram) que refinar o problema em diferentes níveis de abstração auxilia no entendimento do problema, e apenas um participante discordou de que o artefato ajuda a visualizar o problema em diferentes níveis de abstração. Apenas um participante considerou o artefato de fácil preenchimento (o que leva a crer que era o único que relatou ter experiência prévia com a Escada Semiótica durante a experiência), enquanto os outros participantes consideraram o artefato difícil de usar, citando a falta de entendimento sobre o seu funcionamento. Um dos participantes que consideraram o artefato de difícil uso citou que os 24 possíveis pontos de análise (4 estágios do Ciclo de Vida dos Dados por 6 degraus da Escada Semiótica) tornam a atividade complexa, mas robusta caso o projetista tenha uma experiência prévia de uso.

Os casos de estudo demonstraram o potencial de aplicabilidade do HDIdea por projetistas inexperientes em contextos de IHD. O primeiro caso de estudo com os graduandos apoiou-os no projeto final da disciplina, onde foram identificados diversos requisitos funcionais e não funcionais para a ferramenta utilizada como contexto, que cobriram os diversos elementos conceituais da IHD. O segundo caso de estudo com os graduandos estava inserido em uma disciplina de Engenharia de Requisitos e serviu de apoio para o primeiro projeto realizado pela turma, resultando na identificação de requisitos funcionais e não-funcionais que foram inspirados também com o auxílio dos artefatos do HDIdea, além dos artefatos de *Brainwriting* e Cenários.

Ainda considerando a aplicabilidade do HDIdea, pode-se fazer relação ao processo de *Design Thinking*, definido por Carlgren et al. (2016) como uma abordagem para a inovação (i.e.

execução de projetos criativos) com o foco em humanos baseada no pensamento e trabalho dos *designers* (e.g. projetistas no HDIdea). O processo de *Design Thinking* apoia os projetos de *software* ou não em sua ideação, experimentação e execução, fazendo uso de processos e artefatos multidisciplinares. Carlgren et al. (2016) identificaram diversas abordagens para o *Design Thinking* na literatura e propuseram cinco temas centrais para a área: (a) *User Focus*; (b) *Problem framing*; (c) *Visualization*; (d) *Experimentation*; (e) *Diversity*. Na Tabela 5.1 foram mapeados os passos do HDIdea aos temas centrais do *Design Thinking*, indicando que o *framework* conceitual consegue apoiar três do cinco temas e podendo ser incorporado ao processo de *Design Thinking* em conjunto com outros artefatos de acordo com a necessidade dos projetistas, como o Método de Inspeção Semiótica utilizado por Barreto et al. (2018) para avaliar uma solução de IHD.

Tabela 5.1: Passos do HDIdea relacionados com os Temas Centrais do *Design Thinking* de Carlgren et al. (2016).

Tema Central	<i>User Focus</i>	<i>Problem framing</i>	<i>Visualization</i>	<i>Experimentation</i>	<i>Diversity</i>
Significado	Entender o usuário e utilizar esse entendimento no restante do projeto	Entender e expandir o problema	Experiências que ampliem o entendimento do problema	Testar e experimentar soluções para o problema	Integrar visões e opiniões diversas na realização do projeto
Passos relacionados	Passos 1 e 2	Passos 1, 2 e 3	Nenhum dos passos	Nenhum dos passos	Passos 1 e 2 – se executados colaborativamente

Por fim, reconhece-se que os casos de estudo de uso e análise de uso do HDIdea tiveram poucos participantes, o que afetou a diversidade de respostas utilizadas. O último caso de estudo com os pós-graduandos foi realizado de maneira extensiva, prejudicando o entendimento e uso do terceiro passo (Ciclo de Vida dos Dados estendido com a Escada Semiótica) e dificultando uma análise comparativa em relação ao caso de estudo com os graduandos. Outro ponto é que a versão de entrega do HDIdea não foi utilizada por novos graduandos, impossibilitando uma análise de que as mudanças efetuadas no *framework* com base nas observações dos casos de estudo foram úteis para abordar os problemas identificados.

Quanto aos trabalhos futuros, vislumbra-se a experimentação da versão de entrega do HDIdea com o público-alvo. Também considera-se possibilidade de experimentação do HDIdea em contextos profissionais, permitindo uma análise fora do contexto acadêmico e com resultados que possam ser analisados posteriormente em produtos com um maior número de usuários que as ferramentas utilizadas nos casos de estudo deste trabalho. Outro possível trabalho se dá em utilizar o HDIdea na avaliação de projetos em contextos de IHD, verificando se os elementos conceituais e os diferentes níveis de abstração presentes no *framework* são cobertos pelo projeto avaliado.

6 CONCLUSÕES

O HDIdea apoia projetistas (analistas, arquitetos, desenvolvedores, engenheiros) de *software*, principalmente os não especialistas, no entendimento de problemas em contextos de IHD, permitindo identificação de requisitos pertinentes e a criação de projetos mais eficazes em solucionar os anseios das partes interessadas, enquanto também precavê potenciais problemas e indivíduos ou grupos mal-intencionados. A dissertação cobre apenas o entendimento de problema, mas se levarmos em conta que Cabitza e Locoro (2016) consideram que a IHD deve cobrir as fases de projeto, desenvolvimento e avaliação de sistemas, há espaço para estender o trabalho futuramente, ao menos no sentido de uma avaliação de sistemas mais específica.

Processos de Engenharia de *Software* ou como o *Design Thinking* são enriquecidos com o conhecimento adquirido na utilização do *framework* conceitual, melhorando a efetividade das soluções desenvolvidas e diminuindo o retrabalho ao descobrir questões e problemas antes de serem presenciados pelas partes interessadas. Ressalta-se que esses processos são, ao mesmo tempo, fundamentais para a criação bem sucedida de produtos digital e tem propriedade atestada em materializá-los, o HDIdea foca em complementar o entendimento em largura do domínio com o cerne em IHD, principalmente nas questões de natureza social e humana.

A IHD aparece na literatura como uma tentativa de abordar a problemática da interação com dados, englobando não só a visualização e o fazer sentido de dados, mas também as consequências dessa produção, como o consentimento e a privacidade das partes envolvidas ou os modos como as informações são inferidas por algoritmos a partir desses dados. A área começou a ser discutida em 2011, porém não existe um consenso sobre sua definição, nem trabalhos que agreguem ou articulem o que foi produzido até então. Buscando o sentido e formas de entender um grande volume de dados (seja considerando a limitação da cognição humana ou da capacidade das máquinas), bem como as consequências dessa produção, a IHD intenta abordar a problemática sistematicamente, e ainda passa por um processo de consolidação enquanto área de pesquisa, começando a estabelecer termos e especificidades. Muitos trabalhos se encontram em estágio embrionário, se concentrando em discussões de problemas e oportunidades, envolvendo autores de diferentes especialidades (e.g., cientistas sociais, jurídicos e da computação), porém ainda com poucos resultados de investigações.

A proposta aqui apresentada foi elaborada com base na literatura da área de IHD por meio da execução de um Mapeamento Sistemático de Literatura, ausente até então. O *framework conceitual* teve como base uma solução encontrada na literatura da área e agregado com outros artefatos que fazem sentido para a contextualização e entendimento do problema. Ainda foi criado um produto de IHD, a ferramenta MaRS, que serviu como experiência empírica para o autor em um contexto de IHD. A concepção e análise do uso do HDIdea se deu em quatro casos de estudo, realizados com graduandos (público-alvo do *framework*) e pós-graduandos, que resultaram em uma versão de entrega do HDIdea, apresentada no Apêndice B.

O presente trabalho busca preencher uma necessidade imediata de quem trabalha projetando sistemas de IHD. O HDIdea contribuiu conceitual e construtivamente para a área, apoiando casos reais de entendimento de problemas em contextos de IHD, explorando problemas e técnicas da IHC em um processo caracteristicamente da Engenharia de Software. Além disso, o MSL realizado auxilia pesquisadores e partes interessadas que desejam conhecer a área, agregando os trabalhos realizados até então, e auxiliando a disseminação da IHD. Por fim, a criação da ferramenta MaRS contribui com a execução de MSLs por quem desejar hospedá-la, também metodologicamente ao promover a transparência dos MSLs realizados por meio dela.

REFERÊNCIAS

- Ball, A. (2012). *Review of Data Management Lifecycle Models*. University of Bath, UK United Kingdom.
- Barreto, P., Salgado, L. e Viterbo, J. (2018). Transparency communication strategies in human-data interaction. Em *Proceedings of the XIV Brazilian Symposium on Information Systems, SBSI'18*, páginas 66:1–66:8, New York, NY, USA. ACM.
- Brasil (2009). Lei da transparência. <http://www.leidatransparencia.cnm.org.br/>. Acessado em 14/07/2019.
- Cabitza, F. e Locoro, A. (2016). Human-data interaction in healthcare: Acknowledging use-related chasms to design for better health information. *ArXiv*, abs/1602.05751.
- Carlgren, L., Rauth, I. e Elmquist, M. (2016). Framing design thinking: The concept in idea and enactment. *Creativity and Innovation Management*, 25(1):38–57.
- Cavoukian, A. e Chibba, M. (2016). *Cognitive Cities, Big Data and Citizen Participation: The Essentials of Privacy and Security*, páginas 61–82. Springer International Publishing, Cham.
- Chamberlain, A. e Crabtree, A. (2016). Searching for music: understanding the discovery, acquisition, processing and organization of music in a domestic setting for design. *Personal and Ubiquitous Computing*, 20(4):559–571.
- Cohn, M. (2004). *User Stories Applied: For Agile Software Development*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA.
- Cohn, M. (2008). Advantages of the “as a user, i want” user story template. *Mountaingoat Software*. Retrieved on July 2018, 15:2012.
- Demchenko, Y., Grosso, P., de Laat, C. e Membrey, P. (2013). Addressing big data issues in scientific data infrastructure. Em *2013 International Conference on Collaboration Technologies and Systems (CTS)*, páginas 48–55.
- Elmqvist, N. (2011). Embodied human-data interaction. Em *ACM CHI 2011 Workshop “Embodied Interaction: Theory and Practice in HCI*, páginas 104–107.
- Floridi, L. (2003). Two approaches to the philosophy of information. *Minds and Machines*, 13(4):459–469.
- Freitas, V. (2014). Parsifal. <https://vitorfs.com/projects/parsifal/>. Acessado em 21/07/2019.
- Golfarelli, M. e Rizzi, S. (2009). *Data warehouse design: Modern principles and methodologies*, volume 5. McGraw-Hill New York.
- Haddadi, H., Mortier, R., McAuley, D. e Crowcroft, J. (2013). Human-data interaction. Relatório técnico, University of Cambridge, Computer Laboratory.

- Hornung, H., Pereira, R., Baranauskas, M. C. C. e Liu, K. (2015). Challenges for human-data interaction – a semiotic perspective. Em Kurosu, M., editor, *Human-Computer Interaction: Design and Evaluation*, páginas 37–48, Cham. Springer International Publishing.
- Hutton, L. e Henderson, T. (2017). *Beyond the EULA: Improving Consent for Data Mining*, páginas 147–167. Springer International Publishing, Cham.
- Jeffries, R. (2001). Essential xp: Card, conversation, confirmation. <http://ronjeffries.com/xprog/articles/expcardconversationconfirmation/>. Acessado em 14/07/2019.
- Kee, K. F., Browning, L. D. e Ballard, D. I. (2012). Sociomaterial processes, long term planning, and infrastructure funding: Towards effective collaboration and collaboration tools for visual and data analytics.
- Koesten, L. M., Kacprzak, E., Tennison, J. F. A. e Simperl, E. (2017). The trials and tribulations of working with structured data: -a study on information seeking behaviour. Em *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, páginas 1277–1289, New York, NY, USA. ACM.
- Kolkman, M. (1993). *Problem Articulation Methodology*. Tese de doutorado, University of Twente.
- Leone, L. (2017). Beyond connectivity: The internet of food architecture between ethics and the eu citizenry. *Journal of Agricultural and Environmental Ethics*, 30(3):423–438.
- Liu, K. (2000). *Semiotics in information systems engineering*. Cambridge University Press.
- Locoro, A. (2015). A map is worth a thousand data: Requirements in tertiary human-data interaction to foster participation. Em *Proceedings of the Third edition of the International Workshop on Cultures of Participation in the Digital Age: Coping with Information, Participation, and Collaboration Overload co-located with the Fifth International Symposium on End-User Development (IS-EUD 2015), Madrid, Spain, May 26, 2015.*, páginas 39–44.
- Mashhadi, A., Kawsar, F. e Acer, U. G. (2014). Human data interaction in iot: The ownership aspect. Em *2014 IEEE World Forum on Internet of Things (WF-IoT)*, páginas 159–162.
- Mishra, S. e Cafaro, F. (2018). Full body interaction beyond fun: Engaging museum visitors in human-data interaction. Em *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '18, páginas 313–319, New York, NY, USA. ACM.
- Morris, C. W. (1938). Foundations of the theory of signs. Em *International encyclopedia of unified science*, páginas 1–59. Chicago University Press.
- Mortier, R., Haddadi, H., Henderson, T., McAuley, D. e Crowcroft, J. (2014). Human-data interaction: the human face of the data-driven society.
- Nunes, N., Ribeiro, M., Prandi, C. e Nisi, V. (2017). Beanstalk: A community based passive wi-fi tracking system for analysing tourism dynamics. Em *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, EICS '17, páginas 93–98, New York, NY, USA. ACM.

- Oulasvirta, A. e Hornbæk, K. (2016). Hci research as problem-solving. Em *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, páginas 4956–4967, New York, NY, USA. ACM.
- Pacheco, D., Wierenga, S., Omedas, P., Oliva, L. S., Wilbricht, S., Billib, S., Knoch, H. e Verschure, P. F. (2015). A location-based augmented reality system for the spatial interaction with historical datasets. Em *2015 Digital Heritage*, volume 1, páginas 393–396.
- Pereira, R. e Baranauskas, M. C. C. (2015). A value-oriented and culturally informed approach to the design of interactive systems. *International Journal of Human-Computer Studies*, 80:66 – 82.
- Pereira, R., Baranauskas, M. C. C. e da Silva, S. R. P. (2013). Social software and educational technology: Informal, formal and technical values. *Journal of Educational Technology Society*, 16(1):4–14.
- Petersen, K., Vakkalanka, S. e Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64:1 – 18.
- Roberts, J. e Lyons, L. (2017). The value of learning talk: applying a novel dialogue scoring method to inform interaction design in an open-ended, embodied museum exhibit. *International Journal of Computer-Supported Collaborative Learning*, 12(4):343–376.
- Skoutas, D. e Simitsis, A. (2007). Ontology-based conceptual design of etl processes for both structured and semi-structured data. *Int. J. Semantic Web Inf. Syst.*, 3:1–24.
- Stamper, R. (1973). *Information in business and administrative systems*. John Wiley & Sons, Inc.
- Stamper, R. (1993). A semiotic theory of information and information systems. Em Randell, B., editor, *Proceedings of the Joint ICL/University of Newcastle Seminar on the Teaching of Computer Science, Part IX: Information*, páginas 1–33. University of Newcastle.
- Strey, M. R., Pereira, R. e de Castro Salgado, L. C. (2018). Human data-interaction: A systematic mapping. Em *Proceedings of the 17th Brazilian Symposium on Human Factors in Computing Systems*, IHC 2018, páginas 27:1–27:12, New York, NY, USA. ACM.
- Widjojo, E. A., Chinthammit, W. e Engelke, U. (2017). Virtual reality-based human-data interaction. Em *2017 International Symposium on Big Data Visual Analytics (BDVA)*, páginas 1–6.
- Wilke, G. e Portmann, E. (2016). Granular computing as a basis of human–data interaction: a cognitive cities use case. *Granular Computing*, 1(3):181–197.
- Zamboni, A., Di Thommazo, A., Hernandez, E. e Fabbri, S. (2010). Start uma ferramenta computacional de apoio à revisão sistemática. Em *Proc.: Congresso Brasileiro de Software (CBSoft'10)*, Salvador, Brazil.

**APÊNDICE A – MAPEAMENTO SISTEMÁTICO DE LITERATURA SOBRE O
ESTADO DA ARTE DA INTERAÇÃO HUMANO-DADOS**

Segue o artigo completo aceito no XVII Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (IHC 2018).

Human Data-Interaction: A Systematic Mapping

Mateus Rambo Strey
Universidade Federal do Paraná
Curitiba, Paraná, Brasil
mateus.rstrey@gmail.com

Roberto Pereira
Universidade Federal do Paraná
Curitiba, Paraná, Brasil
rpereira@inf.ufpr.br

Luciana C. de Castro Salgado
Universidade Federal Fluminense
Niterói, Rio de Janeiro, Brasil
luciana@ic.uff.br

ABSTRACT

Big Data, e-Science and Internet of Things have contributed to increase the production, processing and storage of data, changing the way people deal and live with data. Although the problem is not new, the "human aspect" of data and the possible impact of Human-Data Interaction (HDI) in human life have been explored and discussed as an emerging research area. On the one hand, HDI offers plenty of opportunities for research and development, and on the other hand it demands characterization, grounding, critical discussions, empirical results and thinking tools to support research and practice. This paper presents a Systematic Mapping of Literature on HDI in Computer Science, identifying the different definitions for the area, elements or objects of investigation, contexts of application, stakeholders, etc. Based on 28 selected papers, results point out to a lack of definition or agreement on what HDI is, but suggest that there are different aspects that can characterize it, and allow identifying concerns and objects of study, such as privacy, ownership and transparency. Results suggest a demand for theoretical and methodological frameworks to support the understanding, design and evaluation of HDI via computing systems.

CCS CONCEPTS

• **Human-centered computing** → **Interaction design theory, concepts and paradigms**; *HCI theory, concepts and models*;

KEYWORDS

Human-Data Interaction, Systematic Mapping Review, Human-Computer Interaction

ACM Reference Format:

Mateus Rambo Strey, Roberto Pereira, and Luciana C. de Castro Salgado. 2018. Human Data-Interaction: A Systematic Mapping. In *17th Brazilian Symposium on Human Factors in Computing Systems (IHC 2018), October 22–26, 2018, Belém, Brazil*. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3274192.3274219>

1 INTRODUÇÃO

Com a disseminação das tecnologias computacionais interativas para todos os aspectos da vida humana, estamos cada vez mais conectados por dispositivos eletrônicos e produzindo dados sobre

toda e qualquer atividade possível. Há uma demanda crescente por pesquisas e soluções que apoiem e possibilitem o uso desses dados, por diferentes partes interessadas, em diferentes contextos, finalidades, recursos, bem como garantam o direito a explicação sobre como e para que fins estes dados são ou serão usados [18][21]. Também existem iniciativas que discutem como incorporar dados no processo de *design* e como projetar soluções que priorizem uma interação efetiva entre humanos e dados em tecnologias de *Internet of Things* [64]. A área de Interação Humano-Dados (IHD) investiga a problemática em suas diferentes facetas, englobando não só a visualização e o fazer sentido do grande volume de dados, mas também as consequências dessa produção, como o consentimento e a privacidade das partes envolvidas ou o modo como as informações são inferidas por algoritmos a partir desses dados [45].

Em 2011, o trabalho de Elmqvist [23] iniciou discussões em IHD, mas embora o termo tenha se popularizado e uma linha de pesquisa tenha se delineado sobre o tópico, ainda não existe um consenso sobre uma definição para IHD, nem trabalhos que agreguem o que foi produzido até então. Se por um lado IHD enquanto área emergente oferece oportunidades de pesquisa e desenvolvimento, por outro lado há uma demanda por contribuições teóricas e metodológicas que ajudem a entender e caracterizar IHD como área de pesquisa, e por ferramentas que apoiem o entendimento, projeto e avaliação de soluções computacionais para apoiar a IHD.

Este artigo apresenta um Mapeamento Sistemático de Literatura (MSL) para identificar o estado da arte dos trabalhos em IHD que apresentam alguma solução conceitual ou prática para a área. Após a execução do processo de mapeamento, foram selecionados 28 trabalhos que foram utilizados como base para a extração de dados de modo a responder 8 perguntas de pesquisa. O artigo apresenta o processo de mapeamento, os dados extraídos dos artigos selecionados, e as discussões sobre as perguntas de pesquisa respondidas com base nos dados extraídos. Como principal resultado, o artigo oferece um panorama sobre a literatura em IHD, destacando desafios e oportunidades de pesquisa.

2 MAPEAMENTO SISTEMÁTICO

O mapeamento sistemático foi conduzido seguindo as diretrizes de Petersen *et al.* [52] para identificar o estado da arte dos trabalhos em IHD que efetivamente apresentem e discutam soluções teóricas ou práticas para avançar na área. As etapas planejadas e executadas estão descritas a seguir.

2.1 Protocolo de Busca

Inicialmente, buscou-se por artigos completos e capítulos de livros relacionados à área de IHD em 4 bases de dados, sendo estas: *ACM Digital Library*, *IEEE Xplorer*, *Springer Link* e *Science Direct*. Durante a calibração da expressão de busca foi identificado que poucos artigos citavam explicitamente IHD, e decidiu-se considerar artigos

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IHC 2018, October 22–26, 2018, Belém, Brazil

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6601-4/18/10...\$15.00

<https://doi.org/10.1145/3274192.3274219>

científicos que tratassem do tema independentemente do tamanho ou formato (e.g. artigos completos, resumos estendidos, trabalhos em andamento).

O mapeamento sistemático de literatura foi executado da seguinte forma:

- (1) Elaboração do plano de estudos: definiu-se os objetivos e passos que seriam conduzidos durante o mapeamento. O protocolo foi elaborado pelo primeiro autor, sendo discutido e revisado com o segundo autor do artigo, especificando:
 - Objetivo da pesquisa.
 - Perguntas de pesquisa a serem respondidas a partir do mapeamento.
 - Bases de dados utilizadas.
 - Palavras-chave utilizadas na busca de trabalhos.
 - Critérios de inclusão e exclusão dos trabalhos encontrados.
 - Perguntas para extração de dados dos trabalhos selecionados.
- (2) Busca dos trabalhos nas bases de dados (trabalhos até 17/04/2018);
- (3) Aplicação dos critérios de inclusão e exclusão dos trabalhos encontrados;
- (4) Atualização dos critérios de inclusão e exclusão devido à baixa quantidade de trabalhos selecionados;
- (5) Extração de dados dos trabalhos por meio das perguntas de extração;
- (6) Adição de novos trabalhos com a técnica de *snowballing* nos trabalhos selecionados;
- (7) Aplicação dos critérios de inclusão e exclusão aos novos trabalhos, bem como a extração de dados mediante perguntas de extração;
- (8) Análise dos dados (perguntas de extração) para responder as perguntas da pesquisa.

Tendo por objetivo mapear sistematicamente o estado da arte na área de IHD e identificar como ela tem sido entendida e investigada, foram definidas as seguintes perguntas de pesquisa (PP):

- PP1: O que os trabalhos mapeados consideram como IHD?
- PP2: Quais são os países, instituições e autores que publicam sobre IHD?
- PP3: Em quais fóruns científicos esses trabalhos têm sido publicados?
- PP4: Que tipos de propostas os trabalhos apresentam?
- PP5: Quem são os *stakeholders* alvo das propostas?
- PP6: Como as propostas são avaliadas?
- PP7: Quais bases teóricas fundamentam as propostas mapeadas?
- PP8: Quais elementos conceituais básicos para IHD foram apresentados nos trabalhos?

A busca foi conduzida nas 4 bases selecionadas utilizando apenas a palavra-chave "*Human-Data Interaction*", com e sem o hífen, nos campos de busca Título, Palavras-chave e *Abstract*. A pesquisa realizada sem a limitação pela palavra-chave inteira resultou em mais de 100.000 resultados de diversas áreas, principalmente na saúde, devido aos termos "*human*" e "*data*". Assim, a pesquisa ficou restrita ao termo "*Human-Data Interaction*" ou "*Human Data Interaction*" nos campos mencionados. A Tabela 1 apresenta o link para execução

das expressões de busca por base e indica a quantidade de 57 trabalhos que foram recuperados até 17/04/2018, quando ocorreu a última busca por trabalhos.

Tabela 1: Bases de dados com a URL de busca e a quantidade de trabalhos encontrados.

Bases de dados e Trabalhos encontrados		
Nome	URL	#
ACM	https://dl.acm.org/results.cfm?query=%22human%20data%20interaction%22%20%7C%7C%20%22human-data%20interaction%22	9
IEEE	https://ieeexplore.ieee.org/search/searchresult.jsp?queryText=.LB..QT.human-data%20interaction.QT.%20OR%20.QT.human%20data%20interaction.QT..RB.&newsearch=true	9
Springer	https://link.springer.com/search?query=%22human+data+interaction%22+%7C%7C+%22human-data+interaction%22	31
SD	https://www.sciencedirect.com/search?qs=%22human+data+interaction%22+%7C%7C+%22human-data+interaction%22&origin=home&zone=qSearch	8
Total		57

A busca considerando apenas a palavra-chave representante da área foi conduzida de modo a obter a maior quantidade possível de trabalhos que explicitamente mencionassem a área. Entretanto, trabalhos que não apresentassem nenhuma proposta e que não tivessem IHD como foco, não estavam no escopo deste mapeamento. Assim, os seguintes critérios de exclusão foram definidos:

- (1) Trabalhos não disponíveis para acesso;
- (2) Trabalhos que não estivessem em Inglês;
- (3) Trabalhos que eram duplicatas de trabalhos já selecionados (apresentavam/discutiam um mesmo projeto ou estratégia), sendo mantido o trabalho completo ou mais recente;
- (4) Tabela de conteúdo de simpósio/conferência/etc;
- (5) Trabalhos não relacionados ao contexto de IHD;
- (6) Trabalhos que não apresentassem solução prática ou conceitual para o contexto de IHD, ou que referenciavam IHD sem abordá-la de fato (e.g. alguns trabalhos em Big Data colocam IHD como uma preocupação de *Data Quality*, sem explorar a problemática).

Ao aplicar os critérios de exclusão (Tabela 2), dos 57 trabalhos que inicialmente foram recuperados, apenas 20 foram incluídos para a extração de dados e análise, sendo que dos 37 trabalhos excluídos, 22 trabalhos foram excluídos na aplicação do critério de exclusão 6 por não apresentarem ou discutirem nenhuma proposta de solução em IHD.

Durante o processo de leitura dos artigos selecionados foi identificado que alguns trabalhos-chave não apareceram no mapeamento e, ao investigar o porquê, constatou-se que estes trabalhos faltantes não foram publicados em periódicos tradicionais, mas sim em publicações livres ou apenas nas páginas pessoais (ou institucionais) dos autores. Desta maneira, iniciou-se um processo de seleção de trabalhos por meio da técnica *snowballing*. Foram encontrados 10

trabalhos, 8 dos quais foram selecionados aplicando os mesmos critérios de inclusão e exclusão, totalizando 28 trabalhos selecionados (Tabela 2).

Tabela 2: Trabalhos selecionados por bases de dados após os critérios de exclusão. Os critérios de exclusão estão ordenados por ordem de aplicação.

Trabalhos Encontrados	Bases de dados e Trabalhos selecionados					
	ACM	IEEE	Springer	SD	Outras	Total
	9	9	31	8	10	67
<i>Critério de Exclusão</i>						
1. Acesso não permitido	0	0	0	1	0	1
2. Não está em inglês	1	0	1	0	0	2
4. Tabela de conteúdo	0	1	0	0	0	1
5. Não relacionado a IHD	0	2	7	1	0	10
3. Duplicatas	1	0	0	0	1	2
6. Sem modelo ou solução	2	0	14	6	1	23
Total	5	6	9	0	8	28

2.2 Extração, Análise e Classificação de dados, e Validade do Estudo

Para extrair os dados dos 28 trabalhos, foi necessário desenvolver o modelo de extração descrito na Tabela 3, no qual cada item especificado possui uma descrição do que ele deveria conter (seu valor). Uma instância do modelo foi criada e preenchida para cada trabalho selecionado.

Petersen e Gencel [52] sugerem 4 tipos de validade em estudos de mapeamento: (1) validade descritiva; (2) validade teórica; (3) generalização; e (4) validade interpretativa.

A validade descritiva (1) refere-se à extensão de como as observações são descritas com precisão e objetividade. No nosso estudo, os dados foram extraídos com base no modelo de extração da Tabela 3 e puderam ser ajustados conforme a demanda, então consideramos essa ameaça sob controle.

A validade teórica (2) refere-se a habilidade do estudo ser capaz de capturar o que o estudo de mapeamento se propõe a capturar, sendo crítico escolher os tópicos que cobrem o que se deseja e onde os trabalhos são procurados. Nesse estudo, selecionamos uma única palavra-chave, ("Interação Humano-Dados" ou "Interação Humano Dados"), pois o objetivo é mapear o estado da arte da área na Ciência da Computação, e elegemos 4 bases de dados relevantes. Considerando que muitos trabalhos não estavam catalogados em bases populares ou nem haviam sido publicados em veículos indexados, decidimos utilizar a técnica de *snowballing* para encontrar trabalhos faltantes, reduzindo o risco de não mapear trabalhos relevantes. O

Tabela 3: Template de dados extraídos dos trabalhos selecionados.

Item	Valor	PPs
Identificação do Estudo	Código único para cada trabalho	
Título	Título do trabalho	
Ano de publicação	Ano de publicação	
Informações do(s) Autor(es)	Nome dos autores, país de origem e instituição	PP2
Domínio	Domínio do repositório do trabalho	PP3
Veículo de publicação	Nome da publicação	PP3
URL	Endereço digital do trabalho	
Definição do conceito de IHD	Como o trabalho define IHD	PP1
Fundamentação teórica	Qual o embasamento teórico da proposta	PP7
Proposta	Se é um framework, interface, guidelines, etc	PP4
Público-alvo da proposta	Se a proposta visa o projetista, designer, cliente, etc	PP4, PP5
Avaliação da proposta	Como a proposta foi avaliada: provas de conceito; uso em contexto real; etc.	PP6
Métricas de avaliação	Métricas de classificação na avaliação da proposta	PP6
Resultados	Resultados obtidos e como foram avaliados	PP6
Elementos conceituais	Possíveis elementos conceituais de IHD	PP8

protocolo de busca foi definido pelo primeiro autor e revisado pelo segundo autor, reduzindo o risco de viés do pesquisador.

A generalização (3) é garantida por meio da formalidade e documentação do protocolo de busca, que pode ser expandida para considerar outras bases de dados, e também perguntas de pesquisa que podem ser respondidas com base no modelo de extração de dados adotado.

A validade interpretativa (4) depende das conclusões do mapeamento serem razoáveis com base nos dados apresentados. O pesquisador pode introduzir um viés no desenvolvimento do protocolo de pesquisa, ou na elaboração e extração dos dados dos trabalhos. A ameaça foi reduzida pela definição do processo, padronização da extração de dados, desenvolvimento das respostas com base nos dados extraídos e revisão com o segundo e terceiro autores.

A repetibilidade do mapeamento sistemático é um aspecto fundamental para a sua qualidade e validade. Esta é garantida pelo detalhamento do processo seguido pelos autores, incluindo as *strings* de busca (disponíveis na Tabela 1) e o modelo com os dados extraídos.

Os dados brutos extraídos encontram-se no link¹ apresentado no rodapé.

3 RESULTADOS E DISCUSSÃO

A área de IHD tem ganhado visibilidade no meio científico e é possível perceber um aumento contínuo no número de publicações selecionadas por ano na Figura 1. Além dos 28 trabalhos selecionados, 23 outros trabalhos também citam IHD mas foram excluídos do mapeamento por não proporem uma solução conceitual ou prática para um problema de IHD (critério de exclusão 6 na Tabela 2).

O que se percebe é a inexistência de um consenso em relação a seu significado, fato que motivou a pergunta *o que os trabalhos mapeados consideram como IHD?* (PP1). Para respondê-la, identificamos e nomeamos 4 categorias (trataremos como "visões") que os autores consideram, seja por defini-las ou usar referências da visão, sendo elas: (a) *Embodied*, (b) *Colaborativa*, (c) *Ampla*, e (d) *Visualização*. Na Tabela 4 estão sumarizadas as visões e, respectivamente, uma definição genérica de IHD e os trabalhos selecionados por base de dados. Os trabalhos que estendem a definição referenciada estão em negritos. Alguns trabalhos ([49], [48], [13], [32] e [56]) citam o termo IHD sem defini-lo, mas apresentam contribuições no contexto da área.



Figura 1: Artigos por ano de publicação.

A visão (a) *Embodied* foi a primeira a aparecer na literatura. Elmqvist [23] definiu IHD como "a manipulação humana, análise e *sensemaking* de dados", sendo suas caracterizadas por larga escala, dados não estruturados, múltiplas fontes e relações complexas. Essa definição, completa ou em partes, é utilizada também em trabalhos de outras visões. Cafaro [12] considera que a IHD lida com o problema de "obter dados personalizados, informados pelo contexto, e compreensíveis de grandes conjuntos de dados". Em relação ao contexto *embodied*, Elmqvist propõe uma IHD *embodied* na qual o processo de análise dos dados é apoiado por auxílios visuais/tangíveis/cognitivos, enquanto Cafaro argumenta que a IHD por si só é *embodied*, pois as pessoas desejam explorar os dados, para além do processo de análise que requer uma ideia ou objetivo prévio.

A visão (b) *Colaborativa* foi definida por Kee *et al.* [32], que considera a "Ciência da interação" (*Science of interaction*) no contexto

de sistemas de dados e análise visual. Na Ciência da interação estão envolvidas a comunicação humano-humano (colaboração sobre dados) e a IHD (colaboração entre dados). O autor não aprofunda a definição de IHD, mas relaciona a área ao desenvolvimento de ferramentas colaborativas que lidem com o processo de tomada de decisões, e.g. na concessão de fundos para projetos científicos. Nenhum outro trabalho usa essa visão especificamente, entretanto o elemento básico *colaboração* aparece em boa parte dos trabalhos mapeados (ver PP8).

A visão (c) *Ampla*, mais popular entre os trabalhos mapeados, foi definida por Haddadi *et al.* como a interação entre humanos e a análise de grandes conjuntos de dados pessoais, combinando tanto os dados quanto os algoritmos usados para analisá-los [29] (versão completa de [44]). A interação em IHD difere da interação em IHC por lidar com cenários mais passivos nos quais as pessoas envolvidas utilizam uma infraestrutura e desconhecem ou ignoram seu funcionamento, bem como na escala dos sistemas, e.g. armazenamento e análise de dados de populações inteiras em termos de *petabytes* ou *quintillion-bytes*. No ano seguinte, Mortier *et al.* [45] (autores da definição original e novos colegas) ampliam a definição proposta de IHD ao entender que dados não necessariamente são pessoais, interpretar os dados como *boundary objects* [59] [36] - as várias formas que múltiplas comunidades se referem e pensam sobre dados, e ao incluir 3 temas centrais: (i) *Legibilidade*, (ii) *Agência* e (iii) *Negociabilidade*.

Os temas são definidos da seguinte forma: (i) *Legibilidade* refere-se a tornar os dados e algoritmos analíticos transparentes e compreensíveis para as pessoas, explicitando quais dados eles estão consumindo ou usando para fazer inferências. Em geral, isso está em conflito direto com o fato de que esses algoritmos representam a propriedade intelectual essencial das empresas que os implementam e executam e, portanto, não podem ser facilmente tornados públicos; (ii) *Agência* refere-se a dar às pessoas a capacidade de agir dentro desses sistemas de dados (de optar ou não) por controlar, informar e corrigir dados e inferências sobre si e assim por diante. Nem todos os usuários querem exercer essa capacidade de forma contínua, mas alguns entusiastas podem e muitas pessoas irão de tempos em tempos fazê-lo, seja por curiosidade ou por algum potencial perigo; (iii) *Negociabilidade* refere-se às múltiplas relações dinâmicas que surgem em torno dos dados e processamento de dados. Esse tema engloba, por exemplo, como a compreensão e as atitudes individuais mudam com o tempo, como a sociedade responde a esses problemas por meio da formação de normas sociais em torno dos dados e do processamento de dados, as diferentes estruturas legais e regulatórias que podem ser aplicadas em diferentes jurisdições a diferentes tipos de dados e tipos de processamento de dados - e as maneiras pelas quais os dados são frequentemente ambíguos em seu assunto e seu significado.

Ainda na visão *Ampla*, outros autores estenderam a definição original. Mashhadi *et al.* [41] consideram que a IHD deve providenciar acesso e entendimento dos dados sobre indivíduos, e como esses dados os afetam. Hornung *et al.* [30] consideram que a IHD deve lidar com dados que afetam pessoas (clarificando o conceito de *personal data* [29]) e incluir as partes interessadas no ciclo de vida dos dados, e.g. geração de dados, coleta e processamento. Mortier *et al.* [46] publicaram um capítulo de enciclopédia, com a co-autoria

¹https://docs.google.com/spreadsheets/d/1VO_Ml2soMePxFNPAmlyD3ssdqP5QdrblvsAiN7GH0fM

Tabela 4: Trabalhos selecionados por visão, com uma descrição genérica da definição de IHD.

		Trabalhos, Visões e Definições de IHD
Visão	Definição Genérica	Trabalhos
Embodied	"A manipulação humana, análise e <i>sensemaking</i> de dados" e "obter dados personalizados, informados pelo contexto, e compreensíveis de grandes conjuntos de dados" com auxílios visuais/táteis/cognitivos	Springer: [56] Outras: [23] IEEE: [49] ACM: [12], [43], [13]
Ampla	A IHD é interação entre humanos e a análise de grandes conjuntos de dados, combinando tanto os dados quanto os algoritmos usados para analisá-los. Possui 3 temas centrais: Legibilidade, Agência e Negociabilidade.	Springer: [15], [25], [38], [16], [20], [31], [63], [30] Outras: [45], [46], [11], [39], [29], [53] IEEE: [41], [42], [19], [17] ACM: [48], [34]
Colaborativa	A IHD, juntamente com a comunicação humano-humano, compõe a "Ciência da interação" no contexto de sistemas de dados e análise visual	Outras: [32]
Visualização	A IHD é análoga à IHC, sendo a interface direta entre um ser humano e uma representação visual de dados.	IEEE: [62]

do sociólogo A. Crabtree, no qual discutem temas relacionados à IHD como *Articulation Work* e Prestação de contas (*Accountability*).

Locoro [39] e Cabitza [11] propõem uma distinção tripartida dos dados na qual: *dados primários* são produzidos por e para uma prática, portanto, não são facilmente destacáveis de seu contexto de criação, manipulação e uso; *dados secundários* englobam os processos pelos quais o pessoal administrativo, gerencial e de pesquisa dão sentido aos dados primários, muitas vezes processando-os também para torná-los mais adequados à sua interpretação profissional; *dados terciários* englobam atividades que dizem respeito, em grande parte, ao uso de dados de categorias sociais de pessoas, como formuladores de políticas, usuários finais, contribuintes e cidadãos. A maioria desses usuários de dados não compartilha os mesmos propósitos e as mesmas capacidades de outras categorias profissionais em lidar com "seus dados". Cabitza [11] considera ainda que a IHD deve cobrir 3 fases: projeto, desenvolvimento e avaliação de sistemas que extraem informações e apoiam o conhecimento em domínios de aplicação com dados intensos.

Por fim, a visão (d) *Visualização* de IHD foi definida por Widjojo *et al.* [62] como análoga à IHC, sendo a interface direta entre um ser humano e uma representação visual de dados. Desta maneira, abrange a definição e o *design* de todos os aspectos relacionados diretamente à manipulação de dados, incluindo atributos de dados, *design* e representação visual, e técnicas de interação. A IHD surge da necessidade de envolver os humanos com grandes conjuntos de dados e fornecer uma experiência de usuário altamente envolvente na exploração de dados - permitindo a criação de sentido de grandes e complexos conjuntos de dados. A área de pesquisa é inerentemente interdisciplinar e encapsula a visualização de dados, o *design* da interface do usuário e a interação, bem como a psicologia, a ciência comportamental e a cognição humana. Os autores sugerem que no caso de conjuntos de dados espaciais, como a maioria dos conjuntos de dados científicos, a percepção espacial humana também poderia ajudar a melhorar a capacidade cognitiva, encorajando o uso de tecnologias de Realidade Aumentada (RA) e Realidade Virtual (RV).

Para responder *quais são os países, instituições e autores que publicam sobre IHD?* (PP2), resumimos na Tabela 5 os países e suas respectivas instituições que possuem trabalhos na área. Percebemos que trabalhos não relacionados entre si de diferentes instituições no mesmo país existem apenas no Reino Unido (com 9 instituições) e nos Estados Unidos (com 5 instituições), pois outras instituições no mesmo país com mais de uma instituição listada (Portugal, Itália, Austrália, Finlândia, Alemanha e Suíça) pertencem aos mesmos trabalhos. Os autores que publicaram mais de 1 trabalho estão dispostos na Tabela 6, demonstrando que o maior número de publicações advém de autores localizados no Reino Unido (A. Crabtree, R. Mortier, T. Henderson, H. Haddadi, D. McAuley, J. Crowcroft, idealizadores da visão *Ampla* de IHD e com maior influência dos trabalhos até agora) e Estados Unidos (F. Cafaro, L. Lyons, J. Roberts, proponentes da visão *Embodied*). Todos os autores mapeados com mais de 1 trabalho mantiveram suas visões entre eles.

Ao mapear a literatura em um determinado tema é importante identificar quais os veículos nos quais os trabalhos têm encontrado espaço. Para responder em *quais fóruns científicos os trabalhos têm sido publicados?* (PP3), criamos a Tabela 7, que apresenta os trabalhos publicados em eventos e afins, e a Tabela 8 que apresenta os trabalhos publicados em periódicos e outras publicações não relacionadas a eventos. Juntas, estas tabelas demonstram uma importante relação: enquanto na Inglaterra e nos Estados Unidos há diversidade e quantidade relevante de publicações (pelas diferentes instituições de pesquisa), outros países, como Portugal, por mais que tenha 2 universidades que publiquem em IHD, tais instituições estão atreladas a um mesmo trabalho, indicando que a pesquisa no tema ainda está em estágio inicial de desenvolvimento.

Respondendo *quais tipos de proposta os trabalhos apresentam?* (PP4), 15 trabalhos foram agrupados por 4 tipos de propostas diferentes: 1. Boas Práticas, Diretrizes e Recomendações de *design*; 2. *Frameworks*; 3. Métodos; 4. *Workflows*. Os trabalhos que oferecem uma proposta de definição da área e possíveis desafios foram tratados na PP1 ([23][32][29][39][11][45][46][62]), com uma exceção [12]. Outros 5 trabalhos são tratados abaixo.

Tabela 5: Países que publicaram sobre IHD e suas instituições.

Países e Instituições	
País	Instituições
Reino Unido	University of Southampton
	University of Nottingham
	University of Cambridge
	University of St. Andrews
	University of London
	The Open University
	University of Warwick
	University of Reading
	The Open Data Institute
	Estados Unidos
Portugal	Universidade da Madeira Madeira-ITI ARDITI University of Lisbon
Itália	Università Cattolica del Sacro Cuore Università degli Studi di Milano-Bicocca Viale
Austrália	CSIRO University of Tasmania
Finlândia	IPR University Center Perspicamus Ab
Alemanha	University of Cologne Lower Saxony Memorial Foundation
Suíça	Universität Bern Lucerne University of Applied Sciences and Arts
Suécia	Södertörn University
Espanha	Universitat Pompeu Fabra
Irlanda	Bell Laboratories, Alcatel-Lucent
Brasil	Universidade Estadual de Campinas
Bélgica	Bell Laboratories, Alcatel-Lucent
Canadá	Ryerson University
Índia	Manipal Institute of Technology

Em *Boa Práticas, Diretrizes e Recomendações de design*, existem 2 trabalhos de visão *Ampla* e 1 trabalho de visão *Embodied*, respectivamente. Mashhadi *et al.* [41] propõem um conjunto de diretrizes que os provedores de dispositivo e serviços IoT possam considerar ao projetar seu modelo de negócios, bem como modelos de propriedade de dados de granularidade variável (*Pay-per-use, Data Market, Open Data*). Hutton e Henderson [31] sugerem boas práticas para que coletores de dados garantam que suas atividades de mineração de dados não violem as expectativas (e consentimento) das pessoas a quem os dados se relacionam. Mishra e Cafaro [43] identificam

Tabela 6: Autores com mais de um trabalho.

Autor	Autores e Trabalhos	
	#	Trabalhos
Andy Crabtree	5	[25], [20], [16], [46], [19]
Richard Mortier	4	[45], [20], [29], [46]
Tristan Henderson	3	[31], [45], [46]
Hamed Haddadi	3	[45], [29], [46]
Derek McAuley	3	[45], [29], [46]
Jon Crowcroft	3	[45], [29], [46]
Francesco Cafaro	3	[12], [13], [43]
Leilah Lyons	2	[13], [56]
Jessica Roberts	2	[13], [56]
Angela Locoro	2	[11], [39]

desafios que os usuários enfrentam ao usar um sistema com instruções mínimas, fornecendo recomendações de *design* para a criação de instalações interativas e envolventes relacionadas a exploração de dados, mesmo sem a intervenção e apoio de intérpretes.

Nos *Frameworks*, existem 5 trabalhos de visão *Ampla* e 1 trabalho de visão *Embodied*, respectivamente. Hornung *et al.* [30] estendem o *framework* Semiótico com o ciclo de vida de dados (e.g. origem de dados, seleção, limpeza, mapeamento, exibição ou interação) e as partes interessadas envolvidas, além de identificar desafios na área. Cavoukian e Chibba [15] propõem o uso do *framework Privacy By Design* para lidar com privacidade no contexto de IHD e cidades cognitivas. Wilke e Portmann [63] propõem a Computação Granular como base teórica, formal e metodológica para a IHD, fazendo uso do *general granular framework*. Leone [38] identifica os desafios éticos decorrentes do cenário IoT no âmbito da agroalimentação e propõe uma "abordagem ética no *design*" (*ethical in-design approach*) para a *Internet-of-Food*. Koesten *et al.* [34] entrevistam pessoas que trabalham com análise de dados e propõem um *framework* para a Interação Humano-"Dados estruturados", trabalho que possui um viés da visão *Visualização*. Finalmente, com a visão *Embodied* e um viés de *Visualização*, Pacheco *et al.* [49] propõem um *framework* que aplica RA para explorar e visualizar dados históricos dentro de um memorial ou museu.

Existem 2 *métodos* identificados nos trabalhos mapeados. Com a visão *Ampla*, e viés de *Visualização*, Kaipainen e Pitkänen [42] propõem um método sistemático, implementado pelos autores em um sistema, para explorar domínios focais alternativos no contexto de análise de cluster (*Cluster Analysis*). Com a visão *Embodied*, Roberts e Lyons [56] abordam desafios relacionados à identificação e medição de *learning talks*, propondo a metodologia SQuILD (*Scoring Qualitative Informal Learning Dialogue*).

Existem 2 *sistemas* identificados nos trabalhos mapeados. O primeiro, com a visão *Embodied*, foi criado por Cafaro *et al.* [13] como uma exibição de museu na qual os visitantes interagem com um grande conjunto de dados de censo demográfico, e possui dois modelos diferentes de interação: Controle Tangível, em que *tags* RFID são incorporadas em alguns artefatos e vários usuários podem controlar as informações na tela movendo esses artefatos; e o Controle *Embodied*, em que as pessoas carregam diretamente uma *tag* RFID e interagem com as informações caminhando dentro do espaço da

Tabela 7: Eventos que publicaram os trabalhos selecionados.

<i>Evento</i>	Eventos e Trabalhos	
	# Trabalhos	Trabalhos
ACM CHI Conference on Human Factors in Computing Systems (CHI)	2	[34], [23]
UbiComp	2	[12], [13]
IEEE International Conference on Cloud Engineering Workshop (IC2EW)	1	[19]
European Conference on Computer-Supported Cooperative Work (ECSCW)	1	[20]
International Conference on Recent Trends in Information Technology (ICRTIT)	1	[17]
IEEE World Forum on The Internet of Things (WF-IoT)	1	[41]
International Symposium on Big Data Visual Analytics (BDVA)	1	[62]
ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS)	1	[48]
International Conference on Tangible, Embedded and Embodied Interactions (TEI)	1	[43]
IADIS International Conference e-Health (IADIS EH)	1	[11]
Cultures of Participation in the Digital Age (CoPDA)	1	[39]
Digital Heritage	1	[49]
Total	14	

Tabela 8: Veículos que publicaram os trabalhos selecionados.

<i>Publicação</i>	<i>Editora</i>	Publicações e Trabalhos	
		Área	Trabalhos
University of Cambridge Computer Laboratory Technical Reports	Cambridge	Computação	[29]
SSRN Electronic Journal	Social Science Electronic Publishing	Ciências Sociais e Humanidades	[45]
Human-Computer Interaction: Design and Evaluation	Springer International Publishing	IHC	[30]
Studies in Big Data	Springer International Publishing	Big Data	[31]
Studies in Systems, Decision and Control	Springer International Publishing	Avanços em sistemas amplamente percebidos, tomada de decisão e controle	[15]
Granular Computing	Springer-Verlag	Computação Granular	[63]
Personal & Ubiquitous Computing	Springer-Verlag	Computação Ubíqua	[16]
Journal of Agricultural and Environmental Ethics	Springer-Verlag	Ciência agrícola e Bioética	[38]
Computer Supported Cooperative Work (CSCW)	Springer-Verlag	CSCW	[25]
International Journal of Computer-Supported Collaborative Learning	Springer-Verlag	CSCL	[56]
NSF science of interaction for data and visual analytics workshop		Computação	[32]
Total			11

exibição. No segundo sistema, com a visão *Ampla* e viés de *Visualização*, Nunes *et al.* [48] desenvolveram e implantaram o *Beanstalk*, uma infraestrutura de rastreamento passivo de baixo custo em Wi-Fi para coleta de dados de forma não intrusiva, com o objetivo de explorar a possibilidade de prover informações para uma comunidade ampla de partes interessadas sobre padrões espaço-temporais dos movimento de pessoas em locais turísticos.

Existem 2 *workflows* identificados nos trabalhos mapeados, ambos com a visão *Ampla*. Chamberlain e Crabtree [16] entrevistam pessoas para identificar *workflows* de atividades relacionadas a

"ouvir música" e seus metadados por meio de uma análise etnometodológica. Fischer *et al.* [25] (também com o autor Crabtree) elaboram métodos colaborativos de trabalho com dados e analisam as implicações que eles têm para o *design* no contexto de tecnologias interativas em IoT.

Existem 5 *outros* trabalhos mapeados, 4 com a visão *Ampla* e 1 com a visão *Embodied*, que apresentam diferentes propostas, especialmente conceituais. Crabtree e Mortier [20] discutem desafios fundamentais que a IHD enfrenta a partir de uma perspectiva interacional, contextualizando com as áreas de Estudos Sociais e CSCW.

Em outro trabalho, Crabtree [19] oferece uma perspectiva sociológica sobre a regulamentação da proteção de dados e sua relevância para o *design* de tecnologias digitais que exploram ou "comercializam" dados pessoais. Pogrebna [53] propõe uma "hipótese de IHD Comportamental", que encara os dados como Tradicionais (e.g. dados de medidores de eletricidade e outros dispositivos "tradicionais"), Invasivos (geralmente influenciam o comportamento humano, e.g. aplicativos *fitness* podem fazer com que um indivíduo se exercite mais) e Inventivos (requer que o usuário adicione contexto aos dados coletados por meio de outros dispositivos). Chowdhury e Dhawan [17] identificam modelos de propriedade de dados para aplicativos em Cidades Inteligentes. E, finalmente, com a visão *Embodied* e tratado na PP1, Cafaro [12] apresenta um trabalho em andamento em que busca projetar conjuntos de gestos baseados em alegorias para sistemas de interação *embodied*.

Para responder *quem são os stakeholders alvo da proposta?* (PP5), agrupamos os trabalhos por visão. Quase todos os trabalhos têm como partes interessadas pesquisadores de IHD. Na visão *Colaborativa*, o único trabalho tem como partes interessadas pesquisadores de IHD e IHC, cientistas e gestores que consideram com fundos para pesquisa. Na visão *Visualização*, o único trabalho considera pesquisadores de IHD, IHC e *Visual Analytics*.

Na visão *Embodied*, quase todos os trabalhos consideram visitantes de museu, exceto Elmqvist [23]. Cafaro [12] também envolve designers de sistema de interação *embodied*, e Roberts e Lyons [56] consideram pesquisadores e *designers* de soluções *Embodied*.

Na visão *Ampla*, além de pesquisadores de IHD: Nunes *et al.* [48] considera partes interessadas relacionadas a economia do Turismo; Leone [38] foca em pesquisadores de Internet-of-Food e Ética, além de cidadãos; Hutton e Henderson [31] consideram coletores de dados e legisladores; Wilke e Portmann [63] consideram pesquisadores de Computação Granular e Cidade Cognitivas; Cavoukian e Chibba [15] consideram pesquisadores de Cidades Cognitivas e Privacidade; Fischer e Crabtree [25] consideram pesquisadores e designers de Sistemas colaborativos em IoT, consultores de energia e seus clientes; Mashhadi *et al.* [41] consideram pesquisadores de IoT e provedores de dispositivos e serviços IoT; Crabtree e Mortier [20] consideram pesquisadores de CSCW e Estudos Sociais; Kaipainen e Pitkänen [42] consideram analistas de dados; Cabitza e Locoro [11] consideram pesquisadores e profissionais da área de Saúde; Chamberlain e Crabtree [16] consideram pessoas que escutam música em dispositivos eletrônicos; Pogrebna [53] considera pesquisadores de Modelos de negócios e *Servitization*.

Identificamos que os trabalhos da visão *Ampla* tendem a focar nos especialistas do domínio tratado, e quase todos incluem o usuário final nas suas preocupações. Porém, poucos os incluem de forma prática, e.g. investigando como trabalham, por meio de entrevistas [34] ou observações *in situ* [25]. Entretanto, todos os trabalhos da visão *Embodied* consideram os usuários finais em suas soluções.

Para identificar *como as propostas são avaliadas?* (PP6), agrupamos os trabalhos em suas visões. Dos 20 trabalhos que ofereceram algo diferente de definir IHD e propor desafios, apenas 7 avaliaram suas propostas, com destaque para a visão *Embodied* e aos autores relacionados à F. Cafaro, com 4 trabalhos.

Existem 5 trabalhos da visão *Embodied* com avaliação da proposta. Mishra e Cafaro [43] fizeram um estudo com 56 visitantes do

museu de ciências de *Discovery Place*, em Charlotte, NC, no qual testaram os pontos de entrada de interação em aplicativo sobre nome de países. Roberts e Lyons [56] avaliaram a codificação do diálogo dos participantes após as sessões na aplicação da proposta, comparando a qualidade do que foi aprendido e a taxa de aprendizado, bem como algumas métricas em relação a interação executada, em uma exposição de museu com controle individual e compartilhado, e múltiplas entradas. Cafaro [12] apresentou um trabalho em andamento no qual avaliou diferentes estilos de animação para a exibição citada no trabalho com 9 visitantes do museu, mas sem explicitar como fez. Em outro trabalho, Cafaro [13] avaliou preliminarmente o sistema ao medir o tempo de leitura das *tags* RFID da exibição em diferentes condições. Pacheco *et al.* [49] selecionaram 11 visitantes para usar a proposta (*tablet* com um aplicativo de RA para mostrar informações dos itens presentes no Memorial), separados em dois grupos. Um grupo teve uma exploração "guiada" pelo aplicativo (5 pessoas) e o outro teve uma exploração "livre" (6 pessoas). No final do experimento, os participantes tiveram que colocar marcadores no lugar em que viram alguns itens.

Existem 2 trabalhos da visão *Ampla* com avaliação da proposta. Fischer e Crabtree [25] realizaram um *workshop* reflexivo no qual os autores e os consultores de energia (*energy advisors*) conversaram sobre a implementação do sistema (CharIoT) nas casas dos clientes e como isso pode ajudar os consultores a trabalhar, auxiliando na identificação do *workflow*. Chowdhury e Dhawan [17] conduziram uma análise com o Método Delphi com uma equipe formada tanto pela comunidade cidadã quanto pela academia, mas não é explicitado como foi feita e quais os resultados.

Percebemos que todos os trabalhos que possuem alguma avaliação da proposta consideram e envolvem o público-alvo. Os trabalhos da visão *Embodied* observam como os usuários finais interagem com suas propostas e depois avaliam as interações, enquanto os trabalhos da visão *Ampla* realizam a avaliação juntamente com os usuários finais por meio de entrevistas.

Para responder *quais bases teóricas fundamentam as propostas mapeadas?* (PP7), separamos os trabalhos por visões. Na visão *Colaborativa*, Kee *et al.* [32] se posicionam sobre planejamento de longo prazo (para a ciência) e financiamento de infraestrutura, referenciando organização de tempo [4] [5] e infraestrutura cibernética [54] [8] [33]. Na visão *Visualização*, Widjojo *et al.* [62] utilizam a área de *Visual Analytics* com RV para propor uma IHD baseada em RV.

Na visão *Embodied*, Cafaro [13] usa colaboração em museus [61] e experimentos/exibições com *tags* RFID para propor o sistema multi-usuário tangível e *embodied*. Elmqvist [23] propõe a IHD *embodied* usando interação *embodied* [22] com cognição externa (*external cognition*) [47] para *sensemaking*. Cafaro [12] utiliza o *embodied schemata* [35] e o *Whole Body Interaction framework* [24] para apoiar o *design* de gestos para interagir com sistemas de IHD. Pacheco *et al.* [49] citam a aprendizagem em psicologia, pesquisa exploratória, aprendizagem construtivista com prática para criar o *framework* e sistema de RA para exploração do museu. Roberts e Lyons [56] propõem um método que identifique e meça conversas produtivas (*productive talk*) em uma exposição de museu aberta e *embodied*, adotando uma perspectiva vigotskiana de que "os seres humanos praticamente sempre aprendem através de diálogos" e *learning talk* [2]. Mishra e Cafaro [43] propõem a personalização como ponto de entrada para a IHD e o uso de silhuetas para apoiar

o engajamento em uma exposição de museu que permite que os visitantes explorem grandes conjuntos de dados.

Na visão *Ampla*, separamos os trabalhos em grupos que priorizam ou definem a área ou que propõem alguma solução. No primeiro, Haddadi *et al.* [29] utilizam referências de *Big Data*, definição de IHD de *elmqvist11* [23], conflitos no uso de dados e expande a definição de IHD para suportar as áreas que os co-autores consideram importantes como: visualização de dados e *sensemaking*; transparência e auditoria; privacidade e controle; *analytics* e comércio; *data to knowledge*. Mashhadi *et al.* [41] utilizam Consentimento em computação ubíqua [40] para propor uma visão bidimensional do espaço da IoT e propor 3 modelos de propriedade de dados (Pay-Per-Use [26], Data Market [28] e Dados Abertos [37]). Mortier *et al.* [45] separam os dados criados por nós (indivíduos) e sobre nós (inferidos sobre nossos dados), propõem dados - dados pessoais - como *boundary objects* [59] [36], e ainda definem os temas centrais da IHD como Legibilidade, Agência e Negociabilidade. Crabtree e Mortier [20] identificam a IHD como uma infra-estrutura socio-técnica [58], sendo que sua interação precisa de mecanismos que apoiem o *Articulation Work* [57] e que tenha a noção de "espaços de informação comuns" (nos quais *boundary objects* [59] são entendidos como "contêineres e portadores" de informações entre atores e organizações [6]). Locoro [39] utiliza dados derivados (*derivative data* [27]) para separar dados humanos primários (*primary human data*, diz respeito ao produtor de dados) dos secundários (*secondary human data*, diz respeito ao refinador/limpador/agregador de dados) e terciários (*tertiary human data*, refere-se ao consumidor de dados). Cabitzza e Locoro (2016) [11] segregam dados e processos relacionados em dados primários [7], dados secundários [1] e dados terciários, para contextos da área de saúde com o apoio da IHD.

Ainda na visão *Ampla*, considerando os trabalhos que propõem alguma solução, Hornung *et al.* [30] se fundamentam na Semiótica Organizacional e sugerem um ciclo de vida de dados para propor uma estrutura conceitual que considere as diferentes dimensões da IHD em diferentes estágios do ciclo de vida dos dados, levando em consideração as diferentes partes interessadas envolvidas nesses estágios. Pogrebna [53] propõe a "hipótese comportamental de IHD", e indica o uso de pesquisas em ciência comportamental, modelos de negócios orientados por dados, mercados multilaterais para apoiar a hipótese. Chowdhury e Dhawan [17] estendem os modelos de propriedade de dados de *mashhadi14* [41]. Wilke e Portmann [63] propõem o uso da Computação Granular e do *general computing framework* [50] para apoiar a IHD, citando os temas centrais de Mortier *et al.* [45]. Cavoukian e Chibba [15] utilizam o framework holístico 'Privacy by design' [14]. Crabtree [19] apresenta a perspectiva sociológica atual da Economia Digital e Regulamentação de Dados Pessoais e propõe que a IHD seja uma referência para novas visões e trabalhos na área. Chamberlain e Crabtree [16] utilizam *workflow* como "o desdobramento da atividade de trabalho ao longo do tempo" [9], literatura de metadados e comparação de planos e atividade situada (*situated activity*) [60] para examinar "as atividades comuns envolvidas no consumo de música doméstica". Fischer e Crabtree [25] utilizam *Articulation Work in situ* [57] com ação colaborativa (*collaborative action*) [10] para elaborar métodos colaborativos de trabalho com dados (*data work*) e suas implicações para o *design* de tecnologias IoT interativas. Kaipainen e Pitkänen [42] utilizam *Cluster Analysis* (algoritmos de base não

referenciados) para permitir uma análise de dados mais interativa e transparente na qual o analista escolhe os *subclusters* de seus dados com a ajuda da metodologia proposta, que sugere as lacunas mais amplas na distribuição das variáveis, encontrando prováveis interesses do analista. Nunes *et al.* [48] utilizam estudos em Mobilidade Urbana e Análise de Redes/Profiling para implementar o sistema de rastreamento de turistas com Wi-Fi passivo. Leone [38] utiliza o conceito de "ethical design" [3] com o *framework 'Big Data Ethics'* (no nível normativo) [55] para propor uma "ethical in-design approach" para a *Internet-of-Food*. Hutton e Henderson [31] apresentam literatura sobre Consentimento (para dados pessoais) e fornecem boas práticas para trabalhar com Legibilidade, Agência e Negociabilidade [46] em processos de mineração de dados. Koesten *et al.* [34] propõe uma estrutura conceitual para IHD trabalhar com dados estruturados, utiliza pesquisa de dados, *sensemaking* de dados e informações.

Consideramos algumas bases teóricas que parecem ser capazes de fundamentar a pesquisa e discussão em IHD. Dados vistos como *boundary objects* [59] são interessantes pois embasam a discussão ao compreender que eles significam muitas coisas para diferentes pessoas, mesmo mantendo integridade em relação ao seu sentido, e.g. dados sobre gastos indevidos de um político corrupto são ruins para sua imagem, enquanto para um auditor são garantias e provas para executar leis orçamentárias. Ainda, a Semiótica Organizacional em conjunto com o ciclo de vida dos dados [30] podem fundamentar a análise do que eles significam nas múltiplas camadas de significado e fases por quais passam. Encarar os dados no contexto de *derivative data* [27] é interessante por delimitar dados produzidos para um determinado fim (e que podem ser utilizados para outros fins). O conceito de *Articulation Work* [57] é pertinente para entender que o trabalho cooperativo é formado por um conjunto de atividades entre vários atores distintos, e.g. dados são coletados por dispositivos IoT, agregados em algum sistema, filtrados por analistas e discutidos pela comunidade interessada neles. A literatura de consentimento [40] [31] e *design ético* [3] é relevante ao considerarmos que os dados têm influência no mundo social e frequentemente são utilizados a despeito dos usuários finais, privilegiando transações entre empresas ou governos. Por fim, para fazer sentido dos dados, as áreas de *Visual Analytics*, *Cluster Analysis* e as relacionadas à interação *Embodied* são muito importantes no *design* de sistemas de IHD que favoreçam o entendimento e exploração dos dados.

Para responder *quais os elementos conceituais básicos para IHD apresentados no mapeamento?* (PP8), utilizamos como base a tabela de valores de Pereira *et al.* [51] e identificamos nos trabalhos os elementos que apareceram de maneira explícita ou condizentes com a descrição de um valor. Incluímos Transparência, Legibilidade, Agência e Negociabilidade no bloco de valores pois, embora não considerados na tabela base adotada, foram considerados importantes pelos trabalhos mapeados. Na Figura 2 estão indicados os valores encontrados e a frequência do encontro está expressa no tamanho da palavra.

Os elementos propostos são uma tentativa de encontrar um conjunto de elementos que represente as preocupações e desafios gerais da área que podem ser caracterizados como focos de pesquisa e investigação em IHD. Embora novos estudos precisem ser feitos para se chegar a um conjunto de elementos representativo e fundamentado, o conjunto identificado já permite verificar que IHD

possui diversos conceitos complexos e abrangentes que podem se desdobrar em subáreas de pesquisa. Os principais elementos são listados a seguir.



Figura 2: Elementos básicos presentes nos trabalhos.

Awareness é o elemento mais frequente entre os trabalhos, aparecendo em 26 trabalhos. Surge da necessidade de tornar os usuários conscientes dos dados incorporados nas soluções de IHD. As soluções atuais de análise de dados não parecem lidar bem com o grande volume imposto pelos cenários de Big Data e IoT, favorecendo apenas usuários técnicos e que fazem ideia do que estão procurando, deixando a desejar em alternativas que fazem a exploração desses conteúdos por partes interessadas leigas. Este elemento está relacionado com outros elementos, como *Autonomia*, *Transparência*, *Visibilidade*, *Usabilidade* e *Legibilidade*, pois estes também influenciam no entendimento dos dados para uma criação de sentido ou tomada de decisão apoiada pela IHD. Não existem soluções triviais nesse contexto e os autores estão buscando diferentes áreas para apoiá-las, como a Computação Granular, *Visual Analytics* e *Semiótica Organizacional*. Assim, acreditamos que este é um dos desafios críticos em IHD.

O elemento *Colaboração* é o segundo mais frequente, aparece em 22 trabalhos e parece influenciar muito a forma como os seres humanos utilizam os dados para fazer sentido deles. Os autores encontrados consideram com áreas como a CSCW, CSCL, Sociologia, Interação Tangível e *Embodied*, e propõem soluções nas quais as partes interessadas precisam colaborar entre si para melhorar seu entendimento dos dados. Apesar da colaboração não ser a melhor maneira de fazer sentido dos dados em qualquer contexto [56], em muitos deles tem se mostrado eficiente ou suficiente. Este elemento está relacionado com outros elementos, como *Conversaço*, *Relacionamentos*, *Compartilhamento* e *Grupos*, pois estes também fazem

parte do processo de colaboração entre pessoas. Como indicado por Kee *et al.* [32], a colaboração pode ser considerada um dos desafios da área de IHD.

Finalmente, o elemento *Privacidade* (12 trabalhos) e *Controle* (11 trabalhos) são importantes para a IHD. Muitos trabalhos da visão *Ampla* discutem a problemática de como os dados pessoais são tratados e comercializados por empresas, sem o consentimento dos usuários e violando questões éticas. Percebemos que os autores europeus trataram mais desse assunto e alguns referenciaram a *General Data Protection Regulation* (GDPR) [18], lei de proteção de dados Europeia que entrou em vigor em 25 de maio de 2018. A tendência é que mais regulamentações como a GDPR apareçam no mundo, estimuladas pelo cenário de produção em massa de dados pessoais. A IHD deve prover suporte teórico e metodológico a essas novas iniciativas, sendo crítico que novos trabalhos da área tratem dessa temática de forma explícita e informada.

4 CONCLUSÃO

A área de IHD aparece na literatura como uma tentativa de abordar a problemática da interação com dados, englobando não só a visualização e o fazer sentido do grande volume de dados, mas também as consequências dessa produção, como o consentimento e a privacidade das partes envolvidas (vide leis e regulamentos que aparecem no Brasil [21] e Europa [18]) ou os modos como as informações são inferidas por algoritmos a partir desses dados.

A IHD começou a ser discutida em 2011, porém ainda não existe um consenso sobre sua definição, nem trabalhos que agreguem ou articulem o que foi produzido até então. Buscando o sentido e formas de entender um grande volume de dados (além das consequências dessa produção), a IHD intenta abordar a problemática da Big Data sistematicamente, e ainda passa por um processo de consolidação enquanto área de pesquisa, começando a estabelecer termos e especificidades. Muitos trabalhos se encontram em estágio embrionário, se concentrando em discussões de problemas e oportunidades, envolvendo autores de diferentes especialidades (e.g., cientistas sociais, jurídicos e da computação), porém ainda com poucos resultados de investigações.

Este artigo contribui na direção de mapear e caracterizar as principais iniciativas em IHD que já estejam apresentando alguma proposta de investigação na área. Para isso, apresentamos um Mapeamento Sistemático de Literatura no qual foram encontrados 28 trabalhos que apresentam alguma solução conceitual ou prática em IHD, e respondidas 8 perguntas de pesquisa que ajudam a caracterizar os trabalhos mapeados. As perguntas de pesquisa escolhidas identificaram as definições propostas para a área, quem publica e onde são publicados os trabalhos, as propostas de trabalho que apareceram até então, os *stakeholders* envolvidos nessas propostas, os tipos de avaliação das propostas e as bases que fundamentam as propostas. O artigo finaliza apresentando um conjunto de conceitos que apareceram nos trabalhos mapeados e que podem ser considerados como focos de investigação para trabalhos futuros, contribuindo para delinear a área.

REFERÊNCIAS

- [1] Mervat Abdelhak, Sara Grostic, and Mary Alice Hanken. 2014. *Health Information - E-Book: Management of a Strategic Resource*. Elsevier Health Sciences. <https://books.google.com.br/books?id=69X1BQAAQBAJ>

- [2] Sue Allen. 2003. Looking for learning in visitor talk: A methodological exploration. In *Learning conversations in museums*. Routledge, 265–309.
- [3] Gianmarco Baldini, Maarten Botterman, Ricardo Neisse, and Mariachiara Tallacchini. 2018. Ethical Design in the Internet of Things. *Science and Engineering Ethics* 24, 3 (01 Jun 2018), 905–925. <https://doi.org/10.1007/s11948-016-9754-5>
- [4] Dawna I Ballard and David R Seibold. 2003. Communicating and organizing in time: A meso-level model of organizational temporality. *Management Communication Quarterly* 16, 3 (2003), 380–415.
- [5] Dawna I Ballard and David R Seibold. 2004. Organizational members' communication and temporal experience: Scale development and validation. *Communication Research* 31, 2 (2004), 135–172.
- [6] Liam Bannon and Susanne Bødker. 1997. Constructing common information spaces. In *Proceedings of the Fifth European Conference on Computer Supported Cooperative Work*. Springer, 81–96.
- [7] Marc Berg and Els Goorman. 1999. The contextual nature of medical information. *International journal of medical informatics* 56, 1-3 (1999), 51–60.
- [8] Ian Bird, Bob Jones, and Kerk F Kee. 2009. The organization and management of grid infrastructures. *Computer* 42, 1 (2009), 36–46.
- [9] John Bowers, Graham Button, and Wes Sharrock. 1995. Workflow from within and without: technology and cooperative work on the print industry shopfloor. In *Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work ECSCW'95*. Springer, 51–66.
- [10] Graham Button, Andy Crabtree, Mark Rouncefield, and Peter Tolmie. 2015. *Deconstructing Ethnography: Towards a Social Methodology for Ubiquitous Computing and Interactive Systems Design*. Springer International Publishing. <https://books.google.com.br/books?id=047DCgAAQBAJ>
- [11] Federico Cabitza and Angela Locoro. 2016. Human-Data Interaction in Healthcare: Acknowledging Use-related Chasms to Design for a Better Health Information. In *Proceedings of the 8th IADIS International Conference on e-Health 2016, Part of the Multi Conference on Computer Science and Information Systems, MCCSIS 2016*. 91–98.
- [12] Francesco Cafaro. 2012. Using Embodied Allegories to Design Gesture Suites for Human-data Interaction. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. ACM, New York, NY, USA, 560–563. <https://doi.org/10.1145/2370216.2370309>
- [13] Francesco Cafaro, Leilah Lyons, Joshua Radinsky, and Jessica Roberts. 2010. RFID Localization for Tangible and Embodied Multi-user Interaction with Museum Exhibits. In *Proceedings of the 12th ACM International Conference Adjunct Papers on Ubiquitous Computing - Adjunct (UbiComp '10 Adjunct)*. ACM, New York, NY, USA, 397–398. <https://doi.org/10.1145/1864431.1864455>
- [14] Ann Cavoukian. 2009. Privacy by design. *Take the challenge. Information and privacy commissioner of Ontario, Canada* (2009).
- [15] Ann Cavoukian and Michelle Chibba. 2016. *Cognitive Cities, Big Data and Citizen Participation: The Essentials of Privacy and Security*. Springer International Publishing, Cham, 61–82. https://doi.org/10.1007/978-3-319-33798-2_4
- [16] Alan Chamberlain and Andy Crabtree. 2016. Searching for music: understanding the discovery, acquisition, processing and organization of music in a domestic setting for design. *Personal and Ubiquitous Computing* 20, 4 (01 Aug 2016), 559–571. <https://doi.org/10.1007/s00779-016-0911-2>
- [17] Srinjoy N Chowdhury and Sania Dhawan. 2016. HDI based data ownership model for smart cities. In *2016 International Conference on Recent Trends in Information Technology (ICRTIT)*. 1–5. <https://doi.org/10.1109/ICRTIT.2016.7569514>
- [18] Council of European Union. 2016. Council regulation (EU) no 679/2016 (General Data Protection Regulation). <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1528874672298&uri=CELEX%3A32016R0679>.
- [19] Andy Crabtree. 2016. Enabling the New Economic Actor: Personal Data Regulation and the Digital Economy. In *2016 IEEE International Conference on Cloud Engineering Workshop (IC2EW)*. 124–129. <https://doi.org/10.1109/IC2EW.2016.18>
- [20] Andy Crabtree and Richard Mortier. 2015. Human Data Interaction: Historical Lessons from Social Studies and CSCW. In *ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work, 19-23 September 2015, Oslo, Norway*, Nina Boulus-Rødje, Gunnar Ellingsen, Tone Bratteteig, Margunn Aanestad, and Pernille Bjørn (Eds.). Springer International Publishing, Cham, 3–21.
- [21] Câmara dos Deputados. 2018. Projeto de Lei da Câmara n° 53, de 2018 (Proteção de dados pessoais). <https://www25.senado.leg.br/web/atividade/materias/-/materia/133486>.
- [22] Paul Dourish. 2004. *Where the action is: the foundations of embodied interaction*. MIT press.
- [23] Niklas Elmqvist. 2011. Embodied human-data interaction. In *ACM CHI 2011 Workshop "Embodied Interaction: Theory and Practice in HCI"*. 104–107.
- [24] David England. 2011. Whole body interaction: An introduction. In *Whole Body Interaction*. Springer, 1–5.
- [25] Joel E. Fischer, Andy Crabtree, James A. Colley, Tom Rodden, and Enrico Costanza. 2017. Data Work: How Energy Advisors and Clients Make IoT Data Accountable. *Computer Supported Cooperative Work (CSCW)* 26, 4 (01 Dec 2017), 597–626. <https://doi.org/10.1007/s10606-017-9293-x>
- [26] Daniel Fitton, Vasughi Sundramoorthy, Gerd Kortuem, James Brown, Christos Efstratiou, Joe Finney, and Nigel Davies. 2008. Exploring the design of pay-per-use objects in the construction domain. In *European Conference on Smart Sensing and Context*. Springer, 192–205.
- [27] Luciano Floridi. 2003. Two approaches to the philosophy of information. *Minds and Machines* 13, 4 (2003), 459–469.
- [28] Derek Foster, Mark Blythe, Paul Cairns, and Shaun Lawson. 2010. Competitive carbon counting: can social networking sites make saving energy more enjoyable?. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*. ACM, 4039–4044.
- [29] Hamed Haddadi, Richard Mortier, Derek McAuley, and Jon Crowcroft. 2013. *Human-data interaction*. Technical Report UCAM-CL-TR-837. University of Cambridge, Computer Laboratory. <http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-837.pdf>
- [30] Heiko Hornung, Roberto Pereira, M. Cecilia C. Baranauskas, and Kecheng Liu. 2015. Challenges for Human-Data Interaction – A Semiotic Perspective. In *Human-Computer Interaction: Design and Evaluation*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 37–48.
- [31] Luke Hutton and Tristan Henderson. 2017. *Beyond the EULA: Improving Consent for Data Mining*. Springer International Publishing, Cham, 147–167. https://doi.org/10.1007/978-3-319-54024-5_7
- [32] Kerk F Kee, Larry D Browning, and Dawna I Ballard. 2012. Sociomaterial Processes, Long Term Planning, and Infrastructure Funding: Towards Effective Collaboration and Collaboration Tools for Visual and Data Analytics. (2012). <http://www1.chapman.edu/~kee/PDF/C14.pdf>
- [33] Kerk F Kee, Lucy Craddock, Bridget Marie Blodgett, and Rami Olwan. 2011. *Cyberinfrastructure inside out: Definitions and influences shaping its emergence, development, and implementation*. Peter Lang. <https://doi.org/10.13016/M27Z8H>
- [34] Laura M. Koesten, Emilia Kacprzak, Jenifer F. A. Tennison, and Elena Simperl. 2017. The Trials and Tribulations of Working with Structured Data: -a Study on Information Seeking Behaviour. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1277–1289. <https://doi.org/10.1145/3025453.3025838>
- [35] George Lakoff and Mark Johnson. 2008. *Metaphors we live by*. University of Chicago press.
- [36] Susan Leigh Star. 2010. This is not a boundary object: Reflections on the origin of a concept. *Science, Technology, & Human Values* 35, 5 (2010), 601–617.
- [37] Seppo Leminen, Mika Westerlund, Mervi Rajahonka, and Riikka Siuruaime. 2012. Towards IoT ecosystems and business models. In *Internet of Things, Smart Spaces, and Next Generation Networking*. Springer, 15–26.
- [38] Luca Leone. 2017. Beyond Connectivity: The Internet of Food Architecture Between Ethics and the EU Citizenry. *Journal of Agricultural and Environmental Ethics* 30, 3 (01 Jun 2017), 423–438. <https://doi.org/10.1007/s10806-017-9675-6>
- [39] Angela Locoro. 2015. A Map is worth a Thousand Data: Requirements in Tertiary Human-Data Interaction to Foster Participation. In *Proceedings of the Third edition of the International Workshop on Cultures of Participation in the Digital Age: Coping with Information, Participation, and Collaboration Overload co-located with the Fifth International Symposium on End-User Development (IS-EUD 2015), Madrid, Spain, May 26, 2015*. 39–44. <http://ceur-ws.org/Vol-1641/paper7.pdf>
- [40] Ewa Luger and Tom Rodden. 2013. An informed view on consent for UbiComp. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 529–538.
- [41] Afra Mashhadi, Fahim Kawsar, and Utku G Acer. 2014. Human Data Interaction in IoT: The ownership aspect. In *2014 IEEE World Forum on Internet of Things (WF-IoT)*. 159–162. <https://doi.org/10.1109/WF-IoT.2014.6803139>
- [42] Olli Pitkänen, Mauri Kaipainen, and Pespicaamus Ab. 2017. Human-controlled iterative subclustering analysis. In *2017 IEEE International Conference on Big Data (Big Data)*. 4754–4756. <https://doi.org/10.1109/BigData.2017.8258532>
- [43] Swati Mishra and Francesco Cafaro. 2018. Full Body Interaction Beyond Fun: Engaging Museum Visitors in Human-Data Interaction. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '18)*. ACM, New York, NY, USA, 313–319. <https://doi.org/10.1145/3173225.3173291>
- [44] Richard Mortier, Hamed Haddadi, Tristan Henderson, Derek McAuley, and Jon Crowcroft. 2013. Challenges Opportunities in Human-Data Interaction. In *Proceedings of DE2013: Open Digital - The Fourth Annual Digital Economy All Hands Meeting*.
- [45] Richard Mortier, Hamed Haddadi, Tristan Henderson, Derek McAuley, and Jon Crowcroft. 2014. Human-data interaction: the human face of the data-driven society. (2014). <https://doi.org/10.1145/2370216.2370309>
- [46] Richard Mortier, Hamed Haddadi, Tristan Henderson, Derek McAuley, Jon Crowcroft, and Andy Crabtree. 2016. Human-Data Interaction. In *The Encyclopedia of Human-Computer Interaction 2 ed*. Interaction Design Foundation, Chapter 41. <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/human-data-interaction>
- [47] Donald A. Norman. 1993. *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*. Addison-Wesley Longman Publishing Co., Inc., Boston,

- MA, USA.
- [48] Nuno Nunes, Miguel Ribeiro, Catia Prandi, and Valentina Nisi. 2017. Beans-talk: A Community Based Passive Wi-fi Tracking System for Analysing Tourism Dynamics. In *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '17)*. ACM, New York, NY, USA, 93–98. <https://doi.org/10.1145/3102113.3102142>
 - [49] Daniel Pacheco, Sytse Wierenga, Pedro Omedas, Laura S. Oliva, Stefan Wilbricht, Stephanie Billib, Habbo Knoch, and Paul F.M.J. Verschure. 2015. A location-based Augmented Reality system for the spatial interaction with historical datasets. In *2015 Digital Heritage*, Vol. 1. 393–396. <https://doi.org/10.1109/DigitalHeritage.2015.7413911>
 - [50] Witold Pedrycz. 2006. Granular computing: an overview. In *Applied Soft Computing Technologies: The Challenge of Complexity*. Springer, 19–34.
 - [51] Roberto Pereira, M Cecília C Baranauskas, and Sergio Roberto P da Silva. 2013. Social software and educational technology: Informal, formal and technical values. *Journal of Educational Technology & Society* 16, 1 (2013), 4.
 - [52] Kai Petersen and Cigdem Gencel. 2013. Worldviews, research methods, and their relationship to validity in empirical software engineering research. In *Software Measurement and the 2013 Eighth International Conference on Software Process and Product Measurement (IWSM-MENSURA), 2013 Joint Conference of the 23rd International Workshop on*. IEEE, 81–89.
 - [53] Ganna Pogrebna. 2015. *Servitization through human-data interaction: a behavioural approach*. Working Paper. University of Warwick, Coventry. <http://wrap.warwick.ac.uk/67283/>
 - [54] David Ribes and Thomas A Finholt. 2009. The long now of technology infrastructure: articulating tensions in development. *Journal of the Association for Information Systems* 10, 5 (2009), 375.
 - [55] Neil M Richards and Jonathan H King. 2014. Big data ethics. *Wake Forest L. Rev.* 49 (2014), 393.
 - [56] Jessica Roberts and Leilah Lyons. 2017. The value of learning talk: applying a novel dialogue scoring method to inform interaction design in an open-ended, embodied museum exhibit. *International Journal of Computer-Supported Collaborative Learning* 12, 4 (01 Dec 2017), 343–376. <https://doi.org/10.1007/s11412-017-9262-x>
 - [57] Kjeld Schmidt and Liam Bannon. 1992. Taking CSCW seriously. *Computer Supported Cooperative Work (CSCW)* 1, 1-2 (1992), 7–40.
 - [58] Susan L Star. 1999. The ethnography of infrastructure. *American behavioral scientist* 43, 3 (1999), 377–391.
 - [59] Susan L Star and James R Griesemer. 1989. Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social studies of science* 19, 3 (1989), 387–420.
 - [60] Lucy A Suchman. 1987. *Plans and situated actions: The problem of human-machine communication*. Cambridge university press.
 - [61] Dirk Vom Lehn, Christian Heath, and Jon Hindmarsh. 2001. Exhibiting interaction: Conduct and collaboration in museums and galleries. *Symbolic interaction* 24, 2 (2001), 189–216.
 - [62] Elisabeth A Widjojo, Winyu Chinthammit, and Ulrich Engelke. 2017. Virtual Reality-Based Human-Data Interaction. In *2017 International Symposium on Big Data Visual Analytics (BDVA)*. 1–6. <https://doi.org/10.1109/BDVA.2017.8114627>
 - [63] Gwendolin Wilke and Edy Portmann. 2016. Granular computing as a basis of human-data interaction: a cognitive cities use case. *Granular Computing* 1, 3 (01 Sep 2016), 181–197. <https://doi.org/10.1007/s41066-016-0015-4>
 - [64] Annika Wolff, Ahmed Seffah, Gerd Kortuem, and Janet van der Linden. 2018. Designing for Effective Interactions with Data in the Internet of Things. In *Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems (DIS '18 Companion)*. ACM, New York, NY, USA, 415–418. <https://doi.org/10.1145/3197391.3197402>

APÊNDICE B – MANUAL DE USO - HUMAN-DATA IDEATION

Segue o manual produzido para apresentar o *framework* conceitual HDIdea para projetistas leigos em Interação Humano-Dados, seu público-alvo.

Human-Data Ideation

Framework conceitual para projetistas de *software* no entendimento de problemas em contextos de IHD

Facilitando o *design* rationale do produto

Créditos editoriais

Editoração, diagramação e arte

Robertha Trevisan Coradassi Buff

Texto

Mateus Rambo Strey

Revisão do texto

Roberto Pereira

Imagens

Capa e cabeçalhos: Foto por Serge Kutuzov, disponível em Unsplash

Foto (p. 4) por Samuel Zeller, disponível em Unsplash



Novo projetista:

75

Vai projetar um sistema de Interação Humano-Dados (IHD) e não sabe por onde começar ou quais elementos deve ser considerados?

Uma solução é aprender a utilizar a metodologia Human-Data Ideation para garantir projetos mais estruturados!

IHD pode ser entendida como área que estuda a manipulação humana, análise e sensemaking (o fazer sentido) de dados, sendo que volume de dados não é relevante se a interação é capaz de ampliar a capacidade humana na realização dessas atividades.

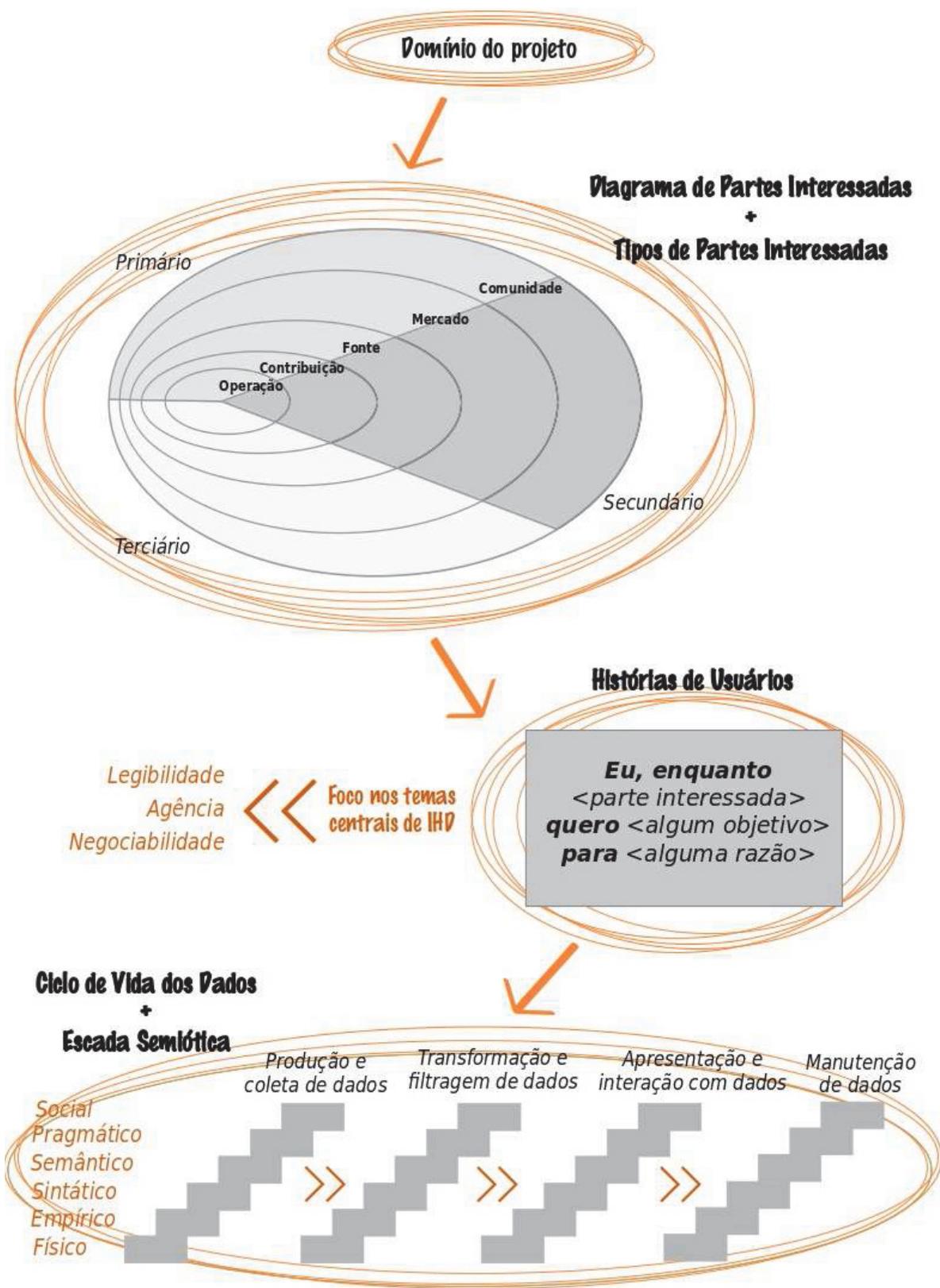
O que é e para que serve?

O *Human-Data Ideation* é um *framework* conceitual desenvolvido para apoiar projetistas de *software* no entendimento de problemas em contextos de IHD. O objetivo é facilitar o *design* rationale do produto, considerando desafios como: usuários comuns corrigindo dados ou compreendendo como os dados são utilizados pelos sistemas, questões de segurança e privacidade, etc.

O uso do *Human-Data Ideation* começa com a identificação das partes interessadas no projeto, reconhecendo o que fazem em contextos de IHD, sucedendo na ampliação de possibilidades e soluções de problemas em diferentes níveis de abstração.

Você pode utilizar o Human-Data Ideation de forma colaborativa para trazer mais dinamismo na elaboração do projeto. Reúna as pessoas, leve post-its, organizem um café e transforme o planejamento do projeto num momento agradável e que deixe os participantes a vontade para contribuírem com suas ideias!

O framework



Domínio do projeto (contexto)

Imagine que você foi contratado pela prefeitura de seu município para reformular o sistema de transparência de contas públicas. O contratante quer ferramentas que melhorem a visualização da informação (geradas a partir dos dados) e que não violem a privacidade das entidades privadas. Você sabe que existe uma secretaria responsável por alimentar os dados no sistema atual.

É importante, antes de iniciar, saber em qual contexto o projeto será construído: qual ambiente, quanto tempo há disponível, qual a finalidade, quem vai usar, etc.



O que deve ser considerado na concepção do projeto?

1. Identificação e classificação de stakeholders (partes interessadas)

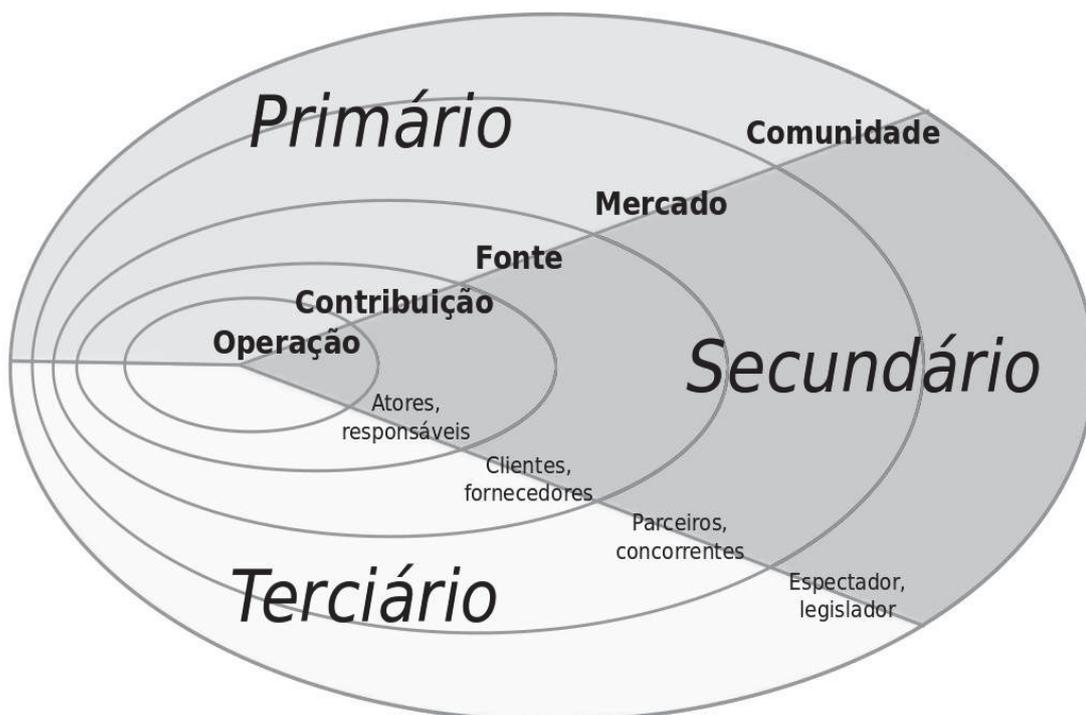
É essencial reconhecer quem está relacionado ao sistema.

No caso de nossa situação hipotética, identificou-se no contexto as seguintes instituições e pessoas que usarão o sistema: a Prefeitura, gestores públicos e cidadãos. Mas será que todas as partes foram identificadas?

a. Para facilitar a identificação e classificação utiliza-se o Diagrama de Partes Interessadas (DPI) e os Tipos de Parte Interessada (TPI).

O DPI é uma ferramenta que auxilia a identificar partes interessadas com diferentes intenções no problema. É válido ressaltar que uma parte pode ter mais de uma intenção. O TPI classifica as partes interessadas que atuarão com dados dentro do problema em 3 tipos.

Na figura abaixo temos um modelo das camadas a serem preenchidas:



As 5 camadas do DPI são:

Operação: o sistema/projeto em si e partes que irão construí-lo (ex.: portal de transparência, projetista, desenvolvedores);

Contribuição: partes que contribuem diretamente para o problema, para a sua solução e/ou são afetados diretamente por ele (ex.: Prefeitura, Secretarias);

Fonte: partes que fornecem dados/informações ou os consomem (ex.: Secretarias, gestores públicos);

Mercado: partes que são parceiros ou concorrentes do projeto (ex.: outros sistemas de transparência de contas públicas, sistema já usado pela prefeitura, grupos de interesse econômico ou político, servidor público acobertando uma fraude);

Comunidade: partes que são espectadores externos, sociedade e legisladores (ex.: cidadãos que querem fiscalizar os gastos públicos, jornalistas investigativos, empresas interessadas em licitações, legislação sobre a transparência das contas públicas).

Os 3 tipos de parte interessadas podem ser:

Primárias: partes que produzem dados a partir e para suas próprias práticas, atividades ou contextos (ex.: secretarias produzindo dados para seu próprio controle);

Secundárias: partes que refinam e agregam o que é feito pelas partes primárias e que estão envolvidas em atividades de gerenciamento, administração ou pesquisa (ex.: gestores públicos, secretaria que agrega os dados das outras secretarias para publicar no portal);

Terciária: fazem uso do que é produzido pelas partes primárias e/ou secundárias utilizando-se dos dados em diferentes contextos (ex.: legisladores, cidadãos que fazem uso dos dados para cobrar ações por parte dos políticos, empresas interessadas em licitações, órgão fiscalizadores, jornalistas investigativos).

b. É importante encaixar todos as partes interessadas nos dois diagramas para que seja possível perceber tanto suas motivações, quanto para entender como ele se relaciona com os dados no projeto!

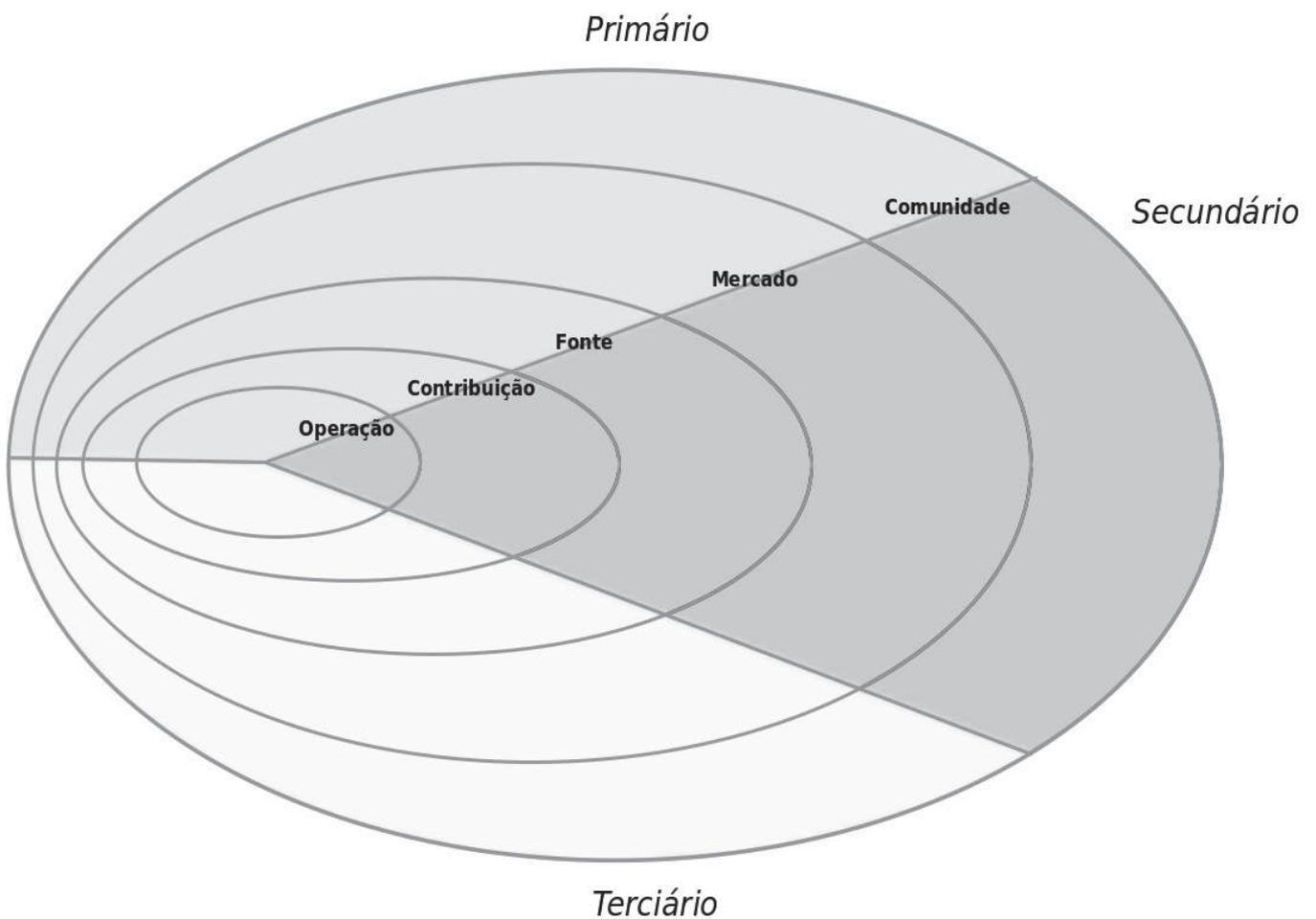
c. O resultado da aplicação do DPI e TPI é uma lista de partes interessadas no projeto.

Agora é sua vez de praticar!

Pense em todas as partes interessadas que o seu projeto pode conter e preencha os diagramas a seguir:

Identificação e classificação de stakeholders

Diagrama de Partes Interessadas e Tipos de Partes Interessadas





2. O que essas partes interessadas fazem em contextos de IHD

Sabendo quem são as partes interessadas, é o momento de compreender suas ações e expectativas em relação ao projeto.

a. Observe os diagramas que você preencheu no passo 1 e reflita: quais das partes interessadas desempenham papéis-chave ou serão mais afetadas pelo projeto? Selecione as principais partes interessadas e, se possível, escolha ao menos 1 de cada tipo, buscando aumentar as diferentes visões e necessidades representadas no entendimento do problema.

No caso do portal de transparência, foram selecionadas como partes interessadas relevantes: (I) Prefeitura, (II) Secretarias, (III) Servidor Público acobertando uma fraude, (IV) Empresas interessadas em licitações, (V) Jornalistas investigativos, (VI) Cidadãos

b. Feita a seleção, a próxima etapa é escrever uma história de usuário.

Um formato sugerido é de Cohn (2008):

"Eu, enquanto parte interessada quero algum objetivo para alguma razão"

Se você já utiliza outros formatos ou encontrar outro que seja adaptado a sua realidade, sinta-se à vontade para substituir essa sugestão.

Ao criar suas histórias de usuário, você deve inserir cada um dos principais aspectos de IHD em ao menos uma história: Legibilidade, Agência e Negociabilidade. Vamos conferir alguns exemplos de como as histórias de usuários podem ser feitas:



Legibilidade: Se preocupa com que termos, siglas, dados e algoritmos analíticos sejam transparentes e compreensíveis para os usuários.

Se questione:

As partes interessadas tem como saber o que aconteceu ou está acontecendo com os dados?

O processo e o fluxo dos dados é transparente e legível?

Eu, enquanto **jornalista investigativo**, quero **receber avisos sobre movimentações suspeitas**, para **publicar matérias e informar o público**.

Os dados revelam algum tipo de relação ou conexão entre entidades e pessoas?

Existe alguma forma de observar essa conexão?

Eu, enquanto **servidor público acobertando uma fraude**, quero **que não seja possível visualizar relações entre as entidades e pessoas envolvidas no esquema**, para **preservar a fraude de uma possível investigação**.

Eu, enquanto **jornalista investigativo**, quero **ver uma rede de conexões (grafo ou diagrama ator-rede) entre empresas beneficiadas em licitações e pessoas físicas associadas a elas**, para **identificar possíveis suspeitos de corrupção**.

Agência: Relaciona as ações dos sujeitos sobre os dados, permitindo ou não a coleta ou interferência nas análises feitas a partir dos dados.

Se questione:

As partes interessadas podem corrigir algum dado sobre elas mesmas?

Eu, enquanto **cidadão**, quero **denunciar que não fui beneficiado com uma diária de viagem**, para **não ser implicado em um caso de corrupção**.

Eu, enquanto **empresa interessada em licitações**, quero **corrigir meu CNPJ erroneamente identificado no sistema**, para **evitar potenciais problemas ao ser escolhido em um edital**.



As partes interessadas conseguem fazer o desejam no sistema de forma autônoma?

Eu, enquanto **Secretaria**, quero **lançar os dados necessários no sistema sem auxílio externo (ex., desenvolvedores do sistema), para trabalhar sem maior burocracia.**

As partes interessadas estão conscientes sobre o que está sendo feito com os seus dados?

Eu, enquanto **cidadão**, quero **encontrar facilmente qualquer informação a meu respeito no sistema, para descobrir se minha privacidade foi violada ou se estou erroneamente indicado em alguma despesa.**

É interessante que as partes interessadas consigam trabalhar em conjunto em determinadas tarefas do sistema?

Eu, enquanto **jornalista investigativo**, quero **salvar e compartilhar os cruzamentos de dados que fiz no sistema, para que meus colegas me ajudem a encontrar possíveis delitos sem maiores dificuldades.**

Negociabilidade: Considera o aspecto social dos dados num contexto espaço-temporal e como lidar com as mudanças de paradigma na propriedade sobre os dados, ou na rentabilidade deles

Se questione:

Qual o aparato legal vigente e como as partes interessadas conseguem ter autonomia na lida com os dados?

Eu, enquanto **Prefeitura**, quero **observar a Lei da Transparência sem expor os dados privados dos meus servidores, para respeitar as leis e meus funcionários.**

As partes interessadas podem reivindicar sua privacidade?

Eu, enquanto **empresa interessada em licitações**, quero **que meus dados sensíveis sejam removidos do sistema caso estejam públicos, para evitar que fraudadores obtenham informações indevidas para praticar delitos.**



As partes interessadas podem ter comportamentos antiéticos com auxílio ou em relação ao sistema?

Eu, enquanto **Servidor Público** acobertando uma fraude, quero **modificar dados lançados no sistema que revelem a fraude, para proteger o esquema de corrupção.**

Eu, enquanto **Prefeitura**, quero **que os dados presentes no sistema possuam um histórico de mudanças, para identificar possíveis servidores mal-intencionados.**

Uma história de usuário exprime o que as partes interessadas querem e/ou precisam do sistema e deve responder três questões: quem, o quê e por quê.

c. Como resultado, serão obtidas histórias de usuários relevantes ao problema e que englobam os vários desafios em IHD, podendo ser refinadas em requisitos funcionais e não-funcionais.

3. Identificação e entendimento de problemas em diferentes níveis de abstração

a. A partir das histórias de usuário identificadas, o projetista deverá escolher ao menos 1 história de usuário para expandi-la no artefato Ciclo de Vida dos Dados (4 estágios) estendido com a Escada Semiótica (6 níveis). O projetista deve tentar identificar questões ou problemas de projeto em todos os 4 estágios do ciclo de vida dos dados, em cada nível da Escada Semiótica.

Os estágios do Ciclo de Vida dos Dados são:

Produção e coleta dos dados: criação e recepção dos dados a partir de sensores, pesquisas, bases, etc;

Transformação e filtragem dos dados: limpeza e seleção dos dados recebidos;

Apresentação e interação com os dados: interações que permitam que as partes interessadas façam sentido e obtenham novas ideias a partir dos dados;

Manutenção dos dados: políticas de armazenamento e exclusão dos dados.



A Escada Semiótica apresenta uma separação conceitual entre o técnico e o social.

A camada técnica possui 3 níveis:

Físico (qual o suporte físico para o *software*?): *demanda de memória, capacidade de processar vídeos, servidores...*

Empírico (quais são os sistemas que comunicam usando o mundo físico?): *protocolos de comunicação, linguagem de programação, otimização de código...*

Sintático (qual é a estrutura do *software*?): *arquitetura de software, estrutura do layout, modelos de navegação...*

A camada social possui outros 3 níveis:

Semântico (o que a interface significa para o público alvo?): *propostas e significados dos elementos da interface, ícones...*

Pragmático (quais tarefas que a interface apoia?): *intenções, comunicações, conversações, negociações...*

Mundo Social (quais os efeitos do uso do *software*?): *crenças, expectativas, funções, compromissos, contratos, lei, cultura...*

b. Recomenda-se o preenchimento do artefato analisando, pelo menos, 1 história de usuário de cada aspecto da área de IHD, pois o artefato pode levantar questionamentos que os passos anteriores não cobrem. Não é necessário preencher todos os campos, apenas os que fazem sentido para a história (mas analise todos os campos!).

c. O resultado do uso do Ciclo de Vida dos Dados estendido com a Escada Semiótica são questões, preocupações e possíveis soluções que favorecem o entendimento do problema, e que podem ser refinados em requisitos funcionais e não-funcionais.

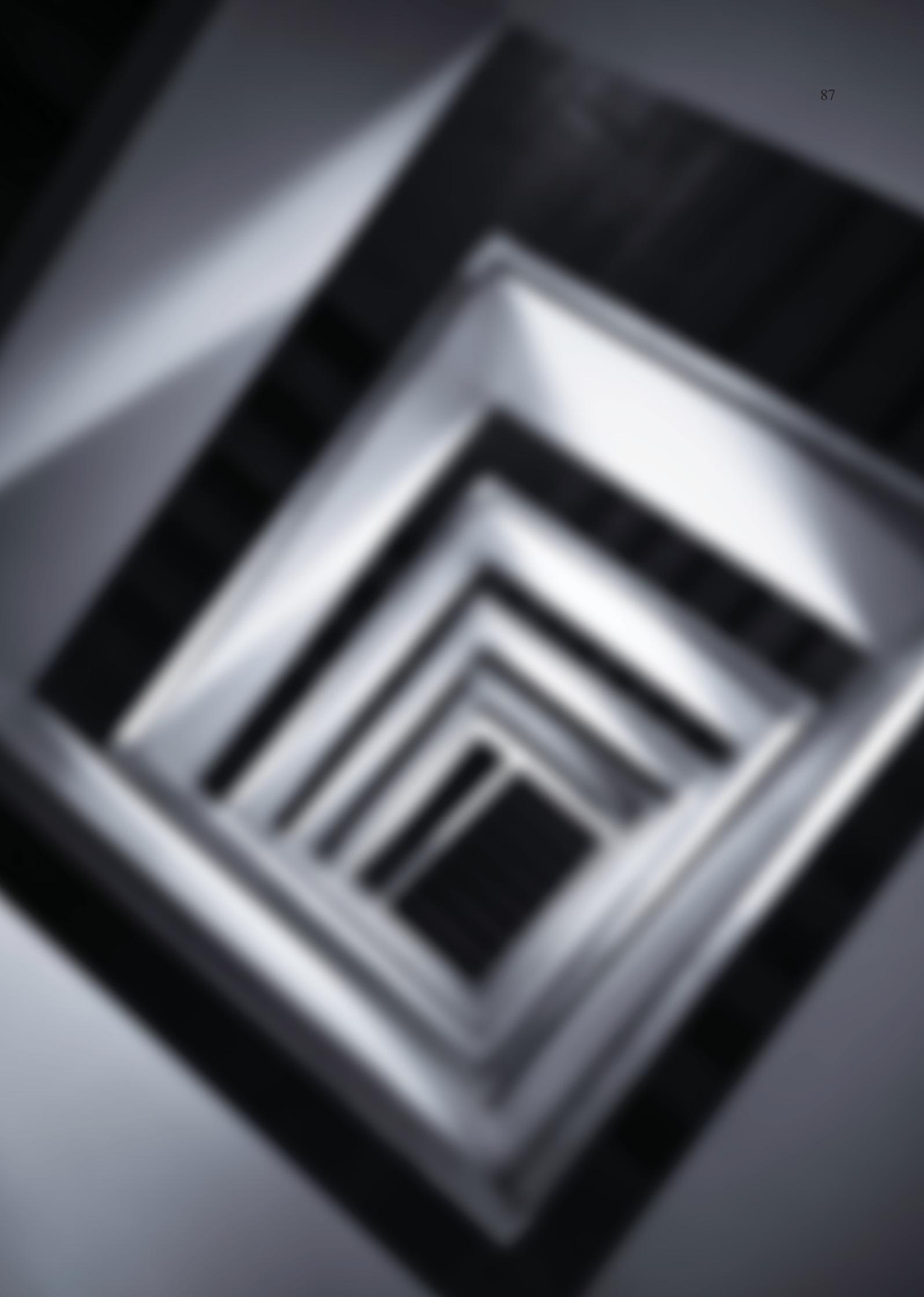
4. Finalizando o processo

A versão atual do *framework* conceitual é composta de 3 artefatos: Diagrama de Partes Interessadas estendido com os Tipos de Parte Interessada; Histórias de Usuário estendida com os aspectos identificados na literatura de IHD; e o Ciclo de Vida dos Dados estendido com a Escada Semiótica.

Após a realização de todas as etapas, será possível encontrar as partes interessadas, quais delas são mais relevantes para a solução do problema, o que fazem e quais seus interesses, além de identificar os problemas em diferentes camadas!

A seguir, é possível observar uma tabela para preencher as histórias de usuário selecionadas.

Escada Semiótica	Ciclo de vida dos dados			
	Produção e coleta de dados (ex. recepção dos dados de sensores, pesquisas ou bases)	Transformação e filtragem de dados (ex. limpeza e seleção dos dados recebidos)	Apresentação e interação com dados (ex. interações que permitam que as partes interessadas façam sentido e obtenham novas ideias a partir dos dados)	Manutenção de dados (ex. políticas de armazenamento e exclusão dos dados)
Social (ex. quais as consequências sociais do uso dos dados no projeto)				
Pragmático (ex. quais as intenções que as partes interessadas tem com os dados)				
Semântico (ex. qual o significado de armazenar um tipo de dado)				
Sintático (ex. quais dados podem ser cruzados)				
Empírico (ex. o ruído na transmissão dos dados)				
Físico (ex. hardware necessário)				



APÊNDICE C – QUESTIONÁRIO APLICADO COM A TURMA DE ENGENHARIA DE REQUISITOS DO DINF NO 2º SEMESTRE DE 2018

Segue os dados planilhados do questionário sobre o uso do *framework* conceitual *Human-Data Ideation* conduzido com os alunos da disciplina de Engenharia de Requisitos do DINF no 2º semestre de 2018, nas Tabelas C.1, C.2, C.3 e C.4.

Tabela C.1: Primeira parte do questionário aplicado com 3 alunos da turma de Engenharia de Requisitos do DINF no 2º semestre de 2018.

	Participante 1	Participante 2	Participante 3
Carimbo de data/hora	<i>24/09/2018 12:05:37</i>	<i>05/11/2018 00:46:01</i>	<i>05/11/2018 16:26:17</i>
Qual a sua idade?	25	22	22
É a primeira vez que você cursa a disciplina de Engenharia de Requisitos?	Sim	Sim	Sim
Avalie a frase: Eu tenho familiaridade com o levantamento (elicitação) de requisitos de software	Discordo fortemente	Concordo	Concordo
Eu tenho familiaridade com o desenvolvimento de software	Indiferente	Concordo	Concordo fortemente
Você já trabalhou em algum projeto de desenvolvimento de software? (Seja na Indústria ou Academia)	Sim	Sim	Sim
Se sim, o que você fez e qual sua opinião sobre desenvolver software?	Desenvolvi um plugin com base em especificações dadas de maneira gradual. O projetista era também o usuário do plugin e tínhamos contato recorrente.	Desenvolvi uma IDE para uma linguagem interna da empresa. Gosto muito de desenvolver.	Tanto back-end para tratamentos do dado para o usuário quanto algumas telas mobile, sendo que ambas são igualmente importantes apesar de termos na universidade uma visão que por vezes leva a desprezar o front-end.
Você participou da atividade de Partes Interessadas em sala	Sim	Sim	Sim
A identificação de partes interessadas ajuda no entendimento do problema	Concordo	Concordo fortemente	Concordo
O Diagrama ajuda a identificar partes interessadas relevantes	Concordo	Concordo fortemente	Concordo fortemente
Foi fácil usar o Diagrama	Indiferente	Concordo	Concordo
Os tipos de parte interessada (Primário, Secundário, Terciário) ajudam a identificar partes interessadas relevantes	Concordo fortemente	Concordo	Indiferente

Tabela C.2: Segunda parte do questionário aplicado com 3 alunos da turma de Engenharia de Requisitos do DINF no 2º semestre de 2018.

	Participante 1	Participante 2	Participante 3
Você conhece outras formas de identificar partes interessadas?	Não	Não	Não
Se sim, quais formas?			
Qual sua opinião sobre este passo?	A atividade realmente ajuda a tentar perceber stakeholders que poderiam não ser percebidos. Ou seja, nos lembra de tentar achar implicações da solução nos mais diversos níveis.		Importante para a elaboração do todo, já que através dele escolheremos as partes que nosso software irá 'focar'.
Você participou da atividade de Histórias de Usuário em sala?	Sim	Sim	Não
As histórias de usuário ajudam a caracterizar as partes interessadas	Concordo	Concordo fortemente	Concordo fortemente
A caracterização das partes interessadas ajuda no entendimento do problema	Concordo fortemente	Concordo fortemente	Concordo fortemente
Foi fácil usar as histórias de usuário	Indiferente	Concordo fortemente	Indiferente
Os aspectos da literatura ajudam na caracterização das partes interessadas	Indiferente	Concordo	Indiferente
Você conhece outras formas de caracterizar as partes interessadas?	Não	Não	Não
Se sim, quais formas?			
Qual sua opinião sobre este passo?			
Você participou da atividade com a Escada Semiótica em sala?	Sim	Sim	Sim
Refinar o problema em diferentes níveis de abstração ajuda no entendimento do problema	Concordo fortemente	Concordo	Concordo

Tabela C.3: Terceira parte do questionário aplicado com 3 alunos da turma de Engenharia de Requisitos do DINF no 2º semestre de 2018.

	Participante 1	Participante 2	Participante 3
A Escada Semiótica (6 níveis) ajuda a visualizar o problema em diferentes níveis de abstração	Concordo fortemente	Concordo	Concordo
O ciclo de vida dos dados ajuda a visualizar o problema em diferentes níveis de abstração	Concordo fortemente	Concordo	Concordo fortemente
Foi fácil usar a Escada Semiótica	Discordo	Discordo	Concordo
Você conhece outras formas de refinar o problema?	Não	Não	Não
Se sim, quais formas?			
Qual sua opinião sobre este passo?	<p>Acredito que essa atividade foi uma das que mais se aproximaram no nível técnico. Portanto, a primeira vista é uma das atividades que mais gera resultados que gerariam ações. No entanto, eu entendo que ela na realidade me ajudou a aprofundar em um determinado caso de usuário, ou seja, ver de diferentes maneiras somente um dos casos de usuário. Ela ajuda a perceber várias funcionalidades e possíveis problemas, mas é difícil se adaptar a preencher a tabela. Talvez essa dificuldade é que traga os resultados dessa atividade.</p>		<p>Poderia ser destrinchado em mais passos de forma a facilitar certas abstrações.</p>
Quais passos foram úteis para entender o problema e levantar (elicitar) requisitos?	<p>Passo 1 - Diagrama de Partes Interessadas com Tipos de Parte Interessada, Passo 2 - Histórias de usuário com aspectos importantes de Interação Humano-Dados (IHD), Passo 3 - Escada Semiótica estendida com Ciclo de vida dos dados</p>	<p>Passo 1 - Diagrama de Partes Interessadas com Tipos de Parte Interessada, Passo 2 - Histórias de usuário com aspectos importantes de Interação Humano-Dados (IHD)</p>	<p>Passo 1 - Diagrama de Partes Interessadas com Tipos de Parte Interessada, Passo 2 - Histórias de usuário com aspectos importantes de Interação Humano-Dados (IHD), Passo 3 - Escada Semiótica estendida com Ciclo de vida dos dados</p>

Tabela C.4: Quarta parte do questionário aplicado com 3 alunos da turma de Engenharia de Requisitos do DINF no 2º semestre de 2018.

	Participante 1	Participante 2	Participante 3
Ao entender o problema e levantar requisitos, é importante conhecer as perspectivas das múltiplas partes interessadas envolvidas	Concordo	Concordo fortemente	Concordo fortemente
O Framework favoreceu o entendimento do problema por múltiplas perspectivas	Concordo fortemente	Concordo fortemente	Concordo
O Framework foi efetivo ao apoiar o levantamento de requisitos do SIM-Transparência	Concordo	Concordo fortemente	Concordo
Quais aspectos foram úteis no contexto de IHD/SIM-Transparência	Legibilidade, Awareness, Transparência, Relacionamentos, Agência, Consentimento Informado, Negociabilidade, Ética	Legibilidade, Awareness, Transparência, Ética	Awareness, Relacionamentos, Agência, Colaboração, Ética
É preciso muito esforço mental para usar o Framework	Concordo	Indiferente	Indiferente
O que foi útil, além do Framework, para você entender o problema e levantar requisitos para o SIMTransparência?			
Eu autorizo a utilização das minhas respostas para análise e disseminação de resultados para fins exclusivamente acadêmicos. (Nenhum dado pessoal será divulgado)	Sim	Sim	Sim

APÊNDICE D – QUESTIONÁRIO APLICADO COM COLEGAS DE PÓS-GRADUAÇÃO DO AUTOR EM 14/12/2018

Segue os dados planilhados do questionário sobre o uso do *framework* conceitual *Human-Data Ideation* no contexto de *redesign* da ferramenta MaRS, conduzido com os colegas de pós-graduação do autor no DINF em 14/12/2018, nas Tabelas D.1, D.2, D.3, D.4, D.5, D.6, D.7, D.8, e D.9.

Tabela D.1: Primeira parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.

	Colega 1	Colega 2	Colega 3	Colega 4	Colega 5
Carimbo de data/hora	14/12/2018 16:28:22	17/12/2018 18:13:10	19/12/2018 13:59:41	19/12/2018 16:44:40	20/12/2018 10:21:30
Qual a sua idade?	30	24	25	31	24
Qual a sua formação?	Mestrado	Graduação	Mestrado	Mestrado	Graduação
Você já participou de alguma atividade relacionada com IHD ou com o Framework Conceitual de IHD?	Não	Sim, uma vez	Não	Não	Não
Qual sua familiaridade com o Mapeamento (ou Revisão) Sistemático de Literatura antes da atividade de Redesign?	Executei mais de 1 mapeamento (ou revisão)	Executei 1 mapeamento (ou revisão)	Executei 1 mapeamento (ou revisão)	Executei 1 mapeamento (ou revisão)	Executei 1 mapeamento (ou revisão)
Qual sua familiaridade com a ferramenta MaRS antes da atividade de Redesign?	Naveguei pelo sistema sem executar um mapeamento (ou revisão)	Estou executando um mapeamento (ou revisão) com auxílio da ferramenta	Estou executando um mapeamento (ou revisão) com auxílio da ferramenta	Naveguei pelo sistema sem executar um mapeamento (ou revisão)	Ouvi falar
Avalie a frase: Eu tenho familiaridade com o entendimento e identificação de requisitos de software	Concordo fortemente	Concordo	Indiferente	Concordo	Concordo
Eu tenho familiaridade com o desenvolvimento de software	Concordo fortemente	Concordo fortemente	Indiferente	Concordo	Concordo

Tabela D.2: Segunda parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.

	Colega 1	Colega 2	Colega 3	Colega 4	Colega 5
Você já trabalhou em algum projeto de desenvolvimento de software? (Seja na Indústria ou Academia)	Sim	Sim	Sim	Sim	Não
Se sim, o que você fez e qual sua opinião sobre desenvolver software?	Analista de sistema (requisitos e diagramas) e designer de interfaces (protótipos)	Desenvolvimento de software para serviços governamentais de previdência. Desenvolvimento de software é uma atividade que envolve muito do conhecimento técnico do programador, dos requisitos de negócio, da criatividade e dos recursos de tempo e pessoal disponíveis.	Desenvolvi alguns projetos para as disciplinas cursadas. O processo necessário para o desenvolvimento precisa ser bem detalhado, uma vez que enquanto desenvolvedor temos uma visão e usuário outra. Vale lembrar também, que a maneira de interação é diferente, então, precisamos sempre refletir sobre o olhar do usuário.	Desenvolvimento de sistema web para gerenciamento de bolsas de Iniciação Científica. Desenvolvimento desde a ideia inicial até o software implantado e funcionando. Minha opinião é: Desenvolver softwares tem momentos legais	
Você participou da atividade de Partes Interessadas?	Sim	Sim	Sim	Sim	Parcialmente
A identificação de partes interessadas ajuda no entendimento do problema	Concordo fortemente	Concordo	Concordo fortemente	Concordo	Concordo fortemente

Tabela D.3: Terceira parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.

	Colega 1	Colega 2	Colega 3	Colega 4	Colega 5
O Diagrama ajuda a identificar partes interessadas relevantes	Concordo fortemente	Concordo fortemente	Concordo	Concordo	Concordo
Foi fácil usar o Diagrama	Concordo	Concordo	Indiferente	Concordo	Concordo
Os tipos de parte interessada (Primário, Secundário, Terciário) ajudam a identificar partes interessadas relevantes	Concordo	Concordo	Indiferente	Concordo	Concordo
Você conhece outras formas de identificar partes interessadas?	Sim	Não	Não	Sim	Não
Se sim, quais formas?	Conhecimento de perfis de usuários, etapas de preparação de desenvolvimento de sistemas, engenharia de software			Observação do contexto ao qual desejo desenvolver o software. Imersão na cultura do usuário.	

Tabela D.4: Quarta parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.

	Colega 1	Colega 2	Colega 3	Colega 4	Colega 5
Qual sua opinião sobre este passo?	Relevante para conhecimento dos usuários finais e demais partes interessadas e seus contextos de usos.	Importante para formar uma caracterização das pessoas afetadas ou interessadas pela tecnologia. Entretanto, deve ser retomada nas etapas posteriores de design para se tornar mais efetiva.	É importante para verificar quem irá utilizar o sistema, e preocupar-se com ele, sua interação, seu conhecimento e as necessidades.	Útil.	Importante, pois facilita a visualização e encontramos os <i>stakeholders</i> de cada grupo.
Você participou da atividade de Histórias de Usuário?	Sim	Sim	Sim	Sim	Sim
As histórias de usuário ajudam a caracterizar as partes interessadas	Concordo fortemente	Concordo	Concordo	Concordo	Concordo
A caracterização das partes interessadas ajuda no entendimento do problema	Concordo fortemente	Concordo	Concordo	Concordo	Concordo fortemente
Foi fácil usar as histórias de usuário	Concordo fortemente	Concordo fortemente	Indiferente	Indiferente	Concordo
Os aspectos da literatura ajudam na caracterização das partes interessadas	Concordo fortemente	Concordo	Indiferente	Indiferente	Concordo

Tabela D.5: Quinta parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.

	Colega 1	Colega 2	Colega 3	Colega 4	Colega 5
Você conhece outras formas de caracterizar as partes interessadas?	Sim	Sim	Não	Sim	Não
Se sim, quais formas?	Template para conhecimento de perfis de usuários da engenharia semiótica.	Mapa da Empatia. Business Model Canvas. Personas.		Observação do contexto ao qual deseja-se projetar o software. Imersão no contexto desejado.	
Qual sua opinião sobre este passo?	Relevante para se obter as características das partes interessadas, que vai além dos usuários finais, englobando assim, todo o contexto social, que é o mundo me que estão inseridos e vivem.	Este passo deixa mais concreto a identificação da parte interessada relevante para o meu projeto. É aqui que o projetista pode dar forma à parte interessada, aproximando-o à sua solução de software.	É um passo importante para identificar as tarefas que os principais usuários podem realizar no sistema, e também, validar as ações.	Útil	Necessário, pois através dele fica mais claro saber o que cada um irá fazer ou necessita fazer como parte de sua função do trabalho.
Você participou da atividade com a Escada Semiótica?	Sim	Sim	Sim	Sim	Sim
Refinar o problema em diferentes níveis de abstração ajuda no entendimento do problema	Concordo fortemente	Concordo	Indiferente	Indiferente	Concordo

Tabela D.6: Sexta parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.

	Colega 1	Colega 2	Colega 3	Colega 4	Colega 5
A Escada Semiótica (6 níveis) ajuda a visualizar o problema em diferentes níveis de abstração	Concordo fortemente	Concordo	Discordo	Concordo	Concordo
O ciclo de vida dos dados ajuda a visualizar o problema em diferentes níveis de abstração	Concordo fortemente	Concordo	Indiferente	Concordo	Concordo
Foi fácil usar a Escada Semiótica	Concordo fortemente	Discordo	Discordo fortemente	Discordo fortemente	Discordo fortemente
Você conhece outras formas de refinar o problema?	Sim	Sim	Não	Sim	Não
Se sim, quais formas?	Artefato quadro de avaliação (QA).	Cenários, Investigação Apreciativa		Observação do contexto ao qual se deseja projetar o software, documentação das observações.	

Tabela D.7: Sétima parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.

	Colega 1	Colega 2	Colega 3	Colega 4	Colega 5
Qual sua opinião sobre este passo?	Relevante para identificar os problemas e pensar em possíveis soluções.	Parece ser relevante e pode ajudar a pensar, mas atividade se tornou difícil de abstrair no preenchimento, por envolver conceitos da semiótica com ciclo de vida de dados. A partir do treinamento do projetista com essa ferramenta/técnica, o preenchimento deve ser feito mais facilmente. Mas é uma atividade robusta, mas complexa de se realizar, com 24 pontos possíveis de análise.	Faltou entendimento e perícia na hora de executar esse passo. Gerou muitas dúvidas na hora de preencher.	Difícil de usar.	Achei muito confuso.
Quais passos foram úteis para entender o problema e levantar requisitos?	Passo 1 - Diagrama de Partes Interessadas com Tipos de Parte Interessada, Passo 2 - Histórias de usuário com aspectos importantes de Interação Humano-Dados (IHD), Passo 3 - Escada Semiótica estendida com Ciclo de vida dos dados	Passo 1 - Diagrama de Partes Interessadas com Tipos de Parte Interessada, Passo 2 - Histórias de usuário com aspectos importantes de Interação Humano-Dados (IHD), Passo 3 - Escada Semiótica estendida com Ciclo de vida dos dados	Passo 1 - Diagrama de Partes Interessadas com Tipos de Parte Interessada	Passo 1 - Diagrama de Partes Interessadas com Tipos de Parte Interessada, Passo 2 - Histórias de usuário com aspectos importantes de Interação Humano-Dados (IHD)	Passo 1 - Diagrama de Partes Interessadas com Tipos de Parte Interessada, Passo 2 - Histórias de usuário com aspectos importantes de Interação Humano-Dados (IHD)

Tabela D.8: Oitava parte do questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.

	Colega 1	Colega 2	Colega 3	Colega 4	Colega 5
Ao entender o problema e levantar requisitos, é importante conhecer as perspectivas das múltiplas partes interessadas envolvidas	Concordo fortemente	Concordo fortemente	Indiferente	Concordo fortemente	Concordo fortemente
O <i>Framework</i> favoreceu o entendimento do problema por múltiplas perspectivas	Concordo fortemente	Concordo fortemente	Discordo	Indiferente	Concordo fortemente
O <i>Framework</i> foi efetivo ao apoiar o levantamento de requisitos da MaRS	Concordo fortemente	Concordo	Discordo	Indiferente	Concordo fortemente
Quais aspectos foram úteis no contexto de IHD/MaRS	Relacionamentos, Autonomia, Colaboração, Privacidade, Ética, Compartilhamento, conversação e acessibilidade.	Awareness, Open Data, Transparência, Agência, Autonomia, Colaboração, Negociabilidade, Privacidade	Relacionamentos, Autonomia, Colaboração	Awareness, Relacionamentos, Autonomia, Colaboração, Privacidade, Ética	Legibilidade, Awareness, Open Data, Transparência, Relacionamentos, Autonomia, Consentimento Informado, Colaboração, Negociabilidade, Privacidade, Ética

Tabela D.9: Questionário aplicado com colegas de pós-graduação do autor em 14/12/2018.

	Colega 1	Colega 2	Colega 3	Colega 4	Colega 5
É preciso muito esforço mental para usar o <i>Framework</i>	Indiferente	Indiferente	Concordo	Concordo	Indiferente
O que foi útil, além do <i>Framework</i>, para você entender o problema e levantar requisitos para a MaRS?	Conhecer sobre os conceitos e características de IHD, pois eu não conhecia essa área.	Foi útil entender qual o propósito da MaRS, e qual sua estratégia para o futuro uso.	Faltou informação sobre o que era de fato o <i>framework</i> e a utilização dele. Talvez um exemplo de como conduzir ajudaria no momento que cada fosse utilizar ele. Ficou vaga a explicação, não apoiando as atividades depois.	Achei interessante a discussão com colegas enquanto usa-se o <i>framework</i> .	
Eu autorizo a utilização das minhas respostas para análise e disseminação de resultados para fins exclusivamente acadêmicos. (Nenhum dado pessoal será divulgado)	Sim	Sim	Sim	Sim	Sim

APÊNDICE E – ARTEFATOS PREENCHIDOS COM A TURMA DE ENGENHARIA DE REQUISITOS DO DINF NO 2º SEMESTRE DE 2018

Segue os dados planilhados dos artefatos "Ciclo de Vida dos Dados estendido com a Escada Semiótica" preenchidos ao fazer uso do *framework* conceitual *Human-Data Ideation* no contexto da ferramenta SIMTransparência, conduzido com a turma de Engenharia de Requisitos do DINF no 2º semestre de 2018.

Ciclo de vida dos dados			
	produção e coleta de dados (e.g. recepção dos dados de sensores, pesquisas ou bases)	transformação e filtragem de dados (e.g. limpeza e seleção dos dados recebidos)	apresentação e interação com dados (e.g. interações que permitam que as partes interessadas façam sentido e obtenham novas ideias a partir dos dados)
Escada Semiótica	As consequências positivas caso seja comprovado o funcionamento eficaz do projeto, o Governo Federal passará a produzir dados de melhor qualidade para facilitar o combate a corrupção e mau uso de dinheiro público. Entretanto, há possibilidade do mesmo ser tratado de forma negativa, reduzindo investimentos e melhorias na qualidade dos dados produzidos	A forma como os dados são filtrados e apresentados implica diretamente no foco em que o mp vai agir. As diferentes construções das informações cria narrativas diferentes impactando diretamente na percepção pública das entidades relacionadas.	manutenção de dados (e.g. políticas de armazenamento e exclusão dos dados)
Social (e.g. quais as consequências sociais do uso dos dados no projeto)			A relação do MP com diversas parcelas da população depende da atuação do mesmo, de acordo com os dados disponíveis pelo órgão.
Pragmático (e.g. quais as intenções que as partes interessadas tem com os dados)	Fiscalização para uma boa manutenção das contas públicas	Construir relatórios eficientes	Fiscalização de irregularidades
Semântico (e.g. qual o significado de armazenar um tipo de dado)	Manter um controle dos gastos.	Limitar a possibilidade de alteração de dados para evitar mascarar um gasto indevido	Ter uma maior eficiência no trabalho do MP e sinalizar o seu interesse na transparência pública
Sintático (e.g. quais dados podem ser cruzados)	Poder cruzar vários tipos de dados para realizar consultas específicas		- Configuração de notificações - Apresentação dos dados em forma de gráfico - Pesquisa de Dados
Empírico (e.g. o ruído na transmissão dos dados)	Cuidado com arredondamentos e ordens de grandezas	Construção de gráficos de forma imparcial, tomando cuidado para não influenciar a narrativa com vies políticos ou pessoais	- Cuidado com arredondamentos e ordens de grandeza - Considerações com taxas de erro aceitáveis
Físico (e.g. hardware necessário)	Banco de Dados e Servidor Web		- Servidor para armazenar os dados - Servidor Web para hospedar a aplicação
<i>Instruções: Escolher uma história de usuário e preencher as células com problemas e questões relativos a história e a linha/coluna.</i>			
História de usuário: Eu como Ministério Público quero ser notificado e ter comprovantes de gastos absurdos de dinheiros público para poder denunciá-los e oprimi-los.			

Ciclo de vida dos dados				
	produção e coleta de dados (e.g. recepção dos dados de sensores, pesquisas ou bases)	transformação e filtragem de dados (e.g. limpeza e seleção dos dados recebidos)	apresentação e interação com dados (e.g. interações que permitam que as partes interessadas façam sentido e obtenham novas ideias a partir dos dados)	manutenção de dados (e.g. políticas de armazenamento e exclusão dos dados)
Escada Semiótica				
Social (e.g. quais as consequências sociais do uso dos dados no projeto)		Dar opções ao usuário portador de deficiência de filtrar os dados de forma rápida e que seja facilmente entendido pelo mesmo dentro de suas respectivas limitações	Fornecer ferramentas que possibilitem o acesso de diferentes necessidades especiais, como um espectro de cores adequado para os diversos tipos de daltonismo ou um texto amigável à leitura de telas.	É possível monitorar se o acesso eficiente realmente é possível por todos os usuários, confirmando que os dados estão armazenados de forma adequada
Pragmático (e.g. quais as intenções que as partes interessadas tem com os dados)		Pesquisas rápidas de modo que, caso o interesse seja despertado, o usuário possa verificar dados específicos sem maiores problemas	Poder de interagir e analisar os dados de forma clara e entendível	Pesquisas rápidas de modo que, caso o interesse seja despertado, o usuário possa verificar dados específicos sem maiores problemas
Semântico (e.g. qual o significado de armazenar um tipo de dado)	Manter um controle dos gastos.	Limitar a possibilidade de alteração de dados para evitar mascarar um gasto indevido	Facilidade de verificação de irregularidades.	Permitir ao usuário realizar consultas de acordo com a narrativa ao qual ele suporta
Sintático (e.g. quais dados podem ser cruzados)	Poder cruzar vários tipos de dados para realizar consultas específicas		Apresentação dos dados em forma de gráfico. Pesquisa de dados.	
Empírico (e.g. o ruído na transmissão dos dados)	Cuidado com arredondamentos e ordens de grandezas		Construção de gráficos de forma imparcial, tomando cuidado para não influenciar a narrativa com vies políticos ou pessoais	Cuidado com a representação interna no banco de dados e na passagem dos mesmos para o Web Server de modo a não perder informações potencialmente importantes devido a tipos de dados não inteiramente portáveis.
Físico (e.g. hardware necessário)			Banco de Dados e Servidor Web	
<i>Instruções: Escolher uma história de usuário e preencher as células com problemas e questões relativos a história e a linha/columna.</i>				
História de usuário: Eu como público em geral quero ter acesso a informação não obstante a limitações físicas para ter acesso a informação.				

Ciclo de vida dos dados			
	produção e coleta de dados (e.g. recepção dos dados de sensores, pesquisas ou bases)	transformação e filtragem de dados (e.g. limpeza e seleção dos dados recebidos)	apresentação e interação com dados (e.g. interações que permitam que as partes interessadas façam sentido e obtenham novas ideias a partir dos dados)
Escala Semiótica	Social (e.g. quais as consequências sociais do uso dos dados no projeto)		manutenção de dados (e.g. políticas de armazenamento e exclusão dos dados)
	Pragmático (e.g. quais as intenções que as partes interessadas tem com os dados)	Os dados coletados devem ser realmente públicos, relevantes e verdadeiros, especialmente no caso de pessoas físicas	Caso aconteça algum relato incorreto, deve existir alguma maneira da correção de dados e relatórios
Semântico (e.g. qual o significado de armazenar um tipo de dado)		Todos os impactados através da mostra de suas contas públicas devem ser tratados de forma idônea de forma a não ser criada uma narrativa que prejudique ou beneficie alguém ou algum grupo	
Sintático (e.g. quais dados podem ser cruzados)	Poder cruzar vários tipos de dados para realizar consultas específicas		
Empírico (e.g. o ruído na transmissão dos dados)	Cuidado com arredondamentos e ordens de grandezas		Cuidado com a representação interna no banco de dados e na passagem dos mesmos para o Web Server de modo a não perder informações potencialmente importantes devido a tipos de dados não inteiramente portáveis.
Físico (e.g. hardware necessário)		Banco de Dados e Servidor Web	
Instruções: Escolher uma história de usuário e preencher as células com problemas e questões relativos a história e a linha/coluna.			
História de usuário: Eu como servidor público quero ter a opção de ser notificado sobre alterações em informações relacionadas a mim, com o objetivo de estar ciente das informações pessoais que são públicas.			

		Ciclo de vida dos dados									
		apresentação e interação com dados		manutenção de dados							
		produção e coleta de dados de sensores, pesquisas ou bases		transformação e filtragem de dados		e.g. limpeza e seleção dos dados		e.g. interações que permitam que as partes interessadas façam sentido e obtenham novas ideias a partir dos dados		e.g. políticas de armazenamento e exclusão dos dados	
Escada Semiótica											
Social (e.g. quais as consequências sociais do uso dos dados no projeto)	Vai permitir ao C3SL saber informações de interesse do usuário. Talvez seja interessante estabelecer o a não divulgação das ferramentas de filtro do usuário. Ou ainda, termos para a proteção do C3SL caso haja o vazamento não intencional de informações.	Uma possível preocupação, ou seja, algo que devemos tomar cuidado é não facilitar somente a filtragem de um tipo de informação ou uma fonte em particular. Devemos tratar os dados de forma isonômica e ou com critérios de seleção bem definidos.	Caso deixarmos certas informações mais fáceis de serem notificadas e ou filtradas, talvez estejamos privilegiando escornder outras. Ou seja, tomar cuidado para ser o mais imparcial na disposição dos conteúdos dos mais diversos setores.	Cria centralização sobre os gastos públicos e tendências de ações governamentais.	Garantir a atualização do usuário sobre o dados atualizado e/ou alterado por meio de e-mails.	Uma pessoa mais informada pode tomar decisões mais acertativas. Por exemplo, votar com maior consciência e manter ciência de possíveis irregularidades.					
Pragmático (e.g. quais as intenções que as partes interessadas tem com os dados)	Garantir a confiabilidade dos dados.	Permitir ao usuário a especificação das informações de interesse.	Garantir mais fácil compreensão e interpretação dos dados.	Apresentar escolhas de unidades alternativas para apresentação de dados. Ex: certos dados talvez possam ser apresentados com o número de salários mínimos.	Garantir o recebimento e envio da notificação.	Explicações sobre as instâncias e métodos	Especificações por meio do número de documento para não haver ambiguidade	... opções, detalhes de como fazer a ação	... forma de organização e estruturação	Permitir a exclusão de usuários	
Semântico (e.g. qual o significado de armazenar um tipo de dado)	Programação em linguagem Web.	A criação de filtros deve ser permitida através de texto (em português), não somente em linguagem matemática.	O C3SL deve formatar os dados de maneira adequada para consumo do banco de dados utilizado.	Permitir o usuário a criar os filtros de forma fácil.	Por meio protocolos web.	Explicações sobre as instâncias e métodos	Especificações por meio do número de documento para não haver ambiguidade	... opções, detalhes de como fazer a ação	... forma de organização e estruturação	Permitir a exclusão de usuários	
Sintático (e.g. quais dados podem ser cruzados)	O C3SL deve criar maneiras seguras de guardar os dados evitando redundâncias desnecessárias. Com backups de segurança com certa frequência. O C3SL deve ter acesso a internet e uma infraestrutura para o recebimento dos dados.	O C3SL deve suportar a carga de informações necessária.	O C3SL deve receber e interpretar os diferentes tipos de dados a serem utilizados.	O C3SL deve configurar maneiras diferentes para receber e interpretar os diferentes tipos de dados a serem utilizados.	O C3SL deve configurar a base de dados com algoritmos que permitam acessos eficientes as informações guardadas para exclusão.	Configuração do tipo de dado a ser recebido: gráfico, texto, porcentagem.	Quais os passos	... medidas e taxasss frequencial/	... opções, detalhes de como fazer a ação	Permitir a exclusão de usuários	
Empírico (e.g. o ruído na transmissão dos dados)	O C3SL necessita de maquinário que suporte geração, processamento e armazenamento de mensagens para notificação em massa. A fim de atender todos os usuários cadastrados.	O C3SL necessita de infraestrutura para o suporte geração, processamento e armazenamento de mensagens para notificação em massa. A fim de atender todos os usuários cadastrados.	O C3SL necessita de infraestrutura para o suporte geração, processamento e armazenamento de mensagens para notificação em massa. A fim de atender todos os usuários cadastrados.	O usuário necessita ter uma conexão com a internet e uma máquina adequada.	A manutenção da notificação será efetuada por meio do e-mail. Ou seja, a notificação será mantida na caixa de mensagens do usuário a partir das configurações deste. Além disso, seriao limitado se houver critério for atendido.	O C3SL precisa de um maquinário que suporte aos dados vindos do MEC.	Quais os passos	... medidas e taxasss frequencial/	... opções, detalhes de como fazer a ação	Permitir a exclusão de usuários	
Físico (e.g. hardware necessário)					Excesso de dados e ou notificações inativas podem gerar excesso de uso de servidor e ou hardware.						
Instruções: Escolher uma história de usuário e preencher as células com problemas e questões relativos a história e a linha/coluna.											
História de usuário:											
US14. Eu como Ministério Público quero ser notificado e ter comprovantes de gastos absurdos de dinheiro público para poder denunciá-los e oprimi-los											
US27. Eu como servidor público quero ter a opção de ser notificado sobre alterações em informações relacionadas a mim, com o objetivo de estar ciente das informações pessoais que são públicas.											

APÊNDICE F – ARTEFATOS PREENCHIDOS COM COLEGAS DE PÓS-GRADUAÇÃO DO AUTOR EM 14/12/2018

Segue os dados planilhados dos artefatos preenchidos ao fazer uso do *framework* conceitual *Human-Data Ideation* no contexto de *redesign* da ferramenta MaRS, conduzido com os colegas de pós-graduação do autor no DINF em 14/12/2018. A Tabela F.1 sintetiza as Histórias de Usuário criadas na atividade, enquanto nas Tabelas F.2, F.3, F.4, F.5, F.6 estão expostos os artefatos do Ciclo de Vida dos Dados estendido com a Escada Semiótica.

Tabela F.1: Histórias de Usuário criadas com colegas de pós-graduação do autor em 14/12/2018.

Eu, enquanto	quero	para
Autor, Supervisor, Reutilizador	conectar perguntas de extração	responder perguntas relacionadas em caso de preenchimento de uma delas
Autor, Supervisor, Reutilizador	inserir tabela GQM no protocolo de pesquisa	auxiliar a elaboração dos objetivos de pesquisa/mapeamento
Autor, Supervisor, Reutilizador	tornar a ferramenta acessível	que pessoas com deficiência possam utilizá-la
Autor, Supervisor, Reutilizador	que as modificações executadas sejam versionadas	ter acesso ao histórico de mudanças
Autor, Supervisor, Reutilizador	conversar com os outros pesquisadores dentro da ferramenta (chat)	trocar informações de maneira síncrona e assíncrona
Leitor	acessar a(s) pergunta(s) de pesquisa	ter acesso aos dados da pesquisa

Tabela F.2: Primeiro artefato preenchido do Ciclo de Vida dos Dados estendido com a Escada Semiótica, criado com colegas de pós-graduação do autor em 14/12/2018.

História de Usuário: <Eu, enquanto> Autor, Supervisor, Reutilizador <quero> formar um grupo de trabalho <para> trabalhar no mapeamento/revisão em equipe				
	Produção e coleta de dados	Transformação e Filtragem de Dados	Apresentação e Interação com Dados	Manutenção de Dados
Mundo Social	Discordância entre opiniões pode gerar briga e desmotivação		Diferentes ideias de apresentação	Erros de sincronismo na manutenção
Pragmático	Agilidade no mapeamento Aumentar o escopo	Agilidade na interpretação Geração de resultados	Riqueza de opiniões na forma de apresentação	Atualização pode sofrer (?)
Semântico	Melhor e mais profundo entendimento da área mapeada	Melhores resultados, maior riqueza (?)	Geração de resultados mais ricos	Manter atualização pode ter diferenças em termos de importância
Sintático	Protocolo para diminuir os vieses	Política de filtragem	Estatísticas, Gráficos e Resumos	
Empírico	Diferentes interpretações de mundo	Dificuldade de alinhamento e interpretação dos dados	Dificuldade de interpretação dos dados	Dificuldade de sincronismo
Físico	Ferramenta computacional para auxiliar (?)	Computador + Sistema		Protocolo de comunicação e atualização

Tabela F.3: Segundo artefato preenchido do Ciclo de Vida dos Dados estendido com a Escada Semiótica, criado com colegas de pós-graduação do autor em 14/12/2018.

História de Usuário: <Eu, enquanto> Autor, Supervisor, Reutilizador, Leitor, <quero> gerar um relatório com o protocolo de pesquisa <para> imprimi-lo e ter acesso ao protocolo fora da MaRS

	Produção e coleta de dados	Transformação e Filtragem de Dados	Apresentação e Interação com Dados	Manutenção de Dados
Mundo Social		Adicionar mais pesquisas primárias	Divulgação das informações levantadas	Evolução do mapeamento
Pragmático	Coletar informação para verificar as pesquisas na área		Verificar/apoiar as ideias de pesquisa Validar os resultados	
Semântico	Para gerar um relatório completo	Transformar um artigo em respostas p/ as perguntas	Com o relatório consegue-se interpretar melhor os dados levantados	
Sintático	Metadados dos artigos	Respostas das perguntas de extração		Adição de novas publicações
Empírico	Cadastro de dados errados	Validação dos dados	Formatação dos dados p/ gerar relatório	Possibilidade de alterar e realizar novos relatórios
Físico	Mecanismos de busca	Formulário de extração		

Tabela F.4: Terceiro artefato preenchido do Ciclo de Vida dos Dados estendido com a Escada Semiótica, criado com colegas de pós-graduação do autor em 14/12/2018.

História de Usuário: <Eu, enquanto> Autor, Supervisor, Reutilizador, <quero> inserir a tabela QM no protocolo de pesquisa <para> auxiliar a elaboração dos objetivos de pesquisa				
	Produção e coleta de dados	Transformação e Filtragem de Dados	Apresentação e Interação com Dados	Manutenção de Dados
Mundo Social	Conscientizar pesquisadores	Favorecer tipo de perfis de usuários	Expandir o conhecimento do pesquisador da comunidade de IHC	Impactos na vida das pessoas, por conta da conscientização dos pesquisadores
Pragmático	Estabelecer os objetivos da pesquisa	Filtrar os objetivos da pesquisa	Contribuir com temas de pesquisa	Inserir ideias novas e remover antigas/sem sentido
Semântico	Avanços nas pesquisas acadêmicas	Estabelecer normas para criar os objetivos	Estabelecer início, meio e fim da pesquisa	Atualizar os objetivos da pesquisa depois de já definidos
Sintático	Os objetivos da pesquisa	Diferentes objetivos de diferentes pesquisadores	Ter acesso a diferentes tipos de arquivos (PDF, DOC, etc)	Permissão ou não para excluir dados de outros pesquisadores
Empírico	Diferentes tipos de dados (DOC, PDF, vídeos)	Opções para seleção do dado (PDF, DOC, etc)	Transmissão de áudio para conversação	Opções de excluir os tipos de mídia
Físico	Espaço de armazenamento grande	Suporte a HD externo	Adaptabilidade (PC, Tablet, Smartphone)	Armazenar diferentes tipos de mídia

Tabela F.5: Quarto artefato preenchido do Ciclo de Vida dos Dados estendido com a Escada Semiótica, criado com colegas de pós-graduação do autor em 14/12/2018.

História de Usuário: <Eu, enquanto> Autor, Supervisor, Reutilizador, <quero> cruzar perguntas de pesquisa (para?)				
	Produção e coleta de dados	Transformação e Filtragem de Dados	Apresentação e Interação com Dados	Manutenção de Dados
Mundo Social	O mapeamento gerará informações novas e conhecimento aplicável em novas pesquisas	Dados podem ser reanalisados, replicados, criando um cenário de transparência	Redução de mapeamentos sem explorar bem os dados	Proveniência dos dados deve estar clara/de maneira exata
Pragmático	Gerar novos conhecimentos Criar argumentos	Intenção de gerar conhecimento rastreável	Dados cruzados serem exportáveis e entendíveis	Atualização rastreável nas mudanças dos dados cruzados
Semântico	Representa conhecimento vindo de várias fontes, devem estar corretas	Fazer sentido de relações de dados ocultos	Podem ser o conhecimento de maior atenção p/ os pesquisadores	
Sintático	Quaisquer dados devem ser cruzados		Dados cruzados apresentados em tabelas ou gráficos	
Empírico	Conhecimento deve ser representável em diversos canais, como texto e imagem	Dados possuem histórico e proveniência	Apresentação com características estéticas de fácil compreensão	Tabela CSV somente dos dados cruzados pode ser gerada e exportada
Físico	Capacidade de processamento necessária	Consultas SQL	Bibliotecas para geração de gráficos, tabelas a partir de dados ou consultas SQL	

Tabela F.6: Quinto artefato preenchido do Ciclo de Vida dos Dados estendido com a Escada Semiótica, criado com colegas de pós-graduação do autor em 14/12/2018.

História de Usuário: <Eu, enquanto> Autor, Supervisor, Reutilizador, <quero> tornar a ferramenta acessível <para> que pessoas com deficiência possam utilizá-la				
	Produção e coleta de dados	Transformação e Filtragem de Dados	Apresentação e Interação com Dados	Manutenção de Dados
Mundo Social	Maior acessibilidade para os usuários	Maior acessibilidade para os usuários	Maior acessibilidade para os usuários	Maior acessibilidade para os usuários
Pragmático	Antes alguns não conseguiam acessar o sistema, agora ele se tornou acessível	Conseguem acessar agora somente informações necessárias	O sistema além de ficar mais acessível, vai ficar mais claro para o usuário	Tornar o sistema acessível
Semântico	Tornar o sistema acessível	Tornar o sistema mais acessível e mais <i>clean</i>		
Sintático	Dados de texto e voz			
Empírico	Ambiente onde o usuário está utilizando o sistema			
Físico	Vai depender da deficiência do usuário, talvez teclas alternativas, fones de ouvido	Vai depender da deficiência do usuário	Vai depender da deficiência do usuário	Vai depender da deficiência do usuário