

THIAGO ADRIANO COLETI

**TR-MODEL: UM PERFIL DE APLICAÇÃO DE
METADADOS PARA TRANSPARÊNCIA DE
DADOS PESSOAIS EM APLICAÇÕES DE
SOFTWARE**

São Paulo
2020

THIAGO ADRIANO COLETI

**TR-MODEL: UM PERFIL DE APLICAÇÃO DE
METADADOS PARA TRANSPARÊNCIA DE
DADOS PESSOAIS EM APLICAÇÕES DE
SOFTWARE**

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção
do Título de Doutor em Ciências.

São Paulo
2020

THIAGO ADRIANO COLETI

**TR-MODEL: UM PERFIL DE APLICAÇÃO DE
METADADOS PARA TRANSPARÊNCIA DE
DADOS PESSOAIS EM APLICAÇÕES DE
SOFTWARE**

Versão Corrigida

Versão original encontra-se na unidade que aloja o Programa de Pós-Graduação

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção
do Título de Doutor em Ciências.

Área de Concentração:
Engenharia da Computação

Orientador:
Prof. Dr. Pedro Luiz Pizzigatti.
Corrêa

São Paulo
2020

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, _____ de _____ de _____

Assinatura do autor: _____

Assinatura do orientador: _____

Catalogação-na-publicação

Coleti, Thiago Adriano

TR-Model: Um perfil de aplicação de metadados para transparência de dados pessoais em aplicações de software / T. A. Coleti -- versão corr. -- São Paulo, 2020.

145 p.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo.
Departamento de Engenharia de Computação e Sistemas Digitais.

1.Engenharia de Computação 2.Ciência da Informação 3.Engenharia de Software 4.Interface-Homem Computador 5.Metadados I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Computação e Sistemas Digitais II.t.

Dedicatória

Dedico esta tese à minha esposa, filha, familiares, amigos.

AGRADECIMENTOS

Agradeço, primeiramente, à Deus e a Nossa Senhora Aparecida, por me darem força, coragem, sabedoria, proteção e paciência para conduzir as tarefas do doutorado.

Agradeço à minha esposa Claudete e à minha filha Tereza, por todo o apoio e paciência nos momentos que precisei; por aceitarem minhas constantes viagens e ausências; e, principalmente, por estarem junto comigo todo o tempo, na busca por esse objetivo.

Agradeço à minha mãe Marilda, ao meu pai José Carlos (*in memoriam*), aos meus avós, à minha irmã, cunhados e sogros, que juntos contribuíram de forma significativa nessa caminhada.

Agradeço aos meus orientadores, professores Marcelo Morandini e Pedro Luiz Pizzigatti Corrêa, pela oportunidade do doutorado, pelos ensinamentos, pela amizade e pela oportunidade de integrar o grupo de pesquisa e, crescer pessoal e profissionalmente.

Agradeço, também, à Professora Lúcia Vilela Leite Filgueiras pelas orientações e pela oportunidade de discutir e divulgar esta pesquisa em suas aulas de MBA e graduação.

Agradeço aos meus colegas de grupo de pesquisa, Dr. André Filipe, Dr. Daniel Lins, Dra. Suelane Fontes e Ricardo Martinelli, pelas conversas, pela parceria, pelos auxílios nas atividades e pela amizade que criamos e manteremos.

Agradeço aos meus colegas professores e funcionários da UENP que não mediram esforços nas substituições e trocas de aulas, para que fosse possível conduzir as tarefas do doutorado.

RESUMO

O uso de recursos como Redes Sociais, Aplicativos Móveis, *Websites* e Redes de Sensores permitiu que quantidades massivas de dados sobre comportamentos e preferências fossem produzidos. Esses dados são chamados de dados pessoais e são usados para identificar ou tornar um indivíduo identificável por meio de alguma característica do dado. Os dados pessoais têm alto valor agregado e são utilizados por organizações para traçar perfis e utilizá-los para os mais diversos propósitos. As inúmeras possibilidades de utilização levantaram questões relacionadas à privacidade, segurança e liberdade devido ao fato que os dados pessoais podem proporcionar muitas informações sobre hábitos e rotinas dessas pessoas. Dentre as preocupações relacionadas ao uso dos dados pessoais está a falta de informações, quais sejam, quais dados serão utilizados, quem os utilizará, como isso será feito e com quais objetivos. Essas informações compreendem a Transparência de Dados Pessoais e são pouco ou nada divulgadas aos usuários. Há iniciativas criadas, mas elas apresentam problemas como: falta de padronização, que pode produzir uma Transparência falha, enviesada, pobre ou com dados desnecessários; falta de um critério de exibição adequado para a Transparência, uma vez que o uso dos dados pode envolver fatores técnicos tornando a compreensão por usuários com pouco letramento comprometida. Assim, esta tese apresenta o TR-Model, um conjunto de diretrizes baseadas em Perfil de Aplicação de Metadados para Transparência de Dados Pessoais. Os metadados têm por objetivo determinar o que deve ser mostrado como Transparência de modo que um conjunto mínimo de informações seja gerado, também, descrever como as informações devem ser apresentadas a fim de apoiar a produção de informação de qualidade e apresentadas de forma adequada para o indivíduo. O TR-Model foi desenvolvido com base na análise de leis e regulamentações de uso de dados pessoais, pesquisas científicas e necessidades/interesses de pessoas em relação ao uso de seus dados; é composto por um modelo de domínio, um conjunto de entidades, metadados, metaeventos e descrições e sua validação foi realizada por meio de testes com usuários, questionários e análise de cenários. Os resultados mostraram que o TR-Model foi eficaz nos aspectos de formato de apresentação de informação e, principalmente, no que tange às dimensões de qualidade de informação para o indivíduo.

Palavras-Chave – Interação-Humano Dados; Perfil de Aplicação de Metadados; Transparência de Dados Pessoais; Visualização de Dados por Indivíduos; Qualidade em transparência.

ABSTRACT

The use of resources as Social Networks, Mobile Applications, Websites, Sensor Networks by people has allowed the production of massive amounts of data about their behaviors and preferences. These data are called personal data and are used to identify or to make an individual identifiable by some characteristic of the data. Personal data have high added value and are used by organizations to identify profiles and behaviors of individuals, as well as for various businesses. The countless possibilities of use have raised questions related to the privacy, security and freedom of individuals due to the fact that personal data can provide a lot of information about their habits and routines. One of the concerns related to the use of personal data is the lack of information to individuals about which ones of their personal data are used, as well as who will use them, how they will be used and for what purpose. This information comprises Personal Data Transparency and it is not clear and accessible to users. There are initiatives developed to provide some kind of Transparency for users, but it presents problems such as: lack of standardization of Transparency information which can produce a failure, bias, poor or unnecessary Transparency; and the lack of adequate presentation criteria for once, data usage may involve technical factors and understanding by lay users may be compromised. Thus, this thesis introduces the TR-Model, which is a set of guidelines based on Metadata Application Profile that aims to present a set of metadata for Personal Data Transparency. Metadata aims to determine what should be presented as Transparency in order to ensure a minimum set of information for individuals, and to describe how information should be presented in order to support the production of information in a friendly and quality format. TR-Model was developed based on the analysis of laws and regulations on the use of personal data, scientific research and people's needs / interests regarding the use of their data. TR-Model consists of a domain model, a set of entities, metadata, metaevents, and descriptions. TR-Model validation was performed through user tests, questionnaires and scenario analysis. The results showed that the TR-Model was effective in the aspects of information presentation format and, especially, regarding the dimensions of information quality for the individual.

Keywords – Human-Data Interaction; Metadata Profile Application; Personal Data Transparency; Personal Infovis; User-friendly Transparency

LISTA DE FIGURAS

1	Modelo de Interação-Humano Dados. Adaptado de Mortier et al. (2016)	23
2	Ciclo de vida dos dados pessoais (ROMANSKY, 2014).	25
3	Exemplo de <i>Personal Visualization</i> no aplicativo Sem Dengue	37
4	Estrutura do TR-Model. Do autor.	54
5	Modelo de domínio para o TR-Model. Do autor.	60
6	Metadados e Metaeventos das entidades do TR-Model. Do autor.	66
7	Relação de leituras das PPS com a preocupação com o uso des dados pes- soais - Perfil A. Do autor.	87
8	Expectativas dos participantes em relação à Transparência. Do autor.	89
9	Interface da entidade <i>Purpose of use</i> . Do autor.	92
10	Interface com informações do controlador do propósito de uso (Relação das entidades <i>Purpose of use</i> e <i>Actors</i>). Do autor.	92
11	Interface dos metadados da entidade <i>Personal Data</i> . Do autor.	93
12	Interface dos metaeventos na entidade <i>Personal data</i> . Do autor	93
13	Interface implementando metadados da <i>Transfer</i> . Do autor.	94
14	Interface implementado demais metadados da entidade <i>Transfer</i> . Do autor.	94
15	Interface com os metaeventos da entidade <i>Transfer</i> . Do autor.	95
16	Interface de verificação dos propósitos de uso. Do autor.	95
17	Interface com informações do metaevento <i>Permission of use</i> da entidade <i>Personal Data</i> . Do autor.	96
18	Interface com lista de atores envolvidos no uso do dado pessoal. Do autor.	97
19	Interface com dados de contato do ator conforme metadados da entidade <i>Actors</i> . Do autor.	97
20	Resultados da Q1 considerando todos os cenários. Do autor.	102
21	Resultados da Q1 organizadas por cenários. Do autor.	102

22	Resultados da Q6 considerando todos os cenários. Do autor.	104
23	Resultados das Q6 organizados por cenários. Do autor.	104
24	Resultados da Q2 considerando a soma de todos os cenários. Do autor. . .	106
25	Resultado da Q2 por cenário. Do autor.	106
26	Resultado da Q3 considerando todos os cenário. Do autor.	107
27	Resultado da Q3 por cenário. Do autor.	108
28	Resultado da Q4 considerando todos os cenário. Do autor.	110
29	Resultado da Q4 por cenário. Do autor.	110
30	Resultado da Q5 considerando todos os cenário. Do autor.	113
31	Resultado da Q5 por cenário. Do autor.	113
32	Resultado da Q8 considerando todos os cenário. Do autor.	115
33	Resultado da Q8 por cenário. Do autor.	115
34	Confiança dos participantes na utilização hipotética do TR-Model para Transparência. Do autor.	116
35	Modelagem UML para a GDPR. Extraído de Tom, Sing e Matulevičius (2018).	145

LISTA DE TABELAS

1	Dimensões de <i>Information Quality</i> . Adaptado de Abib (2010) e Pipino, Lee e Wang (2002)	40
2	Artigos selecionados na RS para análise e síntese dos resultados. Do autor.	44
3	Artigos incluídos manualmente na RS. Do autor.	44
4	Ocorrência de dados para Transparência nos artigos analisados. Do autor. .	46
5	Técnicas de apresentação de Transparência. Do autor.	49
6	Matriz de requisitos de Transparência. Do autor.	58
7	Dimensões de IQ selecionadas para TR-Model. Do autor.	68
8	Metadados e Descrição dos metadados da entidade <i>Actors</i> . Do autor	69
9	Descrição dos Metadados da Entidade <i>Personal Data</i> . Do autor.	72
10	Descrição dos Metaeventos para entidade <i>Personal Data</i> . Do autor	73
11	Descrição de Metadados para a Entidade <i>Purpose of use</i> . Do autor.	75
12	Descrição de Metaeventos para a Entidade <i>Purpose of use</i> . Do autor. . . .	76
13	Descrição de Metadados para a Entidade <i>Transfer</i> . Do autor.	79
14	Descrição de Metaeventos para a Entidade <i>Transfer</i> . Do autor.	80
15	Metadados e Descrições para a entidade <i>Agency</i> . Do autor.	81
16	Cobertura do TR-Model em relação à GDPR. Do autor.	84
17	Cobertura do TR-Model em relação à GDPR. Do autor.	85
18	Pontos positivos e negativos das informações sobre o uso dos dados pessoais. Do autor.	88
19	Questões para a avaliação dos cenários. Do autor.	98
20	Perfil dos participantes da validação. Do autor.	100
21	Resultado da análise da dimensão de Compreensão das informações com elementos de IHC. Do autor.	108

LISTA DE SIGLAS

- GDPR: *General Data Protection Regulation*
- LGPD: Lei Geral de Proteção de Dados
- TETs: Transparency Enhancing Tools
- IP: Endereço IP
- IHD: Interação-Humano Dados
- IHC: Interação-Humano Computador
- Infovis: Visualização de Informação
- RS: Revisão Sistemática
- MAP: *Metadata Profile Application* ou Perfil de Aplicação de Metadados
- SWAP: *Scholarly Work Application Profile*
- VOA3R: *The Virtual Open Access Agriculture - Aquaculture Repository*
- FAO: Organização para a Agricultura e Alimentação (FAO)
- PPS: Política de Privacidade e Segurança
- PIPEDA: Personal Information Protection and Electronic Documents Act
- PV: *Personal Visualization*
- PVA: *Personal Visualization Analytics*
- IQ: *Information Quality* ou Qualidade de Informação

SUMÁRIO

1	Introdução	15
1.1	Contextualização	15
1.2	Objetivo	18
1.3	Motivação	18
1.4	Organização	19
2	Revisão Bibliográfica	21
2.1	Interação-Humano Dados	21
2.2	Transparência de Dados Pessoais	27
2.3	Regulamentação para uso de dados pessoais	31
2.3.1	<i>General Data Protection Regulation (GDPR)</i>	31
2.3.2	Lei Geral de Proteção de Dados	32
2.4	Perfil de Aplicação de Metadados	33
2.5	Visualização de dados por indivíduos	35
2.6	Qualidade de Informação	38
3	Revisão Sistemática	42
3.1	Preparação da Revisão	42
3.2	Condução da Revisão	42
3.3	Análise dos conteúdos	45
3.4	Considerações Finais da Revisão Sistemática	50
4	Construção do TR-Model	52
4.1	Estrutura do TR-Model	52
4.2	Requisitos de Transparência de Dados Pessoais	54

4.3	Modelo de Domínio do TR-Model	59
4.3.1	<i>Actors</i>	61
4.3.2	<i>Purpose of use</i>	62
4.3.3	<i>Personal Data</i>	63
4.3.4	<i>Transfer</i>	64
4.3.5	<i>Agency</i>	64
4.3.6	Considerações sobre as entidades do TR-Model	65
4.4	Definição e Descrição dos Metadados e Metaeventos das entidades	65
4.4.1	Metadados e descrições da entidade <i>Actors</i>	68
4.4.2	Metadados, metaeventos e descrições da entidade <i>Personal Data</i> . .	70
4.4.3	Metadados, metaeventos e descrições da entidade <i>Purpose of use</i> . .	73
4.4.4	Metadados, metaeventos e descrições da entidade <i>Transfer</i>	76
4.4.5	Metadados e descrições da entidade <i>Agency</i>	80
4.5	Considerações finais da construção do TR-Model	81
5	Validação	83
5.1	Análise da Cobertura do TR-Model em relação GDPR	83
5.2	Validação do TR-Model com indivíduos	85
5.2.1	Atividades Pré Teste	86
5.2.2	Validação do TR-Model com cenários	91
5.2.3	Estratégia de análise dos dados	101
5.2.4	Resultados das características de IHC	102
5.2.5	Resultados das dimensões de Qualidade de Informação	105
5.2.6	Análise e discussão da confiabilidade dos participantes e sugestões de mudança	115
5.2.7	Considerações sobre a Validação do TR-Model	118
6	Considerações Finais	120

Referências	124
Apêndice A – Protocolo de Revisão Sistemática	132
A.1 Objetivo	132
A.2 Pergunta de Pesquisa	132
A.3 PICOC	132
A.4 Período de publicação considerado para os artigos	133
A.5 Keywords	133
A.6 Critérios de seleção de base de dados	134
A.7 Idiomas	134
A.8 Fontes de Dados e Métodos de busca e fontes de dados	134
A.9 Critérios de Inclusão e Exclusão	135
A.10 Forma de Análise	135
A.11 Critério de qualidade	135
A.12 Extração de Conteúdos	136
A.13 Síntese dos Dados	136
Apêndice B – Tabela de extração de dados da Revisão Sistemática	137
Anexo A – Diretrizes de Transparência de Dados Pessoais da GDPR traduzidas para o Português.	142
Anexo B – Modelagem de Classes e Atributos para a GDPR	144

1 INTRODUÇÃO

Esta seção apresenta a contextualização, objetivos, justificativa e organização deste texto.

1.1 Contextualização

Em um cenário no qual dados pessoais são coletados, processados e compartilhados, frequentemente a privacidade, segurança e liberdade de pessoas podem ser comprometidas. Tais ações acabam por acontecer de forma desconhecida para os indivíduos¹, pois são pouco divulgadas, ou divulgadas com mecanismos complexos - como textos jurídicos ou algoritmos computacionais - o que dificulta a interpretação por uma pessoa leiga².

A utilização dos dados pessoais despertou a preocupação de entidades governamentais e empresas que buscam incentivar e/ou elaborar estratégias para melhorar a Transparência de Dados Pessoais (Transparência) para os indivíduos.

O conceito de Transparência, aqui adotado, considera o esforço em tonar as ações e agentes envolvidos no uso de dados pessoais acessíveis, claros e compreensíveis para o indivíduo garantindo a participação ativa no fluxo de uso de seus dados e dando-lhes condições de agir, negociar e/ou participar do uso de seus dados (MORTIER et al., 2016). O conceito de Transparência na *General Data Protection Regulation* (GDPR) prevê a necessidade de toda a informação destinada aos indivíduos ser concisa, de fácil acesso e compreensível (VOIGT; BUSSCHE, 2017).

Além das características apresentadas, a Transparência deve ser apresentada com linguagem clara e formato de visualização adequado para o indivíduo, isso porque a informação poderá ser manipulada, tanto por pessoas especialistas, quanto por leigas. Assim, a Transparência pode ser vista como um mecanismo de abstrair complexos processos

¹Indivíduo é toda pessoa que interage com algum ambiente computacional e produz dados pessoais. Nesta pesquisa, o indivíduo também poderá ser tratado como Titular dos Dados.

²Nesta pesquisa o termo *pessoa leiga* refere-se às pessoas com pouca habilidade em análise de dados.

computacionais e situações jurídicas em informações simplificadas o que justifica sua forte relação com a área de Interação-Humano Computador (IHC) (HADDADI et al., 2013; VOIGT; BUSSCHE, 2017).

Além do aspecto humano, características técnicas tais como, o grande número de software diferentes, com áreas de ação diferentes, com perfis de usuários diferentes e que utilizam os dados de formas distintas, também devem ser considerados. Essas características levantam questões como: (1) **com tanta diversidade, o que deve ser mostrado ao titular dos dados como Transparência?**; e (2) **e como deve ser mostrada a Transparência para o indivíduo?**.

Para subsidiar a proposição de respostas a tais questões três abordagens para Transparência foram destacadas: (1) as regulamentações de uso de dados pessoais como a própria GDPR proposta pela União Europeia ou a Lei Geral de Proteção de Dados (LGPD) do Brasil; (2) ferramentas de apoio à Transparência chamadas de *Transparency Enhancing Tools* (TETs) que buscam proporcionar mecanismos para auxiliar os indivíduos na compreensão das ações que ocorrem com seus dados; e (3) Políticas de Privacidade e Segurança que são amplamente aplicadas em software por meio de longos textos.

A necessidade de melhoria identificada na análise dessas abordagens, e que direciona esta pesquisa, assenta-se no fato de que as informações sobre Transparência não são padronizadas, simplificadas e com foco no indivíduo. As abordagens, ferramentas ou regulamentações não são unanimes no que tange a um conjunto mínimo de informações sobre o uso dos dados pessoas que poderia ser apresentado ao indivíduo. Em trabalho mostrado por Murmann e Fischer-Hübner (2017a) foi possível identificar que as TETs mostravam diferentes informações, e, em determinados casos, com forte aspecto técnico, por exemplo: o endereço IP do destinatário, histórico de acesso aos dados, dados de *cache* de navegador e localização dos destinatários dos dados.

Já as regulamentações, se comparadas à GDPR e à LGPD, têm uma estrutura similar e complementar, mas a Transparência é tratada de forma diferente nelas. Enquanto a GDPR é mais detalhada, elencando uma série de itens como Transparência, a LGPD é mais genérica e insere a necessidade de informações dentro de outros conteúdos, o que pode requerer uma inferência por parte da pessoa que faz a leitura a fim de identificar qual informação deve ser fornecida ao indivíduo.

A falta de uma padronização, principalmente nos aspectos voltados para o usuário final, podem levar a problemas como:

- Apresentação de um conjunto de informações enviesadas pela aplicação com o objetivo de mostrar o que é de interesse do controlador³ e não do indivíduo;
- Dificuldade ou incapacidade de informar sobre eventos e agentes envolvidos no uso dos dados, uma vez que as ações são comumente computacionais e técnicas;
- Falta de informações relevantes, objetivas e compreensíveis para o indivíduo o que pode levar à falta de confiança na aplicação;
- Dificuldade do indivíduo em identificar ações contra sua vontade/autorização e, consequentemente, agir em sua defesa.

Assim, o principal problema desta pesquisa é a falta de um conjunto mínimo de informações sobre os agentes e eventos envolvidos no uso dos dados pessoais, bem como diretrizes sobre o formato da informação. Considerando essa necessidade, assumiu-se que o uso de estratégia baseada em Perfil de Aplicação de Metadados (*Metadata Application Profile* ou MAP) poderia proporcionar suporte na resolução desses problemas devido ao fato que um MAP auxilia na criação e descrição de metadados para um domínio específico além de tratar de questões como ambiguidades e forma de utilização.

O MAP proposto nesta pesquisa pode/deve ser utilizado por empresas em ações de desenvolvimento de aplicações que precisam ser transparentes para seus usuários. Entretanto, as especificações das entidades, metadados, metaeventos focaram nas necessidades de informações dos indivíduos em conjunto com as regulamentações e uso de dados pessoais como a GDPR e LGDP. Assim, as características das informações de Transparência para os indivíduos foram priorizados em relação à facilidade de implementá-las por parte das equipes de desenvolvimento.

Destaca-se também que nem toda a utilização de dados pessoais é maléfica ao indivíduo, pelo contrário, infinitos são os benefícios que podem ser alcançados para as pessoas. Entretanto, é direito dos titulares dos dados conhecer quem tem acesso aos seus dados e quais eventos são conduzidos e para quais finalidades.

A próxima seção apresenta os objetivos desta tese.

³ Empresa que determinada como os dados pessoais são usados.

1.2 Objetivo

O objetivo geral desta tese foi o de propor um Perfil de Aplicação de Metadados para dar suporte à Transparência de Dados Pessoais. Este perfil de aplicação foi batizado como TR-Model e, de forma geral, buscou sanar: (1) o que deve ser informado como Transparência; e (2) como apresentar a informação de Transparência.

Para atingir o objetivo desta pesquisa, os seguintes objetivos específicos foram traçados:

- Compreender as definições, elementos, parâmetros e requisitos de Transparência com base na análise de regulamentações de dados pessoais, artigos técnicos/científicos e materiais relacionados ao assunto;
- Construir um modelo de domínio para a Transparência de Dados Pessoais;
- Especificar um conjunto de metadados e metaeventos para determinar quais informações apresentar;
- Elaborar um conjunto de descrições que consideraram fatores de Interação-Humano Computador e Qualidade de Informação para especificar como apresentar as informações para os indivíduos;
- Implementar um protótipo de software com base nas especificações do TR-Model a fim de verificar sua aplicabilidade;
- Validar o TR-Model com base em testes de usuários, resolução de questionários e análise do cenário construído;
- Analisar o resultado da validação para considerar a eficácia do TR-Model.

A próxima seção apresenta a motivação da pesquisa.

1.3 Motivação

Para Schneier (2015) o limite do uso dos dados pessoais é o limite da Internet, da capacidade dos algoritmos e da infraestrutura computacional. Essa definição permite assumir que a utilização dos dados pessoais só tende a aumentar e os propósitos de uso se tornarão cada vez maiores, podendo interferir de forma ainda mais significativa na vida dos indivíduos e, consequentemente, na privacidade, segurança e liberdade.

Uma vez que a interferência pode ser positiva ou negativa, o conhecimento por parte dos indivíduos sobre fatores relacionados à utilização de seus dados pessoais se torna fundamental para garantir sua privacidade, segurança e liberdade, pois pode garantir a capacidade do indivíduo monitorar, atuar e negociar a utilização dos dados.

Considerando que uma parcela significativa dos usuários de aplicações de software não são especialistas em análise de dados e/ou textos jurídicos, as informações de uso dos dados pessoais devem ser disponibilizada em uma abordagem voltada para facilitar a utilização do indivíduo.

Assim sendo, esta tese se justificou pela busca em auxiliar na garantia dos direitos dos titulares dos dados em obter conhecimento claro, objetivo e comprehensível sobre eventos e agentes envolvidos no uso de seus dados pessoais.

A próxima seção apresenta a organização da tese.

1.4 Organização

Essa tese está organizada conforme a estrutura de capítulos:

A Introdução apresenta a contextualização do uso dos dados pessoais, a motivação da pesquisa em Transparência e os objetivos buscados durante a execução dessa pesquisa.

O Capítulo 2 apresenta a revisão bibliográfica dos conceitos de Interação-Humano Dados (IHD), Transparência de Dados Pessoais, Regulamentações de uso de dados pessoais com foco na GDPR, Perfil de Aplicação de Metadados, *Information Visualization* (Infovis) e Qualidade de Informação.

O Capítulo 3 apresenta os resultados da Revisão Sistemática (RS) conduzida a fim de identificar o estado da arte no que tange aos modelos de metadados para Transparência de Dados Pessoais e às formas de apresentação da Transparência de Dados Pessoais.

O Capítulo 4 apresenta o processo de desenvolvimento do TR-Model e seus componentes, tais como Modelo de Domínio, Entidades, Metadados, Metaeventos e Descrição dos metadados.

O Capítulo 5 apresenta o processo de validação do TR-Model, os resultados e as discussões sobre o desempenho do modelo para dar suporte à produção de Transparência de Dados Pessoais.

Por último, o Capítulo 6 apresenta as considerações finais, a conclusão desta pesquisa

e proposição de trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

Este capítulo apresenta a Revisão Bibliográfica a começar pela discussão da área de Interação-Humano Dados. Destaca-se que **o uso dos dados pessoais** é um assunto que desperta o interesse científico, técnico, jurídico e jornalístico. Assim, a busca por bibliografias para essa pesquisa não se limitou às bases de dados científicas, mas outros recursos como *websites*, *blogs*, leis e regulamentações foram consultados. A busca por outras fontes deu-se pelo vasto e rico conteúdo apresentado.

2.1 Interação-Humano Dados

A Interação-Humano Dados (IHD) é uma área de estudos relacionada aos elementos envolvidos no fluxo do uso dos dados pessoais e envolve a interação das pessoas com os dados (HADDADI et al., 2013). O termo IHD é relativamente novo no meio científico, embora trabalhos na área de dados abertos, armazenamento de dados e *data logging* tratem de questões semelhantes ao uso de dados pessoais (GURRIN; SMEATON; DOHERTY, 2014; BARRETO; SALGADO; VITERBO, 2018).

Um **dado pessoal** é um registro que permite identificar ou tornar uma pessoa identificável (SCHNEIER, 2015; VOIGT; BUSSCHE, 2017). A identificação pode ocorrer por um identificador específico ou um conjunto de fatores relacionados às suas ações, personalidade, preferências econômicas e culturais (ROMANSKY, 2014).

A IHD é motivada pela grande interação de pessoas com tecnologias como Redes Sociais, Aplicativos Móveis, *Websites*, Redes de Sensores dentre outros e que produzem grandes volumes de dados sobre suas preferências e comportamentos (ABDULLAH; CONTI; BEYAH, 2008; MORTIER et al., 2016).

Mortier et al. (2016) também destacam que as pessoas passaram a viver em um mundo orientado pelos dados. Essa afirmação é justificada pelo aumento de tarefas feitas pelas pessoas com base em informações de aplicativos e *websites*. Por exemplo: (1) o uso de

mapas em que os dados coletados dos usuários ajudam a indicar para outros usuários pontos de congestionamento ou os sentidos das ruas; ou (2) Redes Sociais que aprendem as preferências de seus usuários para direcionar conteúdos como propagandas, *fan pages* ou perfis de outras pessoas.

As diferentes formas de interação de pessoas com os dados permitiram a formulação de diferentes definições de IHD por diferentes autores conforme a lista a seguir:

- IHD está relacionada à manipulação humana, análise e tomada de decisão em bases de dados grandes, complexas e desestruturadas (ELMQVIST, 2011);
- IHD é a entrega de dados personalizados, dentro de contexto e compreensíveis para os usuários (CAFARO, 2012).
- IHD proporciona suporte para recursos a fim de prover acesso e compreensão de dados aos indivíduos e como os dados afetam seu comportamento (MASHHADI; KAWSAR; ACER, 2014);
- IHD é a capacidade de gerenciar diversas e diferentes bases de dados pessoais e habilitar o usuário para controlar eventos de uso dos dados (MCAULEY; MORTIER; GOULDING, 2011).

Embora com certas diferenças, as definições apresentam elementos comuns que permitem montar uma definição mais ampla e aplicável para todos os tipos de sistemas que trabalham com dados pessoais. Os elementos comuns citados são:

- os indivíduos e os dados produzido por ele/ela (*My Data*) ou sobre ele/ela (*Data about me*) são os principais elementos desse complexo sistema;
- proporcionar mecanismos de uso, controle, visualização, análise, entendimento e tomada de decisões de indivíduos sobre as bases de dados é o principal desafio da área;
- os dados de indivíduos podem estar em uma ou várias bases de dados, compostas somente por dados pessoais ou em conjunto com outros tipos de dados (científicos e corporativos).

Portanto, considerando as definições apresentadas estabeleceu-se a seguinte definição para IHD como base para esta pesquisa: *IDH é a área que estuda fenômenos de interação*

entre indivíduos com grandes bases de dados pessoais ou corporativos com objetivo de prover mecanismos para dar suporte à coleta dos mesmos, transparência, compreensão, visualização de informações de indivíduos, por indivíduos e por empresas, a fim de dar suporte à tomada de decisão pessoal ou empresarial. Por bases de dados pessoais contextualiza-se:

- bases que integram dados pessoais de diversas pessoas a fim de estabelecer comparações, agrupamentos dentre outras informações sobre os indivíduos relacionados;
- bases de dados formadas por dados pessoais de um único indivíduo provenientes de fontes distintas e que são utilizados para aprender sobre um indivíduo específico.

O processo de uso dos dados pessoais pode ser ilustrado pelo modelo mostrado na Figura 1.

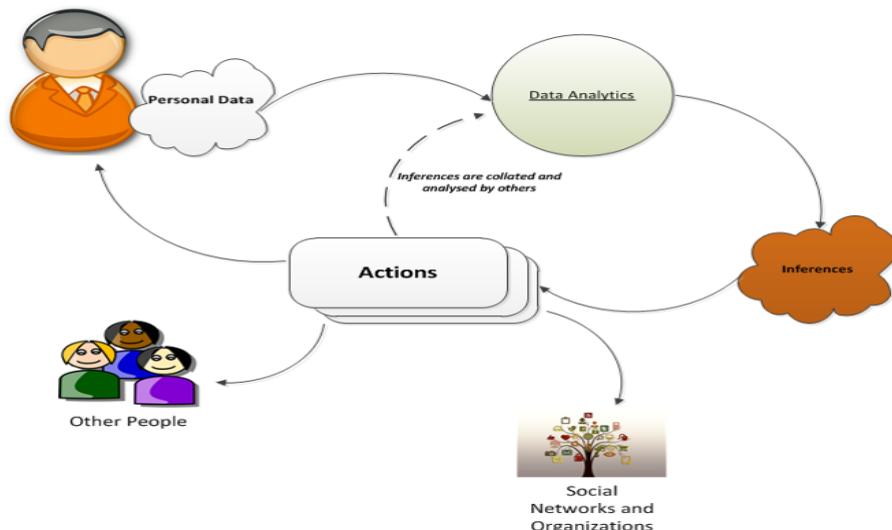


Figura 1: Modelo de Interação-Humano Dados. Adaptado de Mortier et al. (2016)

O modelo contempla: (1) um dado sobre um indivíduo (*Data about me*) ou produzido por um indivíduo (*My Data*) que é constantemente **produzido e coletado** em redes sociais, redes de sensores, sites, aplicativos, sistemas embarcados, vídeo games, sistemas de gestão e outras dezenas de tecnologias e que, de alguma forma, permitem a interação com o indivíduo (MICHAEL et al., 2011); (2) os dados pessoais são **processados** com o apoio de técnicas como *Machine Learning* que auxiliam na identificação, compreensão ou predição de algum comportamento do indivíduo e consequentemente a produção de uma informação sobre ele (LEBO; SUTTI; GREEN, 2016); (3) as informações produzidas são então disponibilizadas e podem ser utilizadas as seguintes maneiras:

1. um indivíduo que produziu um dado sobre sua caminhada, por exemplo, pode utili-

zar a informação para melhorar sua performance, neste caso a abordagem é conhecida como *Personal Informatics* (EPSTEIN, 2015);

2. um indivíduo utiliza a informação produzida por outro indivíduo (CAVANILLAS; CURY; WAHLSTER, 2016). Por exemplo, um indivíduo tira uma foto de um terreno baldio indicando que há um foco de Dengue no local e outro indivíduo utiliza-se dessa informação para decidir se compra ou aluga uma casa naquele bairro;
3. uma empresa utiliza a informação produzida com base nos dados do indivíduo (MORTIER et al., 2016). Essa é uma forma muito comum aplicada pelas mais diversas aplicações. As mesmas aprendem sobre o indivíduo consumidor a fim de oferecer melhorias nos serviços, novos serviços ou produtos e serviços sob demanda.

No fluxo de uso dos dados, o indivíduo atua efetivamente em duas fases: na produção dos dados pessoais e na utilização de informações produzidos pelos seus dados pessoais ou combinados com dados de outros indivíduos. Demais ações são comumente realizadas por ambientes computacionais preparados para tratar os dados pessoais a fim de entregar algum resultado com eles (LEBO; SUTTI; GREEN, 2016).

A existência do dado pessoal dentro do fluxo do uso dos dados pode ser explicado por meio de um ciclo de vida. É comum dentro das áreas de *Big Data* e Ciência dos Dados o uso de um ciclo de vida para os dados. Amaral (2016) propõe um ciclo de vida de dados genérico composto pelas etapas de: Produção, Armazenamento, Transformação, Análise e Descarte. Existem também ciclos de vida que foram criados para atender alguma especificidade na gestão de dados, por exemplo, o ciclo de vida o proposto pelo projeto *Data Observation Network for Earth* (DataONE) (ALLARD, 2012). O ciclo de vida do DataONE possui um número maior de etapas e uma classificação mais rigorosa sobre quais atores atuam em cada etapa.

Para os dados pessoais, o ciclo de vida também sofreu modificações a fim de adequar-se às particularidades desse tipo de dados tais como atender alguma regulamentação de uso dos dados. O ciclo de vida dos dados pessoais proposto por Romansky (2014) pode ser visualizado na Figura 2.

As etapas do ciclo de vida de Romansky (2014) seguem uma abordagem semelhante ao ciclo de vida padrão na Ciência dos Dados e também do *DataONE*. Destaca-se no ciclo de vida dos dados pessoais a interação entre as etapas *Using*, *Preserving* e *Actualization* uma vez que as três etapas atuam em um ciclo fechado após a coleta de dados, diferente dos modelos clássicos em que as etapas atuam em sequência e fecham um ciclo após a

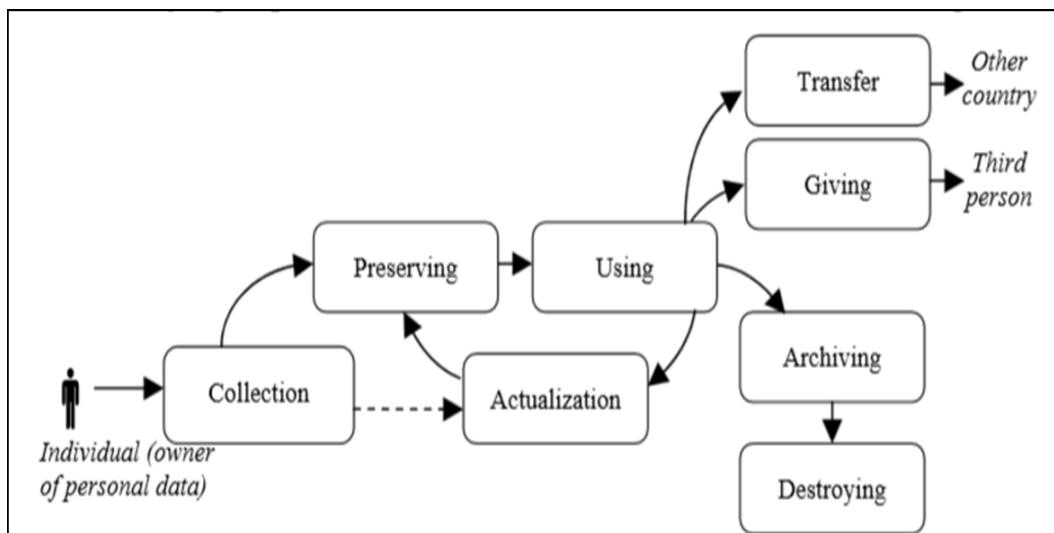


Figura 2: Ciclo de vida dos dados pessoais (ROMANSKY, 2014).

execução da última delas.

Destaca-se também os três diferentes caminhos do dado pessoal após sua utilização pelo controlador que, podem ocorrer pelas seguintes ações:

- Arquivar e destruir (*Archiving and Destroying*): os dados podem ser arquivados, se necessário ou exigido por lei, por um período de tempo. Se os dados não precisam ser mantidos armazenados, os mesmos devem ser eliminados após o cumprimento do objetivo de uso a fim de garantir o titular fique anônimo;
- Transferir (*Transfer*): os dados podem ser transferidos ou compartilhados com controladores de outros países uma vez que os mesmos estejam sob uma regulamentação de uso de dados;
- Disponibilizar (*Given*): os dados podem ser divulgados ou entregues para outras pessoas ou empresas, caso as mesmas também estejam sob vigilância de uma regulamentação

O interesse por processar, aprender e faturar com o uso dos dados pessoais tornou-se tão amplo que, graças às tecnologias existentes como Internet, Redes Sociais, Sistemas Embarcados e Telecomunicação eficaz, é praticamente impossível utilizar algum aplicativo ou um *website* sem ter algum dado pessoal registrado ou utilizado para produzir alguma informação sobre o indivíduo (HADDADI et al., 2013; MAUS, 2015).

Esse cenário não é uma situação restrita, isolada e nem tão pouco temporária. Schneier (2015) relata sobre centenas de iniciativas adotadas há certo tempo para monitorar

e aprender sobre comportamentos das pessoas. Destaca que o uso dos dados pessoais somente tende a crescer e que limita-se, unicamente, aos recursos disponibilizado pela computação e pela Internet.

A particularidade e sensibilidade da IHD fez surgir desafios a serem resolvidos dentro do âmbito de pesquisa, inovação e desenvolvimento. Os desafios estão distribuídos dentre as diversas etapas que envolvem o fluxo dos dados pessoais (BELLAMY; ALONSO, 2016; CRABTREE; MORTIER, 2015). A resolução dos desafios pode estar relacionado a uma única área, mas considerando a multidisciplinaridade de IHD, os avanços nessa área devem ser resultados de integração de áreas como: computação, inteligência artificial, legislação, sociologia e negócios (AMAR; HADDADI; MORTIER, 2016).

Entretanto, desafios são citados por Ackerman e Mainwaring (2005), Crabtree et al. (2016), Haddadi et al. (2013), Mortier et al. (2016) e estão intimamente ligados com a inclusão do indivíduo, como entidade ativa no fluxo do uso dos dados pessoais: Transparência de Dados Pessoais; Capacidade de Ação; e Negociação, Privacidade, Regulamentação e Anonimização.

- **Transparência de Dados Pessoais:** tornar as ações mais claras, acessíveis e compreensíveis de forma que o indivíduo possa compreender o que acontece com seus dados e agir, caso necessário. Dado o viés desta pesquisa, a Transparência será discutida com mais detalhes na Seção 2.2;
- **Capacidade de ação:** refere-se à capacidade de um indivíduo gerenciar a forma como seus dados são utilizados, tratados e inferidos. Não se trata somente de permitir ou não à coleta de dados, mas ao controle e conhecimento sobre a coleta, armazenamento, processamento disponibilizando ao indivíduo e à capacidade de modificar qualquer aspecto relacionado às atividades em seus dados (MORTIER et al., 2016);
- **Negociabilidade:** atribuir poder ao indivíduo de autorizar e retirar parcialmente, ou totalmente, a autorização de uso de seus dados em um determinado processamento. Este aspecto está voltado à questões sociais e de negociação com o utilizador dos dados, dos parâmetros de uso dos dados e da forma como os dados pessoais são integrados à outras bases de dados (HADDADI et al., 2013; MORTIER et al., 2016).
- **Privacidade:** essa é uma das preocupações que mais se destacam e preocupam pessoas, cientistas e desenvolvedores, uma vez que o uso incorreto ou descontrolado dos dados pessoais pode invadir a privacidade e comprometer a segurança e liberdade

dos indivíduos (IACHELLO; HONG, 2007; KUNUNKA et al., 2017; ROMANSKY, 2014);

- **Regulamentação:** a existência das novas regulamentações de uso de dados (a exemplo da GDPR, explicada na Seção 2.3.1) trazem grandes dificuldades no que tange à aceitação, adequação das empresas e fiscalização, a fim de garantir a efetividade da lei/regulamentação (CRABTREE et al., 2016; De Luca, 2019);
- **Anonimização:** com o objetivo de garantir a privacidade, será necessário o desenvolvimento e/ou melhorias de técnicas para garantir a anonimização do indivíduo identificado por um registro de dado (GURRIN; SMEATON; DOHERTY, 2014);

Portanto, são infinitas possibilidades de uso de dados pessoais. A necessidade de *fair use* dos dados por parte das empresas e a necessidade de garantir ao indivíduo titular dos dados maior capacidade de análise, compreensão e ação são preocupações que movem a área de IHD. Os principais conceitos sobre Transparência de Dados Pessoais são discutidos na próxima seção.

2.2 Transparência de Dados Pessoais

O princípio de Transparência de Dados Pessoais, segundo a GDPR, prevê que qualquer informação a respeito de um indivíduo, ou conjunto de indivíduos, deve ser acessível, concisa e fácil de entender (VOIGT; BUSSCHE, 2017). A Transparência deve ser vista com uma propriedade de **visibilidade** onde os obstáculos que impedem o acesso aos detalhes do uso dos dados são mitigados (TURILLI; FLORIDI, 2009).

A Transparência tornou-se um requisito para software que utilizam dados pessoais além de ser exigida por regulamentações como a GDPR (BELLAMY; ALONSO, 2016). Com a Transparência, os elementos e agentes envolvidos no uso dos dados pessoais devem ser acessíveis e compreensíveis aos indivíduos (BONATTI et al., 2017). Assim, a Transparência é fator que estabelece o nível de confiança do indivíduo na aplicação (MURMANN; FISCHER-HÜBNER, 2017a).

Murmann e Fischer-Hübner (2017a) classificam a Transparência em duas categorias. (1) *Ex-ante Transparency*: nessa categoria a Transparência visa garantir ao indivíduo o direito de dar consentimento de uso de seus dados e poder de gerenciar esse consentimento. Essa categoria defende o fato de que o indivíduo deve ter, no momento (ou antes) da coleta dos dados, informações sobre o objetivo de uso dos dados, destinatários dos dados

e seus direitos sobre os dados, inclusive o direito de cancelar ou modificar totalmente ou parcialmente a permissão de uso; e (2) *Ex-post Transparency*: nessa categoria de Transparência, o indivíduo tem o direito de solicitar informações de como seus dados são utilizados, quais processos são realizados. Também pode solicitar cópias de seus dados que devem ser enviados em formato legível por computador. O indivíduo, em posse de cópia de seus dados, pode retransmiti-los para outros controladores conforme seus objetivos.

A preocupação com a Transparência aumenta na mesma proporção em que as preocupações em relação ao uso dos dados pessoais aumentam (CHRISTL, 2017). Em especial, a preocupação com a Transparência deve-se ao fato de que as ações e os agentes envolvidos no uso dos dados pessoais não são divulgadas de forma adequada pelos controladores (MORTIER et al., 2016). Essas ações provocam questionamentos sobre como o uso dos dados pode interferir na privacidade, segurança e liberdade dos indivíduos (DROZD, 2016).

Entretanto, proporcionar Transparência aos titulares dos dados com informações claras, acessíveis e compreensíveis deve passar pela resolução de desafios na área de IHD tais como:

- **Autorização de uso:** os aplicativos e sites solicitam autorização de uso com uma abordagem *black box* na qual o indivíduo deve autorizar acesso a todos os seus dados ou não poderá utilizar a aplicação. A solicitação de autorização é realizada com poucas informações e com conteúdo superficial e/ou técnico, o que induz o indivíduo a aceitar as condições (WATSON, 2016);
- **Complexidade de conteúdos e juridicização da informação:** é o caso das Políticas de Privacidade e Segurança (PPS) que apresentam textos complexos, volumosos, jurídicos e superficiais, o que dificulta a análise por parte de indivíduos não especialistas. Esse cenário advém da necessidade dos controladores garantirem sua segurança jurídica em relação ao funcionamento da aplicação. Tais características do conteúdo levam mais 90% das pessoas a não fazerem uma leitura adequada das PPS (MILNE; CULNAN, 2004; COLETI et al., 2018);
- **Segredos comerciais:** empresas que utilizam os dados pessoais questionam o fato de que, ao divulgar informações sobre o uso de seus dados, poderia divulgar seus segredos comerciais e assim obter prejuízos financeiros consideráveis (BELLAMY; ALONSO, 2016; CHRISTL, 2017);
- **Falta de conhecimento dos indivíduos:** os indivíduos apresentam conhecimento

pobre sobre o uso de seus dados pessoais e pouco sabem sobre o volume de dados que é coletado. Uma pesquisa feita por Filgueiras et al. (2019) utilizou de um aplicativo móvel para registrar e alertar o indivíduo sobre a coleta de seus dados. O volume de dados provocou surpresas nos participantes devido à frequência e ao volume de dados identificados pelo aplicativo;

- **Dificuldade de apresentação visual de conteúdo:** transformar um conjunto de ações conduzidas nos dados pessoais, que comumente passam por ações computacionais (algoritmos, armazenamento em banco de dados etc), em formatos de apresentação comprehensíveis aos indivíduos não é uma tarefa trivial e pode envolver a combinação de diversos elementos de Interação-Humano Computador (IHC) (FAZLIOGLU, 2017; HADDADI et al., 2013);
- **Questões éticas:** há questões éticas envolvidas, a exemplo de: qual informação deve ser mostrada? qual o impacto para os negócios? o indivíduo precisa da informação (TURILLI; FLORIDI, 2009);

De forma geral, o conhecimento sobre uso dos dados pessoais ainda é pequeno por parte da população em geral, uma vez que a interferência não ocorre de forma muito explícita ou é apresentada por meio de algum recurso benéfico oferecido por aplicações (SCHNEIER, 2015). Entretanto, é considerado certo que o aumento do uso dos dados pessoais e suas aplicações devem interferir de forma mais profunda na vida das pessoas podendo prever ações, expor detalhes de comportamentos e preferências das pessoais dentre outras inúmeras possíveis ações (CHRISTL, 2017; SCHNEIER, 2015).

Diante do exposto, iniciativas foram conduzidas em empresas e laboratórios de pesquisas tais como:

O surgimento de leis e regulamentações de uso de dados pessoais como a GDPR, LGPD e PIPEDA contendo um conjunto de diretrizes sobre o uso dos dados pessoais. (Informações sobre as regulamentações serão discutidas na Seção 2.3.1.)

Em relação às **aplicações computacionais**, as *Transparency Enhancing Tools* (TETs), as mesmas são ferramentas que buscam apresentar informações sobre o uso dos dados pessoais (MURMANN; FISCHER-HÜBNER, 2017a; BARRETO; SALGADO; VITERBO, 2018). Hedbom (2009), Janic, Wijbenga e Veugen (2013), Murmann e Fischer-Hübner (2017a) descrevem as TETs como ferramentas que apresentam *insights* sobre os dados pessoais coletados, tais como informações de armazenamento, compartilhamento e processamento. Já Barreto, Salgado e Viterbo (2018) consideram as TETs como ferramentas

de apoio aos indivíduos para questões relacionadas à privacidade e proteção.

As TETs são iniciativas eficazes para auxiliar o indivíduo. Entretanto, as mesmas ainda apresentam um conjunto de informações muito diversificado, o que não garante aos indivíduos um padrão de informações que possa ser considerado relevante. Murmann e Fischer-Hübner (2017a) e Barreto, Salgado e Viterbo (2018) conduziram revisões e avaliações de TETs as quais apresentam informações que compreendiam:

- Localização dos destinatários dos dados, para fins de autorização;
- Dados que seriam coletados, por meio da listagem da *cache* do navegador;
- Lista de serviços com suas ações sobre os dados a fim de permitir o indivíduo escolher o que mais lhe atende;
- Informações de uso da aplicação em forma de *Log files*;
- Endereços e *status* de terceiros identificados e que tenham acesso aos dados;
- Capacidade dos indivíduos intervirem no uso de seus dados indicando quais dados deseja monitorar, dentre outros.

As **Políticas de Privacidade e Segurança (PPS)** também são utilizadas para transmitir informações sobre o uso dos dados pessoais. Entretanto, o formato de exibição e o conteúdo apresentado costuma ser volumoso, complexo, denso, técnico e de difícil interpretação pelo indivíduo o que dificulta a leitura e análise pelos indivíduos (COLETI et al., 2018; EARP et al., 2005). A dificuldade de leitura ocorre muito em virtude da pouca Legibilidade, o que, segundo Cybis, Holts e Faust (2015) afeta negativamente a visualização da informação para os usuários.

Uma iniciativa de interesse, conduzida por empresas, é apresentada por Bellamy e Alonso (2016). Os autores destacam uma ***Roundtable*** entre a *Telefónica*, uma das maiores empresas de telecomunicação do mundo e a CIPL, um centro de pesquisa em privacidade e segurança de dados. Essa parceria discutiu as recentes políticas para uso de dados pessoais, iniciativas e leis com foco em Transparência de Dados Pessoais. Na *roundtable* conclui-se que as principais demandas por Transparência são: (1) controle do usuários; e (2) novas abordagens de Transparência. Além dessas demandas, foram discutidos outros fatores relacionados à Transparência a fim de melhorar aspectos relacionados a essa nova necessidade de software.

Por fim, a Transparéncia de Dados Pessoais é um mecanismo importante para permitir que o indivíduo seja inserido no fluxo do uso de dados pessoais e possa agir, caso necessário, em relação ao uso deles. Na ocasião da condução dessa pesquisa, muitos aspectos inerentes à Transparéncia e ao uso dos dados pessoais eram incipientes e muitos desafios e oportunidades de melhorias nos aspectos de Transparéncia a fim de beneficiar o indivíduo ainda poderiam ser desenvolvidos.

Um mecanismo de Transparéncia é fundamental a fim de garantir ao indivíduo o conhecimento do uso de seus dados, o monitoramento do *fair use* por parte dos controladores e os insumos necessários para intervir caso necessário.

A próxima seção destaca sobre as regulamentações de uso dos dados pessoais com destaque para a GDPR.

2.3 Regulamentação para uso de dados pessoais

Esta seção apresenta duas regulamentações de uso de dados pessoais, a *General Data Protection Regulation* (GDPR) e a Lei Geral de Proteção de Dados (LGPD).

2.3.1 *General Data Protection Regulation (GDPR)*

A necessidade de maior fiscalização e regramento do uso dos dados pessoais fez surgir regulamentações criadas por países, grupos de países e/ou organizações. A mais conhecida, e com uma área de abrangência maior é a GDPR, criada pela União Europeia e que entrou em vigor em Abril de 2018.

O texto desta seção discutirá a respeito da GDPR, uma regulamentação de uso de dados pessoais na qual em seu conteúdo não cabe interpretações pessoais ou subjetivas. Assim, todo o conteúdo relacionado à GDPR teve como única fonte bibliográfica, o conteúdo da própria GDPR evitando assim o uso de artigos que pudessem deturpar seu conteúdo. O conteúdo foi extraído e analisado a partir da página: <https://gdpr-info.eu/>.

A GDPR é uma regulamentação que dispensa muito de seu conteúdo em defesa do indivíduo e preza, principalmente, por garantir os direitos dos mesmos em relação ao uso de seus dados. A regulamentação não prejudica os controladores (empresas que utilizam os dados pessoais), mas cria normas rígidas de uso dos dados pessoais a fim de garantir os direitos dos titulares dos dados (VOIGT; BUSSCHE, 2017).

No que tange à Transparéncia de Dados Pessoais, a GDPR pode ser considerada

completa e objetiva em listar um conjunto de informações que devem ser disponibilizadas aos indivíduos com explicações justificando algumas das informações. Os artigos 13, 14 e 15 da GDPR são voltados fortemente para conteúdos de Transparência. Embora a GDPR não especifique os critérios de apresentação da Transparência, a lista de requisitos de Transparência é considerável para auxiliar na implementação de Transparência em aplicações de *software*.

As diretrizes para Transparência da GDPR são mostradas na Tabela 26 - Anexo A as quais são organizadas em Artigo, Seção e Alínea.¹

A GDPR abrange todas as empresas do território europeu e qualquer empresa fora desse território que tenha algum vínculo com a União Européia, por exemplo, a filial de uma empresa. A obrigatoriedade da GDPR trouxe necessidades de adequação em diversas organizações e a falta do cumprimento da mesma pode ocasionar em multas de valores altos (MURPHY, 2018; ROBOL; SALNITRI; GIORGINI, 2017).

As diretrizes de Transparência dessa regulamentação podem ser utilizadas para apoiar a criação e avaliação de recurso e modelos de Transparência assim como foi utilizada de base para o desenvolvimento desta pesquisa.

Na próxima seção será discutida a Lei Geral de Proteção de Dados.

2.3.2 Lei Geral de Proteção de Dados

A Lei Geral de Proteção de Dados (LGPD) foi sancionada no mês de Agosto do ano de 2018 sob o número 13.709/2018 e regula as atividades de manipulação de dados pessoais realizadas em território brasileiro, com dados coletados dentro do território brasileiro ou que o titular dos dados esteja localizado em território brasileiro (TOLEDO, 2020). A regulamentação atinge ações como coleta, processamento, arquivamento, eliminação, avaliação, acesso, divulgação, transferência ou qualquer ação realizada com os dados pessoais, que na lei é referenciada pelo termo manipulação de dados (MONTI, 2014; TOLEDO, 2020).

Toledo (2020) destaca que a LGPD a dupla função de: (1) fomentar o desenvolvimento tecnológico e econômico; e (2) proteger direitos e liberdades fundamentais. Pioneiro Neto (2020) e Law (2020) destacam que a LGPD contém um conjunto de princípios e melhores práticas, dentre elas, o princípio da Transparência.

¹A Tabela é apresentada no Apêndice por ser longa e necessitar de adequações visuais. Diretrizes diferentes com o mesmo conteúdo são apresentadas na mesma linha da Tabela.

O princípio de Transparência orienta que os titulares dos dados tenham acesso fácil e claro às informações sobre o uso de seus dados pessoais. Pioneiro Neto (2020) destaca a necessidade de informações como: (1) o objetivo do tratamento dos dados pessoais; (2) a adequação do uso dos dados para o contexto de manipulação existente; e (3) a real necessidade do tratamento dos dados para a finalidade apresentada.

Law (2020) destaca a Transparência da LGPD com os seguintes itens:

- Finalidade específica do tratamento dos dados;
- Forma e duração do tratamento dos dados;
- Dados de contato do controlador;
- Informações sobre compartilhamento dos dados como finalidade e contato do destinatário;
- Responsabilidades dos envolvidos no tratamento dos dados;
- Direitos dos titulares dos dados;
- Possibilidade de recusar a autorização do uso dos dados e a consequência dessa escolha.

Destaca-se também a necessidade de acessibilidade e facilidade de compreensão das informações de forma que possam ser utilizadas pelos indivíduos para analisar o tratamento de seus dados.

A próxima seção discorre sobre Perfil de Aplicação de Metadados.

2.4 Perfil de Aplicação de Metadados

Metadados são informações que descrevem os dados, de forma simples, “são os dados dos dados”(CANADA, 2014). Essencialmente, descrevem características dos dados ou das informações, tais como, proprietários, formatos, tamanho, histórico dentre outros, ajudando na organização, na publicação e no suporte (MORAN, 2008).

Os metadados articulam o contexto de um recurso, como uma imagem, um arquivo de texto ou qualquer registro digital, fornecendo descrições que permitem analisar, entender e trocar dados entre diferentes operadores (COYLE; BAKER, 2009). O conceito de metadados não é novo, mas as aplicações modernas ficaram conhecidas por “orientados

à computação” e ganharam espaço nos meados da década de 1990, em conjunto com o crescimento da Internet (COYLE; BAKER, 2009; MORI et al., 2004). A Internet coloca um grande número de documentos à disposição, o que requer uma maneira estruturada de fornecer conteúdo e contexto significativos para utilização por usuários (MORI et al., 2004).

A necessidade de criar metadados baseados na Web que se adaptassem ao domínio surgiu no início dos anos 2000 e ficou conhecido como *Metadata Application Profile* (TENNIS, 2015). *Metadata Application Profile* (MAP) é um conjunto de diretrizes que especifica quais metadados são usados em um domínio, bem como suas regras de uso, formato de dados e outras especificações que permitem a utilização no domínio (COYLE; BAKER, 2009).

A principal justificativa para o uso de MAPs é a capacidade de usar uma combinação de diferentes padrões de metadados, extraíndo o melhor de cada um a fim de atender ao domínio específico (MORGANA; BAPTISTA, 2015). Ainda assim, um MAP não precisa necessariamente ser uma combinação de vários padrões, mas pode ter sua estrutura projetada especificamente para um domínio, caso outros padrões não possam ser reutilizados (TENNIS, 2015).

Dublin Core Application Profile (COYLE; BAKER, 2009)² e Tennis (2015) destacam que um MAP pode conter os seguintes itens:

- **Requisito Funcional:** descreve as ações ou conjunto de ações as quais o MAP deve permitir cumprir;
- **Modelo de Domínio:** especifica as entidades, agentes, relações e atributos do domínio;
- **Conjunto de Descrições:** descreve os termos dos metadados a serem utilizados e suas regras de utilização;
- **Guia de sintaxe o Formato de dados:** definem regras de codificação que serão utilizadas por sistemas computacionais para ler/processar os dados.

Morgana e Baptista (2015) e Malta e Baptista (2014) destacam vários exemplos de MAPs que procuraram abordar uma ampla gama de domínios. Os exemplos citados pelos autores são: *Scholarly Work Application Profile* (SWAP) (UKLON, 2009) usado

²<http://www.dublincore.org/specifications/dublin-core/profile-guidelines/> - Primeiro acesso em Janeiro de 2016

para descrever pesquisas e trabalhos escolares e de pesquisas; e *The Virtual Open Access Agriculture - Aquaculture Repository* (VOA3R) (CORDIS, 2017) desenvolvido pela Organização para a Agricultura e Alimentação (FAO).

Um MAP amplamente conhecido é o Dublin Core, que contém quinze elementos de metadados para identificar e classificar documentos em um contexto específico (APPS; MACINTYRE, 2000). O uso do Dublin Core permite que seus termos sejam refinados por meio do *Qualified Dublin Core* a fim de melhorar a interoperabilidade e a precisão semântica do conteúdo (APPS; MACINTYRE, 2000) (COYLE; BAKER, 2009). Interoperabilidade é a capacidade de trocar informações entre sistemas heterogêneos e distribuídos em localizações físicas distintas (DIALLO et al., 2011); e *Semantic Web* é a capacidade de classificar os dados com base em diferentes contextos e atribuir significado a eles, a fim de melhor entendimento humano, mas também para aprimorar o entendimento das máquinas (ISMAIL; SHAIKH, 2016).

Outro MAP conhecido, e que busca atender à demanda de compartilhamento de dados de maneira rápida, confiável e de qualidade, é o *Darwin Core*. O *Darwin Core* pode ser usado como padrão para o compartilhamento de dados entre sistemas sobre biodiversidade (MCKILLIP; JAYKUS; DRAKE, 1998). Os metadados do *Darwin Core* permitem tratar dados de ocorrências de biodiversidade com informações sobre localização, bioma, contexto, taxonomia e identificação de espécies³.

A quantidade de MAPs existentes ainda é relativamente pequena, conforme discutido por Morgana e Baptista (2015). No entanto, é possível identificar várias iniciativas para o uso de perfis de aplicação em domínios como bibliotecas, coleções de dados, mecanismos de codificação para transmissão, dados abertos e conteúdo multimídia (MORGANA; BAPTISTA, 2015), (SAMPSON; ZERVAS; CHLOROS, 2012).

Acredita-se que o desenvolvimento de novos MAPs pode aumentar as possibilidades do uso correto de metadados e suas descrições em diversos domínios de aplicação (SAMPSON; ZERVAS; CHLOROS, 2012). No contexto desta pesquisa, o uso de dados pessoais é considerado um domínio com ampla possibilidade de uso de MAPs.

2.5 Visualização de dados por indivíduos

A Visualização de dados por indivíduos (*Personal Infovis*) é a capacidade de transmitir uma informação abstrata utilizando meios visuais como textos, gráficos, tabelas, mapas e

³<https://github.com/tdwg/dwc>

imagens para facilitar que um indivíduo compreenda a mensagem transmitida (HUANG et al., 2015). Quando se pensa em informação abstrata, pensa-se em descrever algo que não é físico, por exemplo: análises qualitativas, dados estatísticos, volume de vendas e distribuição de ocorrências de doenças sazonais. Os exemplos citados, embora sejam reais e passíveis de análise, necessitam ser traduzidos em cores, desenhos, formatos, letras e rótulos para seu usuário (FEW, 2016; ILLINSKY; STEELE, 2011).

A necessidade de visualização de informação por indivíduos vem junto com a necessidade dos mesmos em interagir com bases de dados produzidas por eles/elas mesmos, por outras pessoas, por empresas, laboratórios ou outros tipos de entidades (GOMES; GAMA; GONÇALVES, 2010; HSIEH, 2016). O consumo cada vez maior de informações por pessoas tornou o processo de visualização de informações fundamental para proporcionar uma melhor experiência de interação e compreensão entre os usuários e os dados e, assim, a visualização tornou-se um dos principais desafios e/ou área de estudos da IHD (GOMES; GAMA; GONÇALVES, 2010; HADDADI et al., 2013).

Historicamente, os dados comumente foram (e são) armazenados em tabelas, arquivos com formatos diversos ou outros meios para facilitar o processamento por meio de computadores. Entretanto, a compreensão e utilização da informação por pessoas é facilitada, ou até mesmo viabilizada, pelo uso adequado de representações visuais, considerando a habilidade em transmitir uma informação de maneira que os indivíduos possam entender sua mensagem (FEW, 2016).

Um dos grandes desafios da *Personal Infovis* assenta-se no fato de que pessoas diferentes podem ter perspectivas e objetivos diferentes e entender uma mesma informação de forma totalmente distinta. Segundo Gomes, Gama e Gonçalves (2010), as técnicas de visualização não evoluíram de forma a atender as diversas necessidades dos indivíduos, o que produz nos mesmos a necessidade de filtrar ou explorar de forma mais aprofundada a informação produzida a fim de compreender melhor o conteúdo apresentado.

Para melhorar a experiência do usuário na utilização de recursos visuais, novas abordagens de visualização foram propostas e aplicadas, tais como: *Personal Visualization (PV)*; *Personal Visualization Analytics (PVA)*; *Embodied Interaction*; e *Collaborative/Embodied Interaction Visualization Tool* (HUANG et al., 2015).

Personal Visualization (PV) é a capacidade de proporcionar recursos interativos para os indivíduos explorarem os dados em um contexto pessoal de uso sem realizar, necessariamente, análises complexas ou apoiadas por computador (HUANG et al., 2015). Segundo Cafaro (2012) um indivíduo pode ser mais propenso a explorar os dados ao invés de re-

alizar análises complexas e, desta forma, um recurso de visualização deve permitir que o usuário conheça, visualize e explore informações a fim de compreender o conteúdo do *dataset*.

Um exemplo da PV é o aplicativo **Sem Dengue**⁴. no qual a interface principal é mostrada na Figura 3. A aplicação do exemplo não fornece análises complexas para o usuário, mas exibe gráficos e tabelas com valores de ocorrências (totalizadores) em textos formatados para a localização atual do indivíduo e permite a inclusão de denúncias sobre novos focos de dengue.



Figura 3: Exemplo de *Personal Visualization* no aplicativo Sem Dengue

A *Personal Visualization Analytics* (*PVA*) acrescenta na visualização de informações as análises complexas apoiadas por computadores (HUANG et al., 2015). O objetivo da PVA não é somente explorar um *dataset*, mas analisá-lo em busca de melhores soluções para do indivíduo. Para realizar as análises, recursos interativos permitem que indivíduo altere parâmetros de consulta e ajuste para as necessidades de informação desejada. Para Cafaro (2012), a análise exige do usuário conhecimento dos elementos que influenciam a informação final a fim de poder manipular o sistema de forma adequada e produzir resultados coerentes.

Um exemplo da PVA é o aplicativo Waze, que utiliza dados da localização do usuário, dados produzidos por demais usuários, dados de trânsito e rotas para calcular a melhor rota para um destino (GOOGLE, 2017). Nesta abordagem de uso, o indivíduo tem como objetivo buscar uma rota (ou uma rota melhor) para chegar em um destino. O Waze propõe uma rota inicial, considerada a melhor rota. Caso o usuário queira uma nova

⁴<http://www.brasilsem Dengue.org/home.htm> - Acessado em Janeiro de 2018

rota o mesmo pode mudar os parâmetros e assim o Waze recalcula e propõe alternativas considerando as entradas do indivíduo.

Além das abordagens de *Personal Infovis* mostradas, que pertencem a uma categoria mais tradicional de visualização e interação de informação, há também abordagens voltadas para o uso de imagens 3D, interação embarcada, interação colaborativa e *Natural User Interface* (CAFARO, 2012; ELMQVIST, 2011; SALIH, 2010) . Essas abordagens são conhecidas como *Embodiments for Interaction/Data Visualization* e *Collaborative/Embodied Interaction Visualization Tool*.

Essas abordagens são mais complexas e utilizadas para interação com informações fortemente dinâmicas nas quais o comportamento ou necessidade do indivíduo interferem no resultado final (ELMQVIST, 2011). Por não ser aplicada nesta pesquisa, não serão apresentados mais detalhes nessa seção.

As abordagens de visualização de dados apresentadas têm por foco o indivíduo como o usuário da informação. Desta forma é possível ver com os exemplos apresentados que diversas formas de visualização são, e foram, desenvolvidas para facilitar a interação das pessoas com grandes volumes de dados e assim poder usufruir de forma adequada desses elementos que norteiam o dia a dia (HADDADI et al., 2013; MORTIER et al., 2016; HUANG et al., 2015).

A *Personal Infovis* incentiva uma contínua busca por melhorias na capacidade de interação com dados, principalmente pelo fato de que indivíduos não especialistas em um assunto e, com diferentes perspectivas, necessitam explorar, ou mesmo analisar, grandes volumes de dados em busca de metas e objetivos diferentes (HUANG et al., 2015).

Esse conceito apresenta forte relação com a Transparência de Dados Pessoais, uma vez que, segundo Haddadi et al. (2013), as informações sobre o uso dos dados pessoais, que apresentam fortes elementos técnicos, devem ser mostradas por meios visuais a fim de facilitar a compreensão e uso dos indivíduos.

A próxima seção discute sobre Qualidade de Informação.

2.6 Qualidade de Informação

A Qualidade da Informação, ou *Information Quality* (IQ) é um conceito multidimensional que visa dar suporte à produção de informações adequadas para as necessidades dos usuários (WANG; STRONG, 1996). A preocupação com IQ cresceu à medida em

que a busca e o uso de informações da Internet pelas pessoas cresceram. Desta forma, a decisão das pessoas fica cada vez mais atrelada à capacidade de interpretar, analisar e também de confiar na informação acessada (ABIB, 2010).

Uma informação de qualidade é aquela que se enquadra no conceito de *fit for user* (KANDARI et al., 2011a). Esse conceito é amplamente adotado na IQ uma vez que sua literatura destaca a importância de considerar o ponto de vista do consumidor da informação, pois será ele/ela que fará o julgamento da informação consumida (KANDARI et al., 2011a; WANG; STRONG, 1996) .

A IQ busca dar suporte à produção de informação com qualidade com base nos conceitos das dimensões de Qualidade Contextual (Relevância, Completitude e Valor agregado) e Qualidade Representacional (facilidade em compreender, capacidade de interpretar e representação concisa) (WANG; STRONG, 1996).

Determinar a IQ de uma informação depende da análise de fatores como qualidade do produto que entrega a informação, contexto de uso e fatores pessoais (ABIB, 2010). Assim, a IQ depende de um conjunto de dimensões que possibilitem uma melhor mensuração da qualidade e deve considerar as percepções subjetivas dos indivíduos bem como o contexto em que as informações e indivíduos estão inseridos (ABIB, 2010; WANG; STRONG, 1996).

As dimensões de IQ foram estabelecidas por Wang e Strong (1996). Entretanto, diversos trabalhos conduziram estudos para verificar a aplicação em contextos específicos, para refinar/melhorar as dimensões ou para desenvolverem técnicas de aplicação e avaliação de IQ com base nas dimensões. São exemplos desses trabalhos (KANDARI et al., 2011a; LEE et al., 2002; STVILIA et al., 2007).

A Tabela 1 disponível na página 40 apresenta as dimensões propostas por Pipino, Lee e Wang (2002).

Das dimensões apresentadas no trabalhos supracitados, cinco delas foram selecionadas como necessárias para Transparência de Dados Pessoais e são apresentadas na Tabela 7 na página 68 e sua aplicação na pesquisa será discutida posteriormente.

Para Lee et al. (2002) as dimensões de IQ têm sua importância aumentada de forma significativa como mecanismo para auxiliar na organização e na busca dos objetivos de qualidade de informação o que faz com que as mesmas estejam frequentemente no topo das listas de preocupações em projetos envolvendo o uso de informações.

No que tange à avaliação da qualidade de informação com base nas dimensões, Pi-

Tabela 1: Dimensões de *Information Quality*. Adaptado de Abib (2010) e Pipino, Lee e Wang (2002)

Dimensão	Descrição
Acessível	A extensão em que a informação está disponível com acesso rápido e fácil
Quantidade de Informação	A quantidade de informação apresentada deve ser adequado ao indivíduo. No contexto de Transparência, deve limitar-se às informações de uso dos dados pessoais
Credibilidade	Nível de crédito e confiança na informação
Completitude	A informação deve estar completa para dar suporte ao usuário sem a necessidade do mesmo procurar informações complementares em outros fontes
Capacidade de compreensão	A informação deve ser apresentada de forma que o usuário possa compreender o conteúdo transmitido
Representação concisa	A informação deve ser apresentada de forma compacta
Representação consistente	A informação deve ter um determinado padrão
Livre de erros	A informação deve estar correta e confiável
Objetividade	A informação deve ser livre de viés, de parcialidade e de conteúdos prejudiciais
Relevância	A informação deve ser relevante para o indivíduo e deve contribuir para o indivíduo analisar e compreender como seus dados pessoais são utilizados
Reputação	A extensão do quanto a informação é embasada por fontes confiáveis e conteúdo apropriado
Segurança	A extensão do quanto a informação é acessada somente por quem tem autorização
Atualizada	Nível de atualização da informação
Interpretável	A extensão do quanto a informação é facilmente compreendida
Valor agregado	A extensão do quanto a informação é benéfica e utilizable por seus usuários

pino, Lee e Wang (2002) destacam que dois métodos são comumente utilizados: Método Subjetivo e Método Objetivo.

No Método Subjetivo, o usuário avalia a informação com base em sua percepção em relação à sua qualidade. Segundo Pipino, Lee e Wang (2002) essa avaliação pode ser mais complexa de ser analisada, pois envolve questões como: necessidade de uso da informação, costumes do avaliador e forma de utilização da informação. O resultado da avaliação do indivíduo será influenciada por esses parâmetros e o avaliador dever considerá-los na hora de emitir um parecer final.

No Método Objetivo, os avaliadores determinam um conjunto de métricas e produzem meios objetivos e bem definidos do participante avaliar a informação. São consideradas

técnicas como: questões objetivas, mínimo e máximo, escala dentre outras. A grande dificuldade dessa técnica é a escolha das dimensões utilizadas na avaliação, assim como o estabelecimento de quais técnicas são aplicadas (PIPINO; LEE; WANG, 2002).

Uma proposta que utiliza o Método Objetivo é descrita por Lee et al. (2002) onde apresenta uma abordagem em que questões sobre a qualidade de informação são organizadas em quadrantes. Os dados das avaliações permitem fazer um *benchmark* entre a qualidade de informação esperada na organização e qualidade da informação avaliada. A interpretação dos resultados é feita com o auxílio de um componente específico da técnica, chamada de AIMQ.

Já Stvilia et al. (2007) propôs um *framework* com base na compreensão das diversas tipologias de problemas relacionados à qualidade de informação, relacionadas às tarefas e às dimensões. Essas informações foram organizadas em uma estrutura sistemática que considera teorias e práticas. A melhoria proposta pelos autores reside no fato de não considerar somente problemas locais e poucas variáveis, mas em uma estrutura mais completa e organizada. O modelo foi considerado pelos autores como um recurso para apoiar a construção de outros modelos de avaliação de IQ.

Assim, considera-se a IQ como elemento fundamental para a construção de aplicações que trabalham com informações. A observação da IQ, suas dimensões e técnicas de avaliação podem ser o diferencial para a qualidade final da informação.

O próximo capítulo apresenta a Revisão Sistemática.

3 REVISÃO SISTEMÁTICA

Este capítulo apresenta a Revisão Sistemática (RS) conduzida com o objetivo de: *Identificar trabalhos que especifiquem, de forma direta ou indireta, quais informações sobre o uso dos dados pessoais compõem a Transparência de Dados Pessoais em aplicações de software, e como as informações podem ser apresentadas ao indivíduo.*

Entende-se por forma direta, artigos que propuseram um conjunto de metainformações¹ de Transparência; por forma indireta, deve-se compreender artigos que não propuseram metainformações de Transparência, mas as elencaram.

3.1 Preparação da Revisão

A RS foi conduzida com base nas estratégias propostas por Biolchini, Mian e Natali (2005) e Kitchenham (2004) e de acordo com o protocolo disponível no Apêndice A.

Das informações apresentadas no protocolo, destacam-se as questões de pesquisa: (1) quais informações sobre o uso dos dados pessoais são utilizados para compor a Transparência de Dados Pessoais? e (2) quais técnicas são utilizadas para apresentar as informações de Transparência para os indivíduos?

3.2 Condução da Revisão

A seleção de trabalhos nas bases de dados foi realizada no período de Janeiro a Abril de 2019, e as atividades conduzidas foram:

- 1. Obtenção dos artigos:** foram conduzidas as buscas por trabalhos nas bases de dados indexadas com as *strings* de busca apresentadas no protocolo da RS. O total de artigos encontrados nessa etapa foi de 1958.

¹Termo utilizado para referenciar os dados/informações que são disponibilizados sobre o uso dos dados pessoais

2. **Seleção de estudos primários:** foi realizada a leitura do título e do *abstract* com o objetivo de remover trabalhos que não se enquadram nos objetivos da RS. A quantidade de artigos excluídos nessa etapa foi consideravelmente grande, uma vez que somente 61 (sessenta e um) artigos foram selecionados para a extração de conteúdo. Assumiu-se que o grande número de falsos positivos ocorreu porque a *string* de busca foi utilizada no modo *Full Text and Metadata* em todas as bases. Esse meio busca em todas as informações disponibilizadas pelas bases de dados para os artigos. Assim, ocorreu o retorno de trabalhos em que as palavras chave não eram utilizadas dentro do contexto desejado para RS.
3. **Extração de Conteúdo:** Os artigos incluídos foram analisados por completos, classificados nos critérios de Inclusão e Exclusão e então realizada a Extração dos conteúdos. Dos 61 artigos para análise de conteúdo, 11 atenderam aos critérios de inclusão e tiveram seus conteúdos selecionados. Os títulos abreviados e as referências dos trabalhos selecionados são mostrados na Tabela 2 na página 44. Nessa etapa também foram selecionados artigos para inclusão manual via *snowball*, os quais são mostrados na Tabela 3 na página 44.
4. **Tabela de resultados:** os resultados foram tabelados tanto para metainformações de Transparência quanto para formas de apresentação. As tabelas com as informações extraídas são apresentadas e discutidas na continuidade deste texto. A escolha pela organização em forma de tabela deu-se pelo fato de assumir que o resultado pretendido na RS seria obtido ao efetuar uma análise comparativa das informações identificadas em cada artigo, facilitando assim, a identificação e síntese dos resultados.
5. **Análise dos resultados:** com base na análise dos dados extraídos, foi realizada uma discussão dos resultados em relação às metainformações de Transparência e à forma de apresentação.

Tabela 2: Artigos selecionados na RS para análise e síntese dos resultados. Do autor.

Nº	Título abreviado	Autores
1	A Scalable Consent, Transparency ...	(KIRRANE et al., 2018)
2	Can Transparency Enhancing Tools Support ...	(FERREIRA; LENZINI, 2015)
3	Conceptual Representation ...	(TOM; SING; MATULEVIČIUS, 2018)
4	Foundations for Transparency ...	(HOSSEINI et al., 2016)
5	How can cloud users ...	(FISCHER-HÜBNER; ANGULO; PULLS, 2014)
6	Modeling Metrics for ...	(SPAGNUOLO; BARTOLINI; LENZINI, 2017)
7	Transparency, Privacy and Trust ...	(FISCHER-HÜBNER et al., 2016)
8	Nobody puts data in a ...	(CRADOCK; STALLA-BOURDILLON; MILLARD, 2017)
9	PrivacyInsight: The Next ...	(BIER; KUHNE; BEYERER, 2016)
10	Security Analysis and Legal ...	(GUARDA; RANISE; SISWANTORO, 2017)
11	Tools for achieving usable ...	(MURMANN; FISCHER-HÜBNER, 2017a)

Tabela 3: Artigos incluídos manualmente na RS. Do autor.

Nº	Título abreviado	Autores
12	From privacy legislation to ...	(PATRICK; KENNY, 2003)
13	Usable Transparency Enhancing ...	(MURMANN; FISCHER-HÜBNER, 2017b)
14	Transparency Enhancing tools ...	(JANIC; WIJBENGA; VEU-GEN, 2013)
15	Towards displaying privacy ...	(HOLTZ; NOCUN; HANSEN, 2011)

A Tabela 25 disponível no Apêndice B contém a extração das metainformações de Transparência e da forma de apresentação proposta por artigos.

3.3 Análise dos conteúdos

A análise dos resultados foi organizada em duas partes: (1) a primeira buscou responder a questão relacionada às metainformações de Transparência de Dados Pessoais. Nessa etapa foi realizada a identificação e discussão das metainformações apresentadas nos trabalhos; (2) a segunda parte deu suporte para a análise das estratégias de apresentação das informações para os indivíduos, as quais também foram identificadas e discutidas.

Para responder a primeira parte da pergunta da RS, foram identificadas as metainformações de Transparência organizadas em tópicos de acordo com sua similaridade². Para cada tópico foi indicada a ocorrência nos artigos selecionados. O resultado dessa atividade é mostrado na Tabela 4, na página 46.

²Foi considerada similaridade quando não houve subjetividade em relação à definição de duas ou mais palavras que sinalizavam a mesma informação.

Tabela 4: Ocorrência de dados para Transparéncia nos artigos analisados. Do autor.

Informação	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Informações sobre os dados pessoais utilizados	X			X	X		X	X	X	X	X	X	X	X	X
Informações sobre o armazenamento dos dados	X					X			X						X
Informações sobre o processos conduzidos nos dados pessoais	X		X	X	X				X	X		X		X	
Informações sobre o propósito de uso	X	X	X	X					X	X		X			X
Informações sobre a origem dos dados									X		X		X		
Informações sobre compartilhamento, divulgação e/ou distribuição dos dados			X		X		X	X			X		X	X	
Informações sobre pessoas e empresas envolvidas no uso dos dados pessoais como controladores, operadores e destinatários	X	X	X	X	X	X			X		X	X	X	X	X
Leis e direitos dos indivíduos					X		X			X				X	X
Informações produzidas com o uso dos dados pessoais					X					X	X				
Dados sobre objetos e detalhes que compõem os dados pessoais tais como número de cartão, e-mail, grupos, rotinas, endereço IP, dados de rede, dados de cache do navegador etc.									X		X	X		X	
Informação se o software está cumprindo o que foi acordado com o usuário em relação ao uso dos dados pessoais			X			X	X			X				X	X
Informações sobre como os dados são anonimizados		X			X				X			X			X

As análise das informações apresentadas na Tabela 4 permitiu concluir que:

(1) Alguns trabalhos analisados não tinham como foco a proposição de metainformações de Transparência. Entretanto, todos os trabalhos apresentaram, de forma direta ou indireta, metainformações tomadas por base para definir/compor o conceito de Transparência de Dados Pessoais utilizada em cada trabalho.

Os trabalhos apresentaram um conjunto relativamente pequeno de metainformações, uma vez que a classificação resultou em 12 (doze) itens. Destaca-se uma aplicação maior para os itens das 7 (sete) primeiras linhas da Tabela 4, com maior destaque para os dados sobre quais informações dos indivíduos são coletadas e quem acessa e utiliza os mesmos.

O uso dessas metainformações pode ser justificado pela existência de leis/regulamentações de uso dos dados como a GDPR ,que entrou em vigor em 2018, pois são diretrizes requeridas pela regulamentação; além do fato de os artigos analisados datarem a partir desse período.

Já outras metainformações apresentaram menor ocorrências tais como:

- **Informações produzidas com o uso dos dados pessoais:** Essa metainformação pode envolver a divulgação de segredos comerciais o que leva a uma maior resistência (e cuidados) das empresas em divulgá-los;
- **Dados sobre objetos e detalhes que compõem os dados pessoais:** Exemplos - número de cartão, e-mail, grupos, rotinas, endereço IP, dados de rede, Dados de cache do navegador. Essa metainformação pode ter relação com a forma de utilização do dado, pois esse item tem uma grande similaridade com o item **Informações sobre os dados pessoais utilizados** e o artigo pode ter optado por utilizar uma definição mais detalhada. Entretanto, como não foi identificada uma classificação para esse dados, na extração da RS, optou-se por ter dois itens de dados distintos.

(2) **Falta de padronização das metainformações:** Dos 15 (quinze) artigos analisados, não há ocorrências em que conjuntos completos e idênticos de metainformações de Transparência sejam aplicados em dois ou mais artigos distintos. Ocorrem combinações parciais como é o caso dos artigos 11 e 14, provavelmente por serem do mesmo grupo de pesquisa; e os artigos 1 e 9 que utilizam de diretrizes europeias para embasar suas pesquisas.

A falta de padrão na relação de metainformações pode produzir Transparência enviesada fazendo com que o controlador disponibilize somente as metainformações que achar

conveniente para as metas da empresa controladora do software e assim, onerando a capacidade do indivíduo de analisar a Transparência e exercer seus direitos, caso necessário.

(3) **Possibilidade de subjetividade na interpretação da informação:** foi identificada uma subjetividade na definição das metainformações de Transparência e que pode ser ocasionada pela falta de definição clara sobre o significado e conteúdo de cada metainformação. Por exemplo, na GDPR Art 13 Alínea A diz que é necessário apresentar informações de contato do controlador, mas não especifica quais dessas informações devem ser mostradas.

Foram identificadas situações de subjetividade nas metainformações tais como:

- A linha 1 da Tabela 4 apresenta uma definição genérica de **dados pessoais** e não especifica *o que apresentar ou o que deve compor a informação da Transparência*; já o dado da linha 10 da mesma tabela é mais detalhado em relação aos dados que devem ser mostrados para um determinado contexto de Transparência;
- Artigos 5 e 7 com as metainformações: Informações de compartilhamento dos dados e Localização atual do dado pessoal. Ambas as informações podem ter a mesma definição e aplicadas de forma similar, uma vez que a localização para a qual o dado é enviado faz parte de prover as informações de compartilhamento. Além disso, o termo Localização ainda cabe a interpretação uma vez que pode se referir a localização de um controlador ou processador; e
- O uso dos termos dados protegidos e anonimização. A anonimização pode ser considerada uma técnica para proteger os dados pessoais. Assim, ocorre a subjetividade na interpretação, principalmente no termo *dados protegidos*.

A subjetividade na definição e no uso das metainformações pode ocasionar a dificuldade de compreensão por parte do indivíduo, bem como o uso inadequado por controladores para proporcionar a Transparência. A interpretação incorreta por parte do indivíduo pode levar a compreensão incorreta do uso de seus dados pessoais e assim desencadear uma possível ação do indivíduo para garantir seus direitos. A possibilidade de subjetividade nas metainformações pode levar, também, ao problema de enviesamento da Transparência (já discutido) onde controladores mal intencionados divulgam a Transparência de forma intencional legislando em causa própria de modo a se beneficiar da possibilidade de manipulação do conteúdo da metainformação.

Portanto, em relação às metainformações de Transparência, foi possível concluir que os artigos contemplam um grupo de metainformações que derivam, principalmente, de

regulamentações de uso de dados pessoais. O fato dos trabalhos considerarem as metainformações permitiu concluir, que pesquisas com foco em melhorias da Transparência, estão sendo desenvolvidas por centros de pesquisa e empresas.

Pela análise das metainformações ficou evidente que há necessidade de avanços no que tange à definição e especificações de uso das metainformações, principalmente a melhoria da informação para o indivíduo e estruturação de um conjunto de regras para os controladores proporcionarem a Transparência.

Na segunda etapa da análise dos artigos da RS, foram analisados os formatos (*design*) de apresentação da Transparência para os indivíduos. Para essa questão da RS nem todos os artigos incluídos apresentaram métodos ou técnicas voltadas para a apresentação da Transparência. Dos 15 (quinze) artigos, 7 (sete) apresentaram alguma técnica de *design*. As técnicas e suas ocorrências nos artigos são mostradas na Tabela 5.

Tabela 5: Técnicas de apresentação de Transparência. Do autor.

Técnica	Descrição	Artigos
Ícones de Privacidade	A utilização de ícones com desenhos criados especificamente para privacidade e Transparência. Os ícones visam proporcionar um meio mais acessível de visualizar e entender eventos e agentes envolvidos no uso dos dados pessoais.	05, 14 e 15
Mapas e Marcadores	O uso de mapas pode auxiliar na exibição da localização de processadores, controladores, servidores de dados e demais localidades de interesse do indivíduo.	09, 11 e 14
Tabelas, Listas e Painéis	As tabelas, as listas e os painéis podem ser utilizados para apresentar informações estruturadas tais como: dados de contato de agentes, gráficos, animações e lista de processos conduzidos nos dados pessoais.	07, 09 e 14
<i>Traceview</i> e Organograma	Técnicas que podem auxiliar na compreensão do fluxo do dado pessoal nos agentes e eventos. Essa técnica pode simplificar a informação sobre compartilhamento e divulgação dos dados.	07, 11
Mapas de Cores	Indicados para apresentar a Transparência com ênfase de importância ou sensibilidade da informação ou da ação conduzida na mesma.	09 e 14
Linha de tempo (<i>timeline</i>)	Pode auxiliar no entendimento do fluxo de uso dos dados pessoais com uma visão temporal dos eventos.	07
Tutoriais e exemplos	De forma mais didática e estruturada demonstra ao indivíduo eventos envolvidos no uso dos dados pessoais.	12

Com exceção do desenvolvimento de ícones cujo foco é a abstração de informações de privacidade e Transparência, as demais técnicas são reutilizações de abordagens como *Infovis* e *Personal Infovis*, para atender a necessidade de Transparência. O uso de técnicas visuais é considerado por Haddadi et al. (2013) como uma importante ferramenta para abstrair informações de Transparência. Entretanto, a escolha da técnica e a forma de aplicação deve ser conduzida com cuidado, uma vez que a sua eficácia pode ser afetada por questões culturais. Isso porque diferentes culturas podem ter diferentes perspectivas sobre imagens, cores e formatos, a exemplo do que ocorreu na avaliação dos ícones conduzidos por Holtz, Nocun e Hansen (2011) no artigo número 15 (quinze); o mesmo destaca esse resultado, embora os ícones tenham reduzido a complexidade da informação apresentada.

Portanto, foi possível concluir que, em relação à forma de apresentação da Transparência, cabem melhorias em relação ao desenvolvimento de métodos e técnicas de modo a garantir mais eficácia na Transparência. Acredita-se que o foco no indivíduo deve ser ainda maior uma vez que o uso dos dados pessoais apresenta um crescimento considerável e pode atrair a necessidade de Transparência; e, as pessoas devem buscar meios interativos, simples, acessíveis e compreensíveis.

3.4 Considerações Finais da Revisão Sistemática

Essa RS buscou verificar quais metainformações eram consideradas como Transparência e como eram apresentadas aos indivíduos.

Os resultados mostraram que um conjunto de metainformações com base em regulamentações como a GDPR é adotado nas pesquisas, evidenciando um esforço científico para proporcionar a Transparência aos indivíduos. Entretanto, as metainformações requerem melhorias no que tange à padronização a fim de estabelecer um conjunto mínimo dessas metainformações. Essa necessidade foi identificada ao ver que os trabalhos apresentavam metainformações discrepantes em volume e conteúdo. Também foram detectadas situações de subjetividade nos conceitos das metainformações e no conteúdo apresentado por elas.

Assim, as melhorias relacionadas ao conjunto de informações que devem ser disponibilizados aos indivíduos apresenta-se como aspectos a serem discutidos nesta pesquisa e em pesquisas futuras. Definir, classificar ou padronizar metainformações de Transparência pode ser uma tarefa complexa, uma vez que necessita equilibrar os interesses dos controladores de dados, as regulamentações de uso de dados pessoais, as necessidades de

informações dos indivíduos e o contexto de uso dos dados pessoais.

Em relação ao formato de apresentação, foi verificado o uso de padrões de *design* de informação já consolidados na Arquitetura da Informação e IHC, e que são aplicadas na apresentação da Transparência. A criação de ícones voltados para privacidade e Transparência de Dados Pessoais é de extrema relevância e, certamente a base para outras iniciativas para tornar as informações sobre o uso dos dados pessoais mais acessível e compreensível ao indivíduo.

Os desafios na melhoria no formato de apresentação apontam para a necessidade de recursos para transmitir as informações de forma apropriada e orientada aos indivíduos. As informações, costumeiramente apresentadas com linguagens técnicas e jurídicas, não são favoráveis para a análise de pessoas que não procuram esse tipo de conteúdo. Ao mesmo tempo, as opções de ícones específicos para Transparência e/ou privacidade podem ser superficiais e, se não forem adequadamente desenhados, podem gerar interpretações equivocadas.

O próximo capítulo apresenta o TR-Model que é um Perfil de Aplicação de Metadados para Transparência de Dados Pessoais.

4 CONSTRUÇÃO DO TR-MODEL

Esta seção apresenta as atividades conduzidas para a construção do TR-Model. O TR-Model é um conjunto de diretrizes baseado em Perfil de Aplicação de Metadados que implementa um modelo de domínio com entidades, metadados, descrições de uso dos metadados e relacionamentos entre as entidades para apoiar a construção de aplicações de software para Transparência de Dados Pessoais com informações voltadas para análise do indivíduo titular dos dados.

4.1 Estrutura do TR-Model

A Transparência de Dados Pessoais pode ser considerada uma informação produzida por meio da organização de diversos dados relacionados a eventos e pessoas que atuam sobre os dados para um determinado objetivo. A informação pode ser utilizada para fomentar o conhecimento do indivíduo sobre o uso de seus dados e apoiá-lo na tomada de decisões sobre esse contexto.

Nesse aspecto, o TR-Model pode ser definido como um Perfil de Aplicação de Metadados para proporcionar especificações de uso de metadados a fim de disponibilizar informações sobre o uso dos dados pessoais para os titulares de dados de forma que as informações possam ser acessíveis, relevantes e compreensíveis.

O TR-Model apresenta diretrizes para utilização por equipes de desenvolvimento de software que pretendem lançar seus produtos com o atributo de qualidade de Transparência de uso de dados pessoais. As diretrizes orientam sobre quais informações apresentar e quais conceitos utilizar para apresentar.

Destaca-se que o usuário final da informação é o titular dos dados e que o TR-Model foi construído com o objetivo de atender a esse público. Assim, destaca-se que embora esforços de modelagem e especificações tenham sido feitos a fim de facilitar a implementação por equipes de desenvolvimento, o foco no indivíduo foi priorizado em relação à

facilidade de desenvolvimento de/em possíveis ferramentas, deixando a critério da equipe de desenvolvimento a seleção de eventuais recursos para a implementação do TR-Model.

Dentre as preocupações no desenvolvimento do TR-Model, considerou-se que os metadados e suas respectivas especificações pudessem dar suporte à Transparência consistente, utilitário, compreensível, não ambígua e não subjetiva para o indivíduo. Essas preocupações justificam-se pelo fato de que as informações deveriam ser utilizáveis pelos indivíduos para analisar e entender sobre o uso de seus dados pessoais e, se necessário, agir para garantir que seus direitos fossem cumpridos.

Assim, os metadados e as especificações deveriam evitar análises equivocadas por parte dos indivíduos e também orientar, de forma adequada, os desenvolvedores. Para isso, os metadados (e metaeventos - explicados posteriormente) e suas respectivas diretrizes deveriam apresentar objetivo claro para informação e as descrições deveriam contemplar um conjunto finito, porém claramente definido de possibilidades.

Para proporcionar esse tipo de estrutura, decidiu-se pela utilização da técnica de construção de Perfil de Aplicação de Metadados, pois essa técnica permitiria a criação de entidades e metadados tratando problemas de ambiguidade e generalização proporcionando uma estrutura estável e de aplicação estruturada nos software.

A construção do TR-Model foi baseada nas diretrizes de construção de Perfil de Aplicação de Metadados propostos pelo *Dublin Core Metadata Initiative*¹ e compreenderam as seguintes atividades:

1. Definição dos Requisitos Funcionais;
2. Construção do Modelo de Domínio;
3. Definição do Conjunto de Metadados e Metaeventos;
4. Definição das Descrições de Metadados e Metaeventos.

Entretanto, para essa pesquisa, três etapas foram conduzidas com ajustes em relação ao método proposto pelo DCMI a fim de atender a necessidade de construção do TR-Model: (1) não foram reutilizados metadados de outros MAPs, pois a falta de padrões de metadados com foco em uso de dados pessoais e a falta de adequação (ou excesso de ajustes) de outros padrões levaram a decisão de criar todos os metadados a partir do zero; (2) não foram consideradas questões de interoperabilidade de dados, pois o objetivo do TR-Model não é a troca de dados entre sistemas, mas a criação de diretrizes para a produção de

¹<http://dublincore.org/metadata-basics/>

informações; e (3) foi incluído o conceito de metaeventos, que são eventos relacionados às entidades que foram definidas nessa pesquisa como ações que ocorrem nos dados pessoais, que podem ser de interesse do indivíduo e podem influenciar o entendimento do uso dos dados pessoais.

Os componentes do TR-Model são mostrados na Figura 4 e contemplam: o domínio de aplicação de Transparência de Dados Pessoais que representa as regras, atores, relacionamento e demais contextos do domínio da atividade; as entidades que implementam os elementos do modelo de domínio; os metadados e metaeventos que representam quais informações devem ser mostradas como Transparência; e as especificações/descrições do metadados e metaeventos que descrevem como a informação deve ser apresentada ao indivíduo e consideram, ainda conceitos como *Readability*, *Infovis* e *Information Quality*.

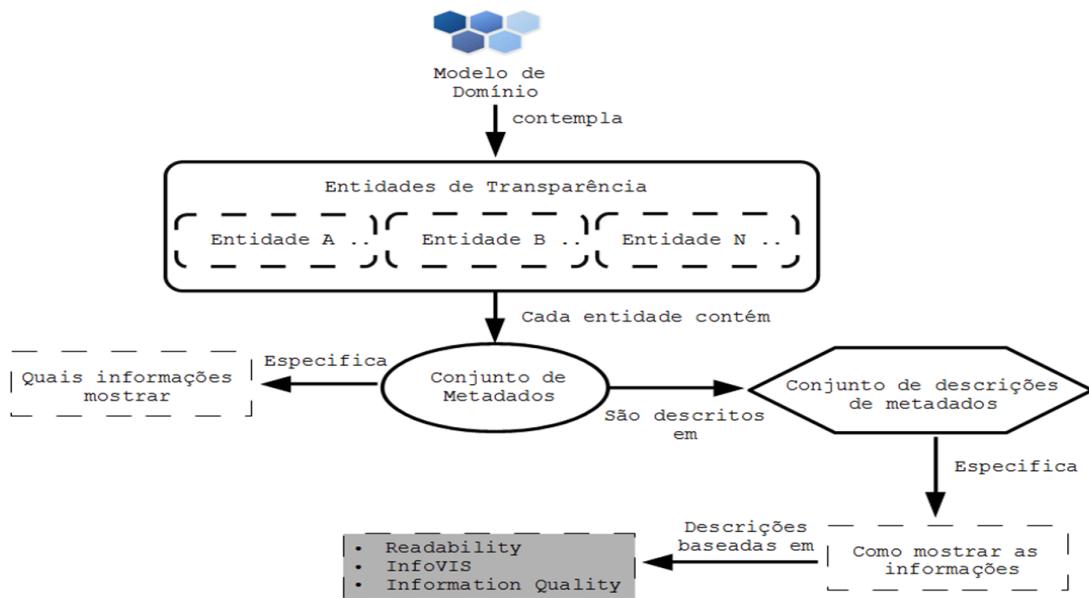


Figura 4: Estrutura do TR-Model. Do autor.

As próximas seções e subseções detalham as ações realizadas nas etapas da construção do TR-Model, a começar pela definição dos requisitos de Transparência de Dados Pessoais.

4.2 Requisitos de Transparência de Dados Pessoais

A construção do TR-Model começou pelo estudo do domínio da aplicação a fim de identificar quais requisitos estavam relacionados e influenciavam a Transparência.

Um dos primeiros requisitos identificados, com base em literaturas existentes, foi o de

que o TR-Model deveria apresentar um equilíbrio entre elementos como: (1) necessidades de usuários; (2) características das aplicações de software; (3) obrigatoriedades de leis e regulamentações; e (4) considerar preocupações das empresas que manipulam os Dados Pessoais e que aplicam conhecimentos e segredos comerciais estratégicos para obter a melhor informação dos Dados Pessoais. Essa decisão foi tomada considerando que o uso dos dados pessoais era uma área multidisciplinar envolvendo a Computação, Interação-Humano Computador, Direito e Sociologia e um elemento complexo: pessoas.

As necessidade dos indivíduos por informações de Transparência melhor apresentadas e relevantes tornou-se a prioridade no projeto do TR-Model. Em todos os momentos, assumiu-se que os indivíduos tinham o direito garantido por lei de ter acesso simples, claro e compreensível às informações sobre o uso de seus Dados Pessoais. Dessa forma, definiu-se as duas principais questões que o TR-Model deveria contemplar em relação à produção de Transparência: **quais informações deveriam ser disponibilizadas para indivíduo?**; e **como deveriam ser apresentadas (conteúdo e modo visual)?**

As necessidades dos usuários poderiam ser facilmente identificadas com o uso de técnicas de elicitação citadas por Benyon (2011) tais como questionários, entrevistas, *brainstorming* e outras comumente utilizadas na elicitação de requisitos de software. Entretanto, logo nos primeiros contatos com os indivíduos, percebeu-se que o conhecimento dos mesmos para questões de Transparência e privacidade de seus dados ainda era vago, ambíguo e incipiente. Atividades com os usuários mostraram que existia a preocupação por parte de um conjunto de pessoas, mas que as mesmas ainda não conseguiam demonstrar claramente suas necessidades, mesmo que o interesse em saber quem utiliza seus dados, e para onde são compartilhados, fosse unânime.

Por exemplo, em um *workshop* no qual foram feitas análises de PPS, um participante, ao analisar uma PPS específica atentou que o texto proporcionava um conteúdo semelhante a este: *serão coletados todos os dados necessários para o correto funcionamento da ferramenta e compartilhados ou vendidos para quem a empresa achar necessário*. Foi questionado pelo participante se essa informação poderia ser considerada uma Transparência adequada. De certa forma, o texto informava os dados que seriam coletados e o que aconteceria, mas de forma genérica e ocultando detalhes importantes que poderiam interferir na decisão de uma pessoa em concordar ou não em utilizar o software com os critérios mostrados. Desta forma, afirmar que a frase apresentada é transparente ou não pode ser uma ação subjetiva e dependente do ponto de vista do indivíduo, mas ao analisar os dizeres da GDPR é possível classificar o texto analisados como **nada ou pouco** transparente, pois a frase deixa de elencar uma série de elementos exigidos nos Artigos

13 e 14 da mesma regulamentação.

Dessa forma, assumiu-se que o levantamento dos requisitos de Transparência com base em informações de indivíduos seria ineficaz para proporcionar requisitos consistentes para o TR-Model. Portanto, decidiu-se efetuar o levantamento de requisitos com base na técnica de Análise de Domínio explicada por Pressman (2014) como uma técnica que não tem uma participação constante do usuário do sistema e considera outras fontes de informações como documentos e relatórios. O usuário, entretanto, continua participando das atividades sempre que possível, validando, discutindo e sugerindo melhorias nos artefatos entregues.

Já as leis e regulamentações foram observadas, pois entravam em vigor em diversos países e assim迫使avam um considerável número de empresas a se adaptarem a elas. Dentro as requisições das regulamentações estava a Transparência que deveria ser observada como um direito do indivíduo. As regulamentações foram utilizadas para fornecer *inputs* para a construção do TR-Model e para que o mesmo se tornasse uma ferramenta de apoio à adequação das regulamentações, pois poderia direcionar as empresas a fornecer todas informações requeridas nas regulamentações. Assim, o modelo garantiria ao indivíduo que o mesmo acessava informações requeridas e as empresas poderiam utilizar o modelo para enquadrar seus aplicativos na regulamentação.

As mesmas empresas que devem enquadrar-se nas regulamentações e garantir o direito dos indivíduos à Transparência apresentaram preocupações relacionadas aos segredos comerciais por trás do uso dos dados. Divulgar detalhes sobre o uso dos Dados Pessoais poderia forçar a divulgação de informações que poderiam comprometer o desempenho financeiro e a segurança da empresa. Assumiu-se então que a preocupação das empresas deveria ser considerada, até mesmo para facilitar a aceitação do TR-Model por elas. Portanto, o TR-Model deveria garantir o direito do indivíduo e ao mesmo tempo fornecer um conjunto de diretrizes aceitáveis pelos controladores.

Portanto, com base nas explanações sobre os artefatos de requisitos, foram adotados os seguintes para o TR-Model:

A GDPR e a LGPD foram as regulamentações de uso de dados pessoais utilizadas. A GDPR apresentava um conjunto de capítulos, artigos e seções que elencavam o que deveria ser considerado como Transparência de Dados Pessoais. Portanto, a mesma foi considerada uma importante fonte de requisitos, pois, no que tange à Transparência, era mais consistente e objetiva que outras regulamentações analisadas como a LGPD e PIPEDA. Mesmo com informações completas e objetivas, a GDPR ainda apresentava re-

quisitos de Transparência em alto nível, sem detalhes de como as informações poderiam ser organizadas e/ou apresentadas. Os textos das regulamentações seguiam uma abordagem semelhante a: *deve ser mostrado qual o tipo de dados são utilizados de forma clara e acessível para o indivíduo*, que por sinal, era uma especificação ainda muito subjetiva. Nesse aspecto, a GDPR foi muito relevante para decidir *o que apresentar*, mas deixou vago *como apresentar*.

Os **Artigos Científicos** foram estudados para identificar contribuições realizadas para a Transparência de Dados Pessoais. Os artigos mostraram que a Transparência era uma preocupação em centros de pesquisa e empresas, mas que seus conceitos, estruturas e formas de aplicação ainda tinham muitos aspectos carentes de definição tais como: um conjunto de informações organizadas a respeito dos agentes e eventos envolvidos no uso dos dados pessoais. Destaca-se nestes artigos trabalhos como Murmann e Fischer-Hübner (2017a), Mortier et al. (2016), Haddadi et al. (2013) e Patrick (2015) este último apresentou um conjunto de princípios de privacidade que amparam fortemente as entidades de Transparência discutidas nesta pesquisa.

Websites sobre o assunto foram utilizados, quando considerados confiáveis, uma vez que o uso dos Dados Pessoais é assunto destaque de vários *blogs*, páginas de notícias e *websites* sobre informática, direito e legislação. Além de websites de assuntos gerais foram considerados *websites* como o <http://hdiresearch.org> ou *website* de notícias como o *Dois na Web* (De Luca, 2019) o qual disponibiliza vasto material sobre leis, regulamentações e ações sobre o uso de dados pessoais.

Os **livros** foram analisados para estudar as referências clássicas dos temas da pesquisa tais como *Big Data*, Ciência dos Dados, Engenharia de Requisitos, *Readability*, IHC e Qualidade de Informação. Com os conceitos clássicos foi possível melhorar o entendimento das áreas de estudos que envolviam a Transparência.

Os **Indivíduos** ou **Titulares dos dados** participaram das atividades de levantamento de requisitos, palestras, *wokshops* e entrevistas nas quais eram discutidas questões sobre Transparências e apresentadas situações para análise e discussão. Embora com conhecimento ainda pouco apurado sobre o uso dos dados pessoais, os indivíduos foram extremamente importantes, uma vez que forneceram expectativas sobre as informações de Transparências.

As necessidades dos indivíduos em conjunto com recomendações levantadas nos demais artefatos produziram a matriz de requisitos apresentada na Tabela 6.

A coluna **indivíduos** destaca as necessidades de Transparência identificadas nos

Tabela 6: Matriz de requisitos de Transparência. Do autor.

Requisito	Indivíduos	GDPR	LGPD	Artigos
Apresentação da informação relevante e com qualidade para os indivíduos	X	X	X	
Clareza na exibição das informações	X	X	X	X
Dados de quem tem acesso aos dados pessoais	X	X	X	X
Estratégia de coleta e uso dos dados (intervalo, dispositivo etc)	X			
Informações sobre a legalidade do uso dos dados	X	X	X	
Informações sobre a anonimização dos dados				X
Informações sobre como cancelar ou mudar a permissão de uso dos dados pessoais	X	X	X	X
Informações sobre como denunciar o mau uso dos dados pessoais	X			X
Informações sobre o compartilhamento ou divulgação dos dados pessoais	X	X	X	X
Informações sobre os processos realizados nos dados pessoais	X	X		X
Facilidade de acesso as informações	X	X	X	X
Objetivo de uso dos dados pessoais	X	X	X	X
Relação de direitos dos titulares dos dados				X
Relação de dados pessoais coletados/manipulados	X			X

workshops, palestras e entrevistas; já as colunas GDPR e LGPD apontam os requisitos destacados nas regulamentações; e a coluna Artigos destaca as necessidades de informações discutidas nos artigos científicos, principalmente aqueles apresentados no capítulo de Revisão Sistemática.

A análise de todos os artefatos e a definição da matriz de requisitos ² permitiram estabelecer o requisito funcional que nortearia o desenvolvimento das entidades, atributos e descrições do TR-Model como: *Proporcionar um conjunto de informações sobre o uso dos Dados Pessoais elencando agentes e eventos envolvidos na coleta, processamento, utilização e distribuição dos dados pessoais. As informações devem ter formato de apresentação apropriado, além de acessíveis, concisas, completas, compreensíveis e relevantes para os indivíduos.* .

²Os websites apresentavam notícias e materiais complementares, mas não foram utilizados para auxiliar na montagem da matriz de requisitos.

No que tange aos elementos do requisito funcional, os agentes e eventos foram implementados pelas entidades, metadados e metaeventos. Já os formatos de apresentação e elementos de IQ como objetividade, concisão, completitude e relevância foram definidas pelas descrições dos metadados e metaeventos, conforme será apresentado nas seções seguintes.

Com o requisito funcional do TR-Model definido, passou-se para a etapa de criação do Modelo de Domínio, que será apresentada na próxima seção.

4.3 Modelo de Domínio do TR-Model

O Modelo de Domínio do TR-Model contempla o conjunto de entidades que descrevem agentes e eventos envolvidos no uso de dados pessoais e que podem ser considerados como Transparência de Dados Pessoais. As entidades foram criadas para implementar os elementos do domínio de aplicação e então atender ao requisito funcional apresentado na seção anterior. Cada entidade proposta representou um agente, evento ou elemento considerado relevante para entender o fluxo de uso dos dados pessoais.

Todas as entidades identificadas tinham, pelo menos, uma relação que complementava suas informações com outra entidade; ou complementava outra entidade com suas informações e tornava a Transparência mais completa e coerente.

O modelo de domínio do TR-Model é mostrado na Figura 5, na página 60.

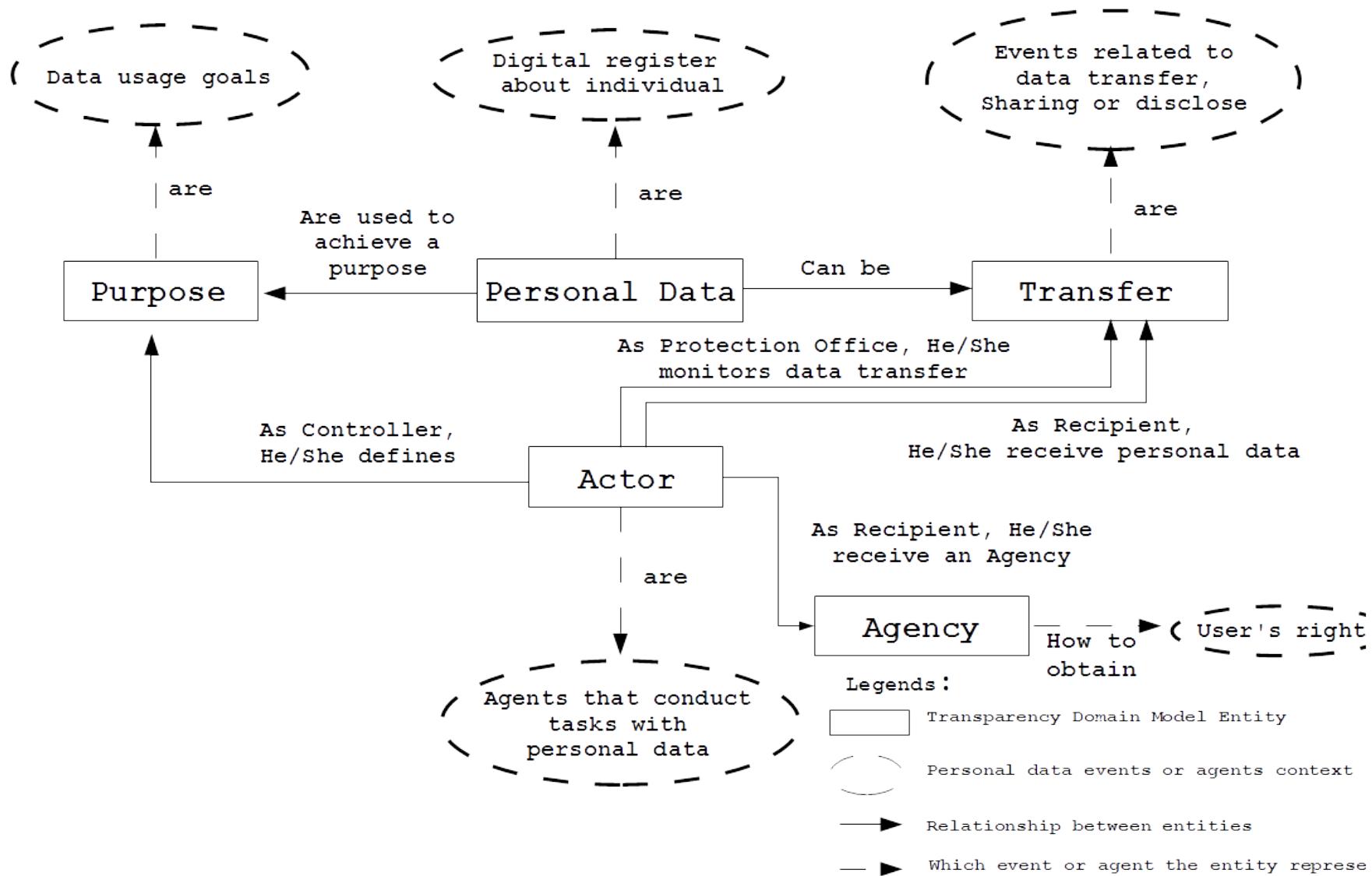


Figura 5: Modelo de domínio para o TR-Model. Do autor.

As entidades do modelo de domínio e os elementos que cada uma delas representam são explicados nas próximas subseções.

4.3.1 *Actors*

Esta entidade organiza informações sobre pessoas, empresas ou organizações legalmente constituídas que estão envolvidas no uso dos dados pessoais. No contexto de Transparência os atores desempenham funções distintas como: coleta dos dados; definição de regras de uso; processamento; armazenamento; distribuição/compartilhamento; aquisição e fiscalização.

A GDPR, por meio de seus Artigos 13, 14 e 15 estabelece regras bem definidas em relação aos direitos e deveres dos atores em relação às ações realizadas e às responsabilidades com os dados pessoais, ou em relação aos direitos adquiridos ao produzir e entregar os dados para alguém utilizar.

Considerando as possibilidades de atuação dos atores no ciclo de vida dos dados pessoais, e a relação de responsabilidade estabelecida para cada um, foi necessário criar uma classificação para os mesmos de forma a facilitar a identificação e atribuição de responsabilidades dos atores para com as suas ações. As classificações herdaram nomenclaturas e características apresentadas no Artigo IV (*Definitions*) da GDPR e são:

- **Indivíduo ou Usuário ou Titular dos dados:** pessoa que, por meio da interação com um dispositivo tecnológico como um *smartphone*, uma rede social ou uma rede de sensores produz dados a seu respeito. Por representar a pessoa, esse ator aparece nas regulamentações como sendo detentor de um conjunto de direitos tais como: acessar informações sobre o uso dos dados pessoais; requerer detalhes do processamento; obter cópia dos dados; ou de realizar alguma denúncia ou intervenção quando necessário. As diretrizes da GDPR, em sua maioria, procuram estabelecer ações de proteção e defesa do indivíduo a fim de evitar abusos com o uso de seus dados pessoais;
- **Controlador:** organização pública ou privada que determina os propósitos de uso dos dados pessoais e estabelece critérios de processamento, armazenamento e distribuição. O controlador deve exercer suas funções de acordo com leis e regulamentações de proteção de dados. O controlador é o mais cobrado em relação às responsabilidades uma vez que cabe a ele garantir que os direitos dos indivíduos sejam preservados. Também cabe ao Controlador responder por qualquer situação

considerada anormal ou irregular no uso dos dados;

- **Processador/Operador:** organização pública ou privada responsável por realizar atividades computacionais referentes ao processamento ou armazenamento dos dados. O Processador pode (ou não) ser a mesma empresa a que o controlador representa, pois um controlador pode não deter recursos para executar atividades computacionais ou não ter interesse em realizá-las e, consequentemente, terceiriza essa tarefa. Por compartilhar as ações de uso dos dados, principalmente no aspecto de acesso e processamento, o processador responde solidariamente ao controlador por quaisquer danos causados ao indivíduo, principalmente em relação à segurança e privacidade. Em relação aos detalhes do processamento, as leis asseguram a garantia de segredo comercial, mas recomendam ao processador que disponibilize informações mínimas que permitam ao indivíduo entender como seus dados são tratados, principalmente se o processamento levar a uma tomada de decisão realizada pelo computador sem a supervisão humana;
- **Escritório de Proteção de Dados:** organização pública independente estabelecida por um país ou organização de países que é responsável por fiscalizar e garantir a correta utilização dos dados pessoais conforme as regulamentações vigentes. No contexto de Transparência, os escritórios devem garantir que o indivíduo tenha acesso garantido, claro e simplificado às informações de uso dos dados pessoais assim como acesso aos mecanismos de denúncia ou requisições;
- **Destinatário:** O destinatário pode atuar no fluxo de uso dos dados pessoais em duas perspectivas: (1) Destinatário dos dados: um controlador que recebe dados pessoais compartilhados, divulgados ou transferidos pelo controlador que os obteve; e (2) Destinatário de uma requisição: o ator recebe alguma solicitação ou reclamação do indivíduo.

Considerando as ações que cada tipo de ator desempenha e que consequentemente impacta na Transparência, as informações dos atores são utilizadas para completar e dar coerência e relevância para informações de outras entidades.

4.3.2 *Purpose of use*

Esta entidade tem por objetivo fornecer Transparência sobre o propósito de uso dos dados pessoais. O propósito (ou objetivo de uso) é sempre definido pelo Controlador ou por um conjunto de controladores, chamados na GDPR como *Jointly* de controladores.

Obter informações sobre o propósito de uso foi uma das principais queixas de indivíduos durante os *workshops*. No que tange às regulamentações, informações sobre o propósito de uso são garantidos na GDPR nos Artigos 13 e 14 Seção 1, Alínea c.

A Transparência sobre o propósito de uso é pouco (ou nada) divulgada, pois são ocultadas ou disfarçadas em complexos textos nas PPS ou inseridas com viés intencional em telas de concessão de autorização sob uma perspectiva de *ou autoriza tudo ou não usa nada*. Assim, não é incomum um indivíduo ser surpreendido por algum evento de um aplicativo ou *website* ocasionado pela coleta e processamento dos dados como: direcionamento de informação, propagandas ou previsão de necessidades dos indivíduos, pois os mesmos autorizaram sem ter um conhecimento bem formado sobre o porquê do uso dos dados. Há casos em que a dificuldade em acessar o objetivo de uso dos dados está disfarçada na necessidade de segredos comerciais/industriais dos controladores.

Assim, a entidade de Propósito de uso foi criada para abstrair informações a fim de esclarecer *o porquê* do uso dos dados pessoais, quem é responsável pelo uso e as características do propósito. Com essa entidade acredita-se ser possível esclarecer ao indivíduo o motivo do uso de seus dados, proporcionar maior confiança na aplicação e evitar, ou minimizar, surpresas em decorrência do mau uso dos dados.

4.3.3 Personal Data

A entidade **Personal Data** é o centro de todo o modelo de domínio uma vez que os dados pessoais são os principais objetos de uso para produção de informações sobre os indivíduos. A entidade *Personal Data* foi criada para melhorar a compreensão sobre quais dados são utilizados e como eles são obtidos pelo controlador.

Relatos já citados em seções anteriores mostram que recursos já existentes informam sobre os dados pessoais, mas de forma técnica, o que dificulta o entendimento do indivíduo sobre *o que é coletado e como é coletado*. Exemplo disso são TETs que fornecem o número do IP da máquina do indivíduo ou o endereço MAC. Além de ser uma informação pouco (ou nada) útil para análise, tal informação é complexa para ser compreendida por algum leigo na área de informática.

Por ser uma informação extremamente sensível ao contexto de Transparência e de indispensável relevância para o titular dos dados, buscou-se propor uma entidade com informações objetivas, claras e relevantes aos indivíduos. Assim, essa classe foi criada para informar quais os dados pessoais são coletados e utilizados. A informação deve ser

mostrada em uma linguagem menos técnica, voltada ao indivíduo leigo e priorizar os seguintes aspectos: (1) quais dados são coletados; (2) de onde são coletados; (3) como são coletados; e (4) o que é feito com eles.

4.3.4 Transfer

A entidade **Transfer** busca proporcionar Transparência em uma das principais preocupações dos indivíduos dos dados, a distribuição dos dados pessoais e seu eventual impacto na privacidade, segurança e controle do titular dos dados.

As preocupações dos indivíduos foram consideradas na GDPR, que exige informações sobre o compartilhamento dos dados pessoais com terceiros, com ênfase na distribuição para outros países. Algumas TETs apresentam informações como Localização do destinatário em um mapa; Endereço IP do destinatário ou permitem ao usuário selecionar o perfil de quem pode receber os dados. A localização e a seleção dos destinatários são aspectos relevantes, diferentes do endereço IP. Entretanto, acredita-se que tais informações são insuficientes para proporcionar uma Transparência completa que transmita conteúdo suficiente para auxiliar o indivíduo na identificação de distribuição indevida de seus dados para que o mesmo possa intervir ou negociar a respeito.

Assim, essa entidade foi criada para permitir ao indivíduo conhecer o destinatário, entender o objetivo da distribuição dos dados e se a mesma é amparada por artefatos legais.

4.3.5 Agency

A entidade **Agency** foi criada para prover Transparência sobre quais ações o indivíduo deve conduzir para agir e/ou negociar o uso de seus dados pessoais a fim de garantir que seus direitos sejam cumpridos.

Destaca-se que esta entidade visa **informar** como realizar a ação e não necessariamente disponibilizar o recurso para realizar a ação que pode ser feita pelo Controlador, Processador ou Escritório de Proteção de Dados em recursos como o *website* da empresa, formulário específico, endereço de *e-mail* ou outros meios que o ator venha disponibilizar.

4.3.6 Considerações sobre as entidades do TR-Model

O número de entidades do TR-Model foi de 5 (cinco) entidades. As entidades apresentadas abstraem informações sobre agentes, eventos ou elementos envolvidos no uso dos dados pessoais e que podem ser caracterizados como informações de Transparência.

Assume-se que, com essas entidades e seus respectivos metadados, metaeventos e descrições (apresentadas nas seções seguintes) possam ser suficientes e apropriadas para um indivíduo visualizar, entender e decidir sobre o uso de seus dados pessoais.

Existiu (e foi discutida) a possibilidade de criar mais entidades para exibição de um conjunto maior de informações. Entretanto, assumiu-se que as informações discutidas seriam de caráter técnico e somariam no aspecto de quantidade de informação, mas não foi possível identificar contribuições significativas em relação à qualidade e relevância do conteúdo para o titular dos dados.

A próxima seção apresenta a definição dos metadados e metaeventos.

4.4 Definição e Descrição dos Metadados e Metaeventos das entidades

A definição dos metadados, metaeventos e descrições das entidades do TR-Model foi realizada com o objetivo de definir: (1) **quais informações seriam proporcionadas aos usuários como Transparência de Dados Pessoais** no caso do metadados e metaeventos e; (2) **como as informações deveriam ser apresentadas aos indivíduos**, nesse caso, as descrições dos metadados.

As entidades do TR-Model contemplam o modelo de domínio para Transparência de Dados Pessoais, mas para fornecer informações sobre o uso dos dados era necessário definir as características de cada entidade. O modelo de domínio mostrado na Figura 5 na página 60 permite inferir certas características que poderiam ser representadas, mas essa pesquisa buscou, justamente, evitar que diferentes interesses ou diferentes interpretações proporcionassem informações incompatíveis com a necessidade do indivíduo.

As características foram classificadas como **Metadados** referindo-se às características das entidades que compõem o fluxo de uso dos dados pessoais; e **Metaeventos** para fazer referência aos eventos realizados pelas entidades e que influenciam na utilização dos dados pessoais. Os metaeventos foram inseridos em classes, as quais o entendimento dos eventos torna-se fundamental para compreender por completo como os dados são coletados,

utilizados e/ou distribuídos. De forma geral, os metadados e metaeventos definem quais informações devem compor uma entidade de Transparência de Dados Pessoais.

Os metadados e metaeventos foram propostos com o objetivo de balancear os interesses dos indivíduos (informações sobre o uso de seus dados); os interesses dos controladores/processadores: manter detalhes técnicos como segredos comerciais; e as obrigatoriedades de Transparência requeridas pela GDPR.

Os metadados e metaeventos definidos para as entidades do TR-Model são mostrados na Figura 6 e explicados nas próximas subseções em conjunto com as explicações sobre as descrições dos metadados e metaeventos.

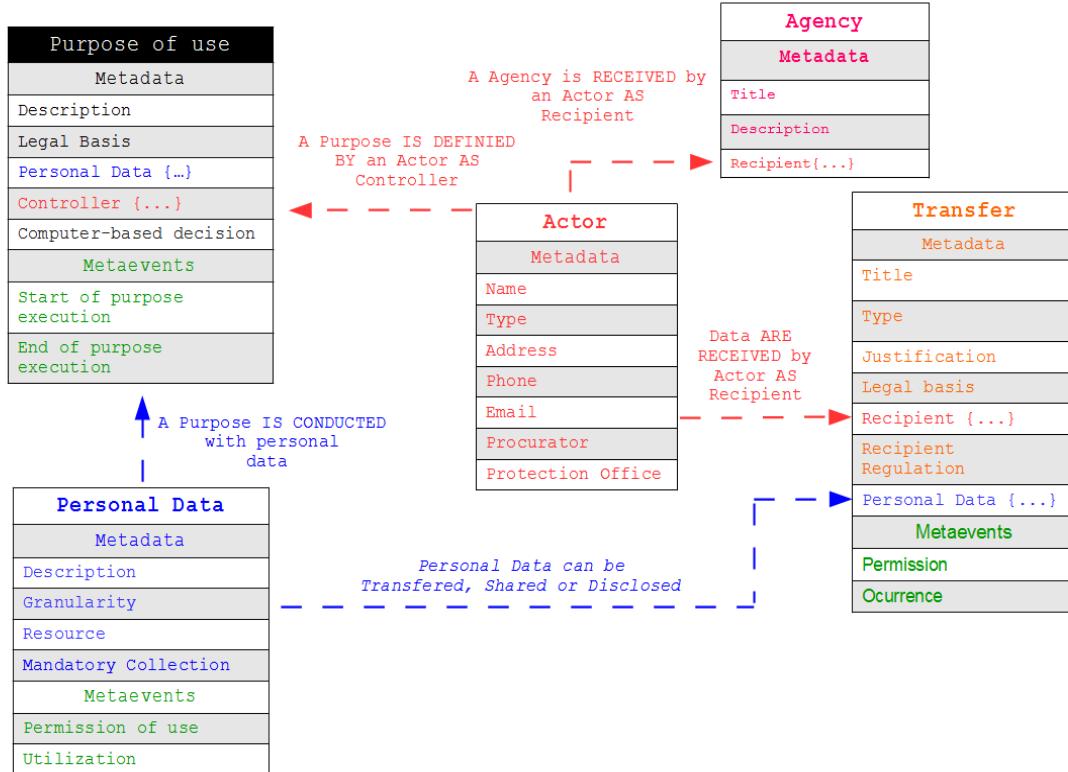


Figura 6: Metadados e Metaeventos das entidades do TR-Model. Do autor.

As descrições dos metadados e metaeventos orientam como cada informação de Transparência deve ser apresentada para o usuário. As descrições buscaram estabelecer **como apresentar a Transparência para o indivíduo**.

As descrições foram criadas para atender aos titulares dos dados, pois as informações produzidas com base nelas deveriam ser utilizáveis, comprehensíveis, relevantes e agradáveis

para os mesmos. Para apoiar a criação das descrições foram considerados os conceitos de IHC como o *Readability* e *Infovis* além do conceitos e dimensões de Qualidade de Informação (IQ).

Os conceitos de *Readability* foram aplicados a fim de propor textos curtos e focados na necessidade do indivíduo uma vez que um dos problemas da Transparência em técnicas como a PPS é o grande volume de textos apresentados. Com o uso da característica de produção de **senteças curtas**, assumiu-se que a *Readability* poderia auxiliar no quesito forma de apresentação da Transparência e também na facilidade de localização e identificação das informações.

Embora o uso de textos pode não ser a melhor opção para todos os tipos de usuários, assumiu-se que o uso de sentenças curtas e bem elaboradas pode ser eficaz na transmissão da informação. Outro fator considerado foi a sensibilidade de algumas informações as quais poderiam ser interpretadas (ou até não interpretadas) se apresentadas com outras técnicas.

A *Infovis* foi utilizada em descrições de metadados e metaeventos, os quais julgou-se necessário que as informações fossem apresentadas com recursos como imagens, gráficos, linhas de tempo, infográficos, mídias, tabelas e demais *design patterns* tais como os apresentados por Neil (2014) e Mew (2016), uma vez que poderia facilitar ou melhorar a capacidade de entendimento e compreensão dos indivíduos.

Já a IQ foi considerada a fim de garantir que determinadas dimensões fossem aplicadas nas descrições. Problemas como a falta de interesse do indivíduo na informação, informações incompletas ou subjetivas além da falta de relevância na informação para o indivíduo deveriam ser mitigadas com o suporte das dimensões de IQ selecionadas.

As dimensões de IQ selecionadas para essa pesquisa são mostradas na Tabela 7. No contexto do TR-Model, o termo informação pode ser associado à Transparência de Dados Pessoais, pois a mesma trata de **informações** sobre dados pessoais.

Por fim, os elementos de IHC e a IQ apoiaram a construção das descrições com o objetivo de proporcionar informações de Transparência com formatos orientado à utilização dos indivíduos de forma geral, além de proporcionar informações relevantes, completas, comprehensíveis e objetivas.

Os metadados, metaeventos e descrições serão apresentados em tabelas com as seguintes colunas e conteúdos:

- **Nome:** Nome do metadado ou metaevento. Pode ser usado como rótulo da in-

Tabela 7: Dimensões de IQ selecionadas para TR-Model. Do autor.

Dimensão	Descrição
Quantidade de Informação	A quantidade de informação apresentada deve ser adequado ao indivíduo. No contexto de Transparência, deve limitar-se às informações de uso dos dados pessoais.
Capacidade de compreensão	A informação deve ser apresentada de forma que o usuário possa compreender o conteúdo transmitido.
Objetividade	A informação deve ser livre de viés, de parcialidade e de conteúdos prejudiciais.
Completitude	A informação deve estar completa para dar suporte ao usuário sem a necessidade do mesmo procurar informações complementares em outras fontes.
Relevância	A informação deve ser relevante e deve contribuir para o indivíduo analisar e compreender como seus dados pessoais são utilizados.

formação em uma interface de software;

- **Descrição do Metadado ou Metaevento:** Apresenta a descrição do metadado ou metaevento e um exemplo de aplicação. Seu conteúdo é subdividido em:
 - Tipo: Formato de apresentação da informação: Texto, Figura, Mapas etc. O Formato sugere o meio pelo qual a informação pode ser apresentada, mas o *design pattern* a ser adotado fica a cargo do *designer*;
 - Conteúdo: Qual conteúdo sobre a Transparência dos Dados Pessoais deve ser mostrado.

As próximas seções apresentam os metadados, metaeventos e descrições para as entidades do TR-Model.

4.4.1 Metadados e descrições da entidade *Actors*

Durante os *workshops* e palestras, quando questionado aos indivíduos sobre quais informações sobre seus dados pessoais eles gostariam de saber, uma das respostas era: *Gostaria de saber quem utiliza meus dados*. Entretanto, os mesmos não detalhavam quais dados, em específico, gostariam de saber. Acredita-se que essa informação era omitida, pois sabendo quais empresas ou pessoas seriam, procurar mais informações a respeito na Internet seria uma tarefa simples.

Já a GDPR , Capítulo 1, Artigos 13 e 14 descrevem a exigência de informações para permitir o **contato com o ator** que realiza uma tarefa com os dados, assim como na

LGPD que destaca a necessidade pelo **contato do controlador ou do operador**. O termo *dados de contato* é colocado de forma subjetiva na GDPR o que facilita interpretações variadas.

Assim, concluiu-se que as informações de contato seriam necessárias para complementar a Transparência de outras entidades a serem discutidas neste texto, como *Purpose of use*, para identificar o controlador e *Transfer* para identificar o destinatário e o órgão de proteção de dados. Decidiu-se, portanto, estabelecer um conjunto de metadados para permitir ao indivíduo obter informações de identificação nos moldes de uma ficha de identificação cadastral bem como informações de localização e meios para estabelecer algum contato.

Os metadados propostos acabaram por deixar a entidade com um aspecto semelhante ao de uma ficha cadastral, muito comum em software. O perfil da entidade dispensou a necessidade de metaeventos, pois não foram identificadas ações com as informações dos atores que interferissem na compreensão do uso dos dados pessoais.

As descrições dos metadados de atores são relativamente simples uma vez que as regulamentações não deixam claro o que deve ser considerado como *dados de contato*. Assim, buscou-se simplificar o máximo para o indivíduo nessa entidade.

Os metadados e as respectivas descrições para a entidade *Actors* são mostradas na Tabela 8.

Tabela 8: Metadados e Descrição dos metadados da entidade *Actors*. Do autor

Metadado	
Nome	Descrição do Metadado
Name	Tipo: Texto; Conteúdo: Nome comercial do ator
Type	Tipo: Texto ou Imagens; Conteúdo: Especificar qual papel o ator desempenha, dentre as opções: (1) Controlador; (2) Escritório de Proteção de Dados; (3) Processador ou Operador; e (4) Destinatário
Address	Tipo: Texto ou Imagem de mapa com marcador de localização; Conteúdo: Deve conter as informações: (1) Nome da rua; (2) Número ou complemento; (3) Nome da cidade; (4) Estado ou província; (5) Código Postal; e (6) País.
Phone	Tipo: Texto; Conteúdo: Código do País + Número do telefone.
Email	Tipo: Texto; Conteúdo: Endereço de e-mail associado ao ator.
Procurator	Tipo: Texto; Conteúdo: Nome da pessoa que responde pelo ator (em caso de empresas).
Protection Office	Tipo: Amostra da entidade <i>Actor</i> ; Conteúdo: Informações de um ator com o metadado <i>Type</i> com valor igual a Escritório de Proteção de Dados .

A próxima subseção apresenta os metadados, metaeventos e descrições da entidade *Personal Data*.

4.4.2 Metadados, metaeventos e descrições da entidade *Personal Data*

Os dados pessoais são o centro das informações de Transparência. O volume e a forma como os dados são obtidos geram preocupações como: quais dados são coletados; como são coletados; o que é feito com os dados; foi autorizado fazer; a empresa que coleta os dados tem direito de fazer isso.

As regulamentações não são específicas em relação ao que deve ser mostrado sobre os dados pessoais; a preocupação maior está em relação ao que é feito com os dados (situação discutida nas próximas subseções). Entretanto, durante entrevistas, palestras e conversas com indivíduos foi detectado uma preocupação muito grande dos mesmos em relação às características de seus dados pessoais utilizados.

A preocupação é justificada por situações vividas pelos mesmos que causaram surpresas, preocupações ou transtornos, por exemplo: pessoas que falaram sobre algum assunto em uma conversa informal e, de repente, em uma rede social ou *website* de busca começou a direcionar conteúdos sobre o assunto; ou empresas que tinham acesso aos dados de alguma pessoa, mas a pessoa nunca esteve na empresa.

Considerando tais fatos, foi necessário produzir para essa entidade tanto metadados quanto metaeventos. Os metadados descreveram as características do dado pessoal coletado, assim como informações dos recursos tecnológicos que obtinham os dados. Já os metaeventos foram estabelecidos, pois foram identificados eventos que ocorriam com os dados tais como permissão de uso ou formas de processamento que eram de interesse dos indivíduos e influenciavam a compreensão do uso dos dados.

Como características dos dados, percebeu-se a necessidade pela descrição (nome do dado pessoal); sua granularidade (nível de detalhamento); qual recurso era utilizado para coleta do dado; e a obrigatoriedade da coleta. Destaca-se entre os metadados a **granularidade**.

O metadado de granularidade foi estabelecido para facilitar a Transparência quando o dado pessoal é formado por muitos fragmentos e, ao mesmo tempo, facilitar a forma de transmissão da informação ao indivíduo. A granularidade permite abstrair elementos

menores de dados que, de certa forma, compõem o dado pessoal, o que acredita-se facilitar o entendimento do usuário.

Por exemplo: Considerando um aplicativo de mapas (GPS) que registra regularmente a posição em que o usuário se encontra. O dado pessoal pode ser considerado como **Localização**. Uma localização pode ser composta por informações como Latitude, Longitude, Data e Hora. Os elementos que compõem, ou seja, formatam o dado pessoal são definidos como a granularidade do mesmo.

A descrição dos metadados dessa entidade requereu certos cuidados uma vez que a compreensão do indivíduo de quais dados pessoais são usados e como eles produzem e fornecem é essencial para entender demais ações. Além disso, alguns metadados têm aspecto técnico e precisaram ser abstraídos para um linguagem mais popular, como o metadado *Source*.

A relação dos metadados e suas descrições são mostrados na Tabela 9. Já a relação e a descrição dos metaeventos da entidade *Personal Data* são mostrados na Tabela 10 na página 73.

Tabela 9: Descrição dos Metadados da Entidade *Personal Data*. Do autor.

Nome	Descrição dos Metadados
Description	Tipo: Texto; Conteúdo: Descrição ou identificação do dado pessoal. Exemplo: (1) <i>Localização</i> ; (2) <i>Registro de compra do cartão de crédito</i> ; ou (3) <i>Registro da atividade física</i> .
Granularity	Tipo: Texto ou Imagem explicativa; Conteúdo: Combinação de itens que compõem o dado pessoal. Por exemplo: <i>O dado pessoal Localização é composto por: Latitude + Longitude + Data + Hora</i> .
Source	Tipo: Texto ou Imagem ou Vídeo explicativo; Conteúdo: O software deve apresentar uma quantidade mínima de frases concisas e objetivas descrevendo os recursos utilizados para criar os itens de dados. Se o recurso for incomum para as pessoas, o sistema deve complementar as informações com uma breve descrição do recurso. Exemplo: <i>Os dados de latitude e longitude são registrados pelo sistema GPS do smartphone. O GPS é um recurso interno de hardware que não pode ser visualizado (apenas acessado via software) pelos usuários</i> . Como alternativa orientada a <i>Infovis</i> , o software pode utilizar imagens, vídeos ou animações para mostrar os recursos. Exemplo: Uma foto de um celular a qual destaca-se a câmera e o microfone pode ajudar um leigo a entender quais componentes coletam suas imagens e áudio.
Mandatory Collection	Tipo: Texto; Conteúdo: Deve apresentar se a coleta de dados é obrigatória. Caso a informação seja SIM, a mesma deve ser seguida por uma sentença que justifique o motivo da coleta compulsória. Se a coleta não é compulsória, o software deve exibir uma sentença explicando o que pode ocorrer se o dado pessoal não for coletado.

Com os metadados e metaeventos selecionados para essa entidade, assumiu-se que os indivíduos têm condições de analisar quais dados são coletados, como são coletados, como são usados e obter uma visão completa do ciclo de vida do dado pessoal. Com a entidade *Personal Data*, o indivíduo poderá compreender como uma aplicação de software selecionou ele/ela para uma promoção; ou como aprendeu determinado comportamento; como um software ou empresa sabe sobre ele/ela.

O entendimento completo de como os dados pessoais são utilizados dá-se com a utilização das informações da entidade *Personal Data* em outras entidades que destacam ações específicas dos dados pessoais como *Purpose of use* e *Transfer* que, respectivamente tratam do propósito de uso e de ações de destruição dos dados (entidades explicadas nas próximas subseções).

A próxima subseção apresenta os metadados, metaeventos e descrições para a entidade *Purpose of use*.

Tabela 10: Descrição dos Metaeventos para entidade *Personal Data*. Do autor

Nome	Descrição dos metaeventos
Permission of use	Tipo: Imagem ou Vídeo; Conteúdo: Este evento deve considerar o momento ou a ação que o indivíduo realizou para conceder permissão de uso para seus Dados Pessoais. Para apresentá-lo ao usuário, é sugerido um conjunto de imagens da interface usada para dar o consentimento, a fim de lembrar / explicar ao usuário sobre o momento ou ação realizada para permitir o uso de seus Dados Pessoais. O uso de animações, cores, destaque podem ser utilizados para auxiliar na identificação da informação. As imagens podem ser complementadas por textos como: <i>Veja quando / como você autorizou o uso dos seus dados</i> ou <i>Veja quando / como você permitirá o uso dos seus dados</i> .
Utilization	Tipo: Texto ou Animações explicativas; Conteúdo: Esse metaevento pode fornecer transparência sobre como os dados serão usados e quais informações sobre o usuário podem ser produzidas pelo uso dos dados. A descrição das informações deve usar uma linguagem sem o uso de conceitos técnicos. A explicação da complexidade do uso de dados seja abstruída para uma linguagem acessível às pessoas e seguir uma estrutura como: <i>Os dados pessoais serão usados da seguinte maneira {descreva a forma de uso} e as informações produzidas ou as perguntas respondidas são: lista de informações produzidas ou perguntas respondidas sobre a pessoa. Nos casos com vários usos diferentes dos dados pessoais, o sistema deve repetir as especificações para cada uso.</i>

4.4.3 Metadados, metaeventos e descrições da entidade *Purpose of use*

Essa entidade proporciona informações sobre o objetivo do uso dos dados pessoais. O propósito de uso é determinado pelo controlador que, assim, estabelece nessa entidade um vínculo com a entidade *Actors*, pois trata-se de um tipo de *Actor* envolvido no uso dos dados pessoais e, de certa forma, a mais influente além do indivíduo propriamente dito.

As características do propósito de uso foram definidas como a descrição do propósito de uso; informações sobre sua legalidade; quais dados pessoais contemplavam o propósito; e a informação sobre decisões feitas exclusivamente por computador (sem supervisão de um humano).

Os dados pessoais são informações fornecidas pela entidade *Personal Data* que estabelece um vínculo similar ao citado anteriormente para a entidade *Actor*. Para cada dado pessoal utilizado no propósito de uso é estabelecido um vínculo com a entidade *Personal Data*.

Já o atributo nomeado como *Computer-based decision* atende a GDPR que orienta a

exibição de uma informação a respeito da capacidade de tomada de decisão sem a supervisão humana. A necessidade dessa informação é justificada no Artigo 22 que estabelece que nenhum indivíduo é obrigado a ser submetido a qualquer tipo de ação feita única e exclusivamente por computador. Se necessário a execução, há necessidade de consentimento explícito do indivíduo. Assim, esse atributo atua como um mecanismo de checagem do indivíduo a fim de garantir que seus direitos estão preservados pelo propósito de uso e que alguma informação incorreta seja produzida sobre o indivíduo.

Como metaeventos, identificou-se que as ações relacionadas ao propósito de uso que poderiam interferir na análise do uso dos dados seriam: o início e o término da execução do propósito. Essas informações são justificadas por: (1) o indivíduo ter conhecimento do período em que o uso dos dados poderá acontecer; (2) ter conhecimento de qual ou quais ações ele realizou dentro do período e que podem ter sido observadas com os dados; (3) garantir que o uso dos dados seja feito dentro do período do consentimento.

A descrição dos metadados (Tabela 11) e metaeventos (Tabela 12) dessa entidade foram propostos com base em textos (*readability*) isso porque as informações envolvem aspectos legais, descrição de leis, além da necessidade da explicação clara e simples sobre o objetivo do uso dos dados. Portanto, a descrição dos metadados e metaeventos exigiram um equilíbrio fino entre a necessidade de informar claramente o indivíduo com textos e ao mesmo tempo, manter a simplicidade e objetividade para o indivíduo.

Tabela 11: Descrição de Metadados para a Entidade *Purpose of use*. Do autor.

Metadata	
Nome	Descrição dos metadados
Purpose description	Tipo: Texto; Conteúdo: Deve descrever o nome do objetivo de uso e os interesses legítimos do uso do Dados Pessoal pelo controlador. Para fazer isso, deve-se utilizar uma única frase especificando o propósito de uso, conforme a estrutura: "O propósito de uso para seu dado pessoal é {descrição do propósito}". Exemplo: <i>O objetivo dos seus dados pessoais é conhecer seus principais destinos e rotas durante a semana para oferecer alternativas melhores ou, O objetivo do uso de seus dados pessoais é conhecer suas preferências de marcas e modelos.</i>
Legal Basis	Tipo: Texto ou Infográfico; Conteúdo: Uma frase que reafirme o propósito de uso dos dados seguida de informações sobre a lei/regulamentação que garante a legalidade do propósito apresentado. Exemplo: <i>O uso de informações pessoais para conhecer seus principais destinos durante a semana para oferecer as rotas alternativas está de acordo com a Lei Geral de Proteção de Dados (GDPR), Artigo número 89 item A disponível em https://gdpr-info.eu/art-89-gdpr/</i>
Personal Data	Tipo: Amostra da entidade <i>Personal Data</i> ; Conteúdo: Deve especificar cada dado pessoal utilizado para atender o propósito de uso. As descrições de metadados e metaeventos devem ser os da entidade <i>Personal Data</i> .
Controller	Tipo: Amostra da entidade <i>Actor</i> ; Conteúdo: Deve descrever os atores responsáveis pelo propósito de uso. As descrições dos metadados devem ser os da entidade <i>Actors</i>
Computer-based Decision	Tipo: Texto; Conteúdo: Deve informar ao indivíduo se, para o propósito de uso, existirá alguma decisão tomada exclusivamente por computador (sem a supervisão humana), com base na análise de seus dados pessoais. Caso sim, deve-se apresentar uma sentença que justifique tal iniciativa de uso dos dados pessoais.

Tabela 12: Descrição de Metaeventos para a Entidade *Purpose of use*. Do autor.

Metaevents	
Nome	Descrição dos metaeventos
Start of purpose execution	Tipo: Texto ou Imagens; Conteúdo: Este metaevento deve apresentar informações sobre o momento ou ação que dá início ao uso dos Dados Pessoais é iniciado. Um momento refere-se a uma data, hora ou período específico. Exemplo: <i>Início da execução do propósito de uso: 01/01/2020 às 00:00</i> . Uma ação é uma interação do sistema ou do usuário. Exemplo: <i>Início da execução do propósito de uso: Confirmar o preenchimento do cadastro inicial no aplicativo</i> .
End of purpose execution	Tipo: Texto ou Imagens; Conteúdo: Este metaevento deve apresentar informações sobre o momento ou ação os quais o uso dos Dados Pessoais é finalizado. Um momento refere-se a uma data, hora ou período específico. Exemplo: <i>Término da execução do propósito de uso: 31/12/2020 às 23:59</i> . Uma ação é uma interação do sistema ou do usuário. Exemplo: <i>Término da execução do propósito de uso: Desinstalação do aplicativo</i> .

A próxima seção apresenta os metadados, metaeventos e descrições da entidade *Transfer*.

4.4.4 Metadados, metaeventos e descrições da entidade *Transfer*

Por ser considerada uma das maiores preocupações dos indivíduos em relação ao uso dos dados pessoais, a distribuição para terceiros por meio de Transferência, Compartilhamento ou Divulgação levaram à necessidade de uma atenção especial e uma análise mais rigorosa dos elementos envolvidos nesse contexto a fim de criar os metadados e metaeventos.

Os indivíduos que participaram da análise dos requisitos, e as regulamentações, apresentaram preocupações em relação à distribuição, o que deu à entidade *Transfer* destaque nessa pesquisa. Por parte dos participantes, eles cogitaram a necessidade de saber para quem os dados eram compartilhados, por que e quando foi autorizada tal ação. Já as regulamentações apresentam preocupações com a necessidade de informações sobre as leis que amparam a divulgação dos dados e a preocupação em garantir que o destinatário responda a alguma regulamentação de uso de dados pessoais a fim de que o indivíduo seja protegido no destinatário de sua informação.

Assim, definiu-se como características da entidade *Transfer* um conjunto de informações

que representavam as preocupações dos indivíduos quanto ao destino e a legalidade da transferência, por exemplo: o perfil da divulgação dos dados, a justificativa, garantias legais na origem e destino, e informações de contato do destinatário e do órgão de proteção de dados do mesmo.

Um metadado criado para essa entidade foi chamado de *Type* e faz referência ao tipo de distribuição de dados feita. A necessidade desse metadado foi constatada por identificar formas distintas de uso de dados por terceiros. O metadado *Type* apresenta qual tipo de ação de distribuição de dados é realizada. A seleção das palavras e seus significados foi feita após analisar nas bibliografias, nas PPS e em aplicações de software como os Dados Pessoais poderiam ser distribuídos para vários controladores em parceria. Após identificadas as ações, foi realizada uma pesquisa em dicionários a fim de encontrar os termos mais apropriados para representar as ações.

Esses metadados podem ajudar o usuário a identificar se a distribuição de seus dados é benéfica, uma vez que a parceria com outros controladores pode ser feita para melhorar ou complementar uma informação e assim produzir um resultado ainda melhor para a pessoa. Entretanto, a distribuição pode ser feita para ações que não são de interesse do indivíduo e assim a privacidade, segurança e liberdade ficam comprometidas.

Foi determinada também, como característica da distribuição, quais dados pessoais são distribuídos. Essa informação está relacionada à entidade *Personal Data* na qual deve existir uma instância para cada dado distribuído. A Transparência sobre a distribuição dos dados não seria completa, tão pouco coerente e relevante, se a mesma não relacionasse os dados pessoais envolvidos na ação.

Como metaeventos, decidiu-se pelas informações do momento da permissão do compartilhamento, que pode ser diferente da permissão de uso/coleta dos dados pessoais e de informações de quais ações atuavam como um gatilho para acionar a divulgação dos dados. Os gatilhos de distribuição tem por objetivo auxiliar o indivíduo em entender quais ações feitas em um aplicativo ou *website* pode desencadear um envio de dados para terceiros e caso seja de interesse (e possível) evitar tais ações.

Com a construção de uma estrutura que considera destinatário, os dados distribuídos, a legalidade no envio e no recebimento pelo destinatário, informações do consentimento para distribuir e os gatilhos que efetuam o envio dos dados, acredita-se que a entidade *Transfer* seja capaz de proporcionar informações que garantam uma compreensão segura do fluxo dos dados pelos diversos e possíveis controladores uma vez que a quebra de qualquer elemento envolvido na estrutura pode desencadear a necessidade de intervenção

do usuário.

Também deve-se considerar que, com o conhecimento do destinatário de seus dados pessoais, o indivíduo pode compreender eventos como contato de empresas com as quais o indivíduo nunca se relacionou, conhecimento de perfil do indivíduo por empresas diversas e o uso em ações de vendas ou direcionamento de conteúdo comumente feitas por aplicações de *e-commerce* e redes sociais.

No que tange à descrição dos metadados (Tabela 13) e metaeventos (Tabela 14), essa entidade também priorizou um perfil voltado para textos. Embora, ao analisar as descrições seja possível visualizar eventuais técnicas de *Infovis* para implementar, decidiu-se pela descrição em forma de textos devido ao fato de que adotar técnicas de *Infovis* eficazes para a entidade *Transfer* poderia ser dependente da habilidade e experiência de um *designer*. O uso dos textos para essa Transparência já pode ser adotada de forma simplista mesmo que um desenvolvedor não tenha ampla experiência em aspectos visuais.

Tabela 13: Descrição de Metadados para a Entidade *Transfer*. Do autor.

Metadata	
Nome	Descrição dos metadados
Título	Tipo: Texto; Conteúdo: Título da distribuição
Type	Tipo: Texto e/ou Ícones; Conteúdo: Especifica que tipo de distribuição é feita. As opções são (1) Transfer: Os dados são transferidos para terceiros. O terceiro pode usar os dados para finalidades diferentes, mesmo que não estejam relacionadas ao propósito de uso apresentado pelo controlador que coletou os dados; (2) Sharing: os dados são compartilhados para um destinatário que trabalhará com o controlador para melhorar, suplementar, fornecer mais dados que, de outra forma, estejam relacionados ao propósito de uso apresentado pelo controlador que coletou os dados; (3) Disclose: O controlador disponibiliza os dados para acesso público àqueles que assim o desejarem.
Justification	Tipo: Texto; Conteúdo: Sentença que descreve o motivo da distribuição dos Dados Pessoais. A frase pode seguir a seguinte estrutura: Dados pessoais {nome do dado pessoal} são/é {transferido/compartilhado/divulgado} devido a {motivo da distribuição}. Por exemplo: <i>Dado pessoal: Localização é divulgado para a polícia e outras autoridades (se necessário) para apoiar qualquer tipo de investigação criminal.</i> A sentença pode fazer referência a todos os dados, desde que a ação seja realizada dessa maneira. Por exemplo: <i>A transferência de dados pessoais citada aqui é devido ao compartilhamento de recursos de computação.</i>
Legal basis	Tipo: Texto ou Infográfico; Conteúdo: Descrição da lei/regulamentação que garante a transferência, compartilhamento ou divulgação dos dados pessoais, pode ser estruturada da seguinte forma: (1) - Nome da lei/regulamentação; (2) - Número do artigo ou seção; e (3) - Nome/número de item específico. Por Exemplo: <i>A divulgação está de acordo com o Regulamento Geral de Proteção de Dados (GDPR), número do artigo 48 disponível em https://gdpr-info.eu/art-48-gdpr/.</i>
Recipient	Tipo: Amostra da entidade <i>Actor</i> ; Conteúdo: Deve especificar o destinatário dos dados pessoais. As descrições dos metadados devem ser os da entidade <i>Actors</i> .
Recipient regulation	Tipo: Texto ou Infográfico; Conteúdo: Deve apresentar a lei/regulamentação de uso dos dados pessoais à qual o(s) destinatário(s) são submetidos.
Personal Data	Tipo: Amostra da entidade <i>Personal Data</i> ; Conteúdo: Deve apresentar quais dados pessoais são distribuídos. As descrições dos metadados e metaeventos devem ser os da entidade <i>Personal Data</i> .

Tabela 14: Descrição de Metaeventos para a Entidade *Transfer*. Do autor.

Metaevents	
Nome	Descrição dos metaevents
Permission	Tipo: Imagens; Conteúdo: Este evento deve informar sobre o momento ou a ação à qual o indivíduo concedeu autorização para a distribuição de seus dados. Para apresentar a informação ao usuário, o sistema deve usar um conjunto de imagens apresentando a interface utilizada para conceder autorização de distribuição dos dados. O uso de animações, cores, destaque podem ser utilizados para auxiliar na identificação da informação. As imagens podem ser complementadas por frases como <i>Veja quando/como/onde você autorizou a transferência de seus dados pessoais</i> .
Ocurrence	Tipo: Texto ou Imagens ou Infográficos; Conteúdo: Esse metaevento deve fornecer informações sobre a estratégia de distribuição dos dados pessoais. Os dados pessoais são transferidos, compartilhados ou divulgados por um período de tempo ou com base em alguma ação do indivíduo ou da aplicação. Sugere-se a estrutura: <i>O dado pessoal {descrição do dado} é/será { transferido/compartilhado/divulgado} { estratégia de compartilhamento}</i> . Por exemplo: <i>Seus dados pessoais serão transferidos toda vez que você utilizar o aplicativo</i> ou <i>Seus dados pessoais serão compartilhados a cada trinta minutos</i> .

A entidade *Transfer* propôs informações que permitem ao indivíduo identificar o destino de seus Dados Pessoais, bem como a legalidade das ações. Com esse conhecimento a pessoa pode tomar decisões sobre como continuar (ou não) usando o serviço ou aplicativo e, consequentemente, distribuir seus dados a terceiros, além de fornecer subsídios para uma eventual denúncia.

A próxima seção apresenta os metadados e as descrições para a entidade *Agency*.

4.4.5 Metadados e descrições da entidade *Agency*

A entidade *Agency* é a menor e com metadados com descrições mais abrangentes, porém seu conteúdo não é menos relevante que o de outras entidades. Essa entidade visa prover informações de como o indivíduo deve agir para exercer seu direitos. Os direitos dos indivíduos que são garantidos nas regulamentações são diversos e já citados anteriormente. A responsabilidade de disponibilizar meios para o indivíduo exercer seus direitos é do controlador e o mesmo pode fazer da maneira que achar necessário, podendo inclusive, cobrar por ele.

Por essa razão, decidiu-se que os metadados dessa entidade precisam orientar os in-

divíduos a exercerem seus direitos. Portanto, essa entidade precisou somente de metadados para representar as características do direito do indivíduo. Não foram incluídos metaeventos, pois, assim como na entidade *Actors*, as ações relacionadas à existência da informação na entidade *Agency* não foram consideradas relevantes para a compreensão do uso dos dados pessoais.

Os metadados limitaram-se a: (1) título da ação; (2) descrição de como o indivíduo deve agir, um metadado aberto que proporciona flexibilidade ao controlador para informar a maneira mais viável de como o indivíduo deve agir. O uso de uma abordagem de tutorial é recomendada na GDPR ; e (3) informação do destinatário que se trata de uma instância da entidade *Actor* que deve informar sobre quem receberá a requisição, denúncia ou ação do indivíduo.

Os metadados da entidade *Agency* e suas descrições são mostrados na Tabela 15.

Tabela 15: Metadados e Descrições para a entidade *Agency*. Do autor.

Nome	Descrição dos metadados
Agency's Name	Tipo: Texto; Conteúdo: Frase simples que descreve o nome da ação/negociação, por exemplo: (1) Solicite uma cópia dos dados pessoais; ou (2) Relate o uso incorreto dos seus dados.
Description	Tipo: Livre; Conteúdo: Esse metadado deve fornecer informações sobre as ações necessárias para o indivíduo agir ou negociar o uso de seus dados. O aplicativo pode exibir uma lista de eventos, ou conjunto de imagens, ou um tutorial ou um endereço de e-mail o indivíduo possa utilizar. Não é obrigação da ferramenta de Transparência fornecer o recurso ao usuário, mas informar como acessar o recurso fornecido pelo site do Controlador ou do Escritório de Proteção
Recipient	Tipo: Amostra da entidade <i>Actor</i> ; Conteúdo: Deve descrever o destinatário das requisições dos participantes. As descrições dos metadados devem ser os da entidade <i>Actors</i> .

A próxima seção apresenta as considerações finais da construção do TR-Model.

4.5 Considerações finais da construção do TR-Model

Esse capítulo apresentou as atividades conduzidas para construir o TR-Model que contemplou a definição do modelo de domínio, a criação das entidades e a definição e descrição dos metadados e metaeventos.

A construção do TR-Model utilizou de técnicas de levantamento de requisitos a fim de identificar as características de Transparência de Dados Pessoais, as necessidades dos

indivíduos, e as exigências de regulamentações para esse assunto; de técnicas para construção de Perfil de Aplicação de Metadados propostas pelo *Dublin Core Metadata Initiative* a fim de produzir um modelo consistente, aplicável e não ambíguo; e de conceitos de qualidade de informação e *readability* para descrever os metadados para dar suporte a produção de Transparência com foco no indivíduo.

O próximo capítulo apresenta a validação do TR-Model.

5 VALIDAÇÃO

Esta seção apresenta as ações realizadas para validar a eficácia do TR-Model em dar suporte à produção de informações de Transparência de Dados Pessoais. Por eficácia considerou-se: a capacidade de atender as requisições de Transparência da GDPR e da LGPD; e a capacidade de entregar Transparência para apoiar a análise dos indivíduos.

Assim, decidiu-se por realizar a validação utilizando o seguinte conjunto de técnicas:

1. **Cobertura das informações:** verificar se as informações propostas atendiam as instruções elencadas nas regulamentações;
2. **User Evaluation:** essa etapa da validação com os indivíduos foi realizada com combinação de testes controlados com resolução de questionários. Os dados coletados foram analisados em duas frentes: (1) mensuração do quanto o TR-Model atendia as expectativas dos indivíduos em relação às informações de Transparência; e (2) qualidade dos elementos de IHC e das dimensões de IQ nas disponibilização das informações para os indivíduos.

As próximas seções apresentam as validações realizadas, começando pela análise da cobertura do TR-Model em relação à GDPR.

5.1 Análise da Cobertura do TR-Model em relação GDPR

Nesta etapa foi realizada uma comparação entre as entidades, metadados e metaeventos do TR-Model com as requisições de Transparência elencadas na GPDR e na LGPD.

A análise da GDPR é apresentada na Tabela 16 na qual considerou-se os Artigos 13 e 14 da regulamentação.. Foi realizada uma compilação das diretrizes de Transparência da GDPR e apontado qual metadado, metaevento ou entidade no TR-Model atendeu (ou não) cada item. Decidiu-se fazer a compilação uma vez que artigos diferentes apresentavam

exatamente a mesma diretriz. Também foram ajustados alguns textos os quais tiveram algumas sentenças suprimidas por não ter relação com a Transparência.

Tabela 16: Cobertura do TR-Model em relação à GDPR. Do autor.

Artigo, Seção e Alínea da GDPR	Entidade, metadado ou metaevento do TR-Model
Art. 13 and 14/1/a.	Entidade <i>Actor</i>
Art. 13 and 14/1/b	Entidade <i>Actor</i>
Art. 13 and 14/1/c	Entidade <i>Purpose of use</i> nos metadados <i>Description</i> e <i>Legal Basis</i>
Art. 14/1/d	Entidade <i>Personal Data</i> com os metadados <i>Description</i> and <i>Granularity</i>
Art. 13/1/d and Art. 14/2/b	A combinação de informações das entidades <i>Purpose of use</i> e <i>Personal Data</i> podem proporcionar as informações de Transparência
Art. 13/1/e and Art. 14/1/d	Metadado <i>Controller</i> na entidade <i>Purpose of use</i> e o metadado <i>Recipient</i> na entidade <i>Transfer</i> , pois fazem referência aos dados dos atores
Art. 13 and 14/1/f	Metadado <i>Recipient Regulation</i> na entidade <i>Transfer</i>
Art. 13 and 14/2/a	Metaevento <i>Utilization</i> na entidade <i>Personal Data</i>
Art. 13/2/b and Art. 14/2/c	Entidade <i>Agency</i>
Art. 13/2/e	Metadado <i>Mandatory Collection</i> na entidade <i>Personal Data</i>
Art. 13/2/f and Art. 14/2/g	Metadado <i>Computer-based decision</i> na entidade <i>Purpose of use</i>
Art. 14/2/f	Metadado <i>Resource</i> na entidade <i>Personal Data</i>

Já para a LGPD foram consideradas as especificações propostas por Law (2020) e já discutidas na subseção 2.3.2 na página 32. A análise da cobertura é mostrada na Tabela 17.

Com base nas informações da Tabela 16 é possível concluir que o TR-Model atende as requisições da GDPR em relação a *quais informações disponibilizar para o indivíduo*. De fato, a GDPR foi um dos principais artefatos para compreender o domínio de Transparência e fornecer insumo para os requisitos. Assim, pode-se afirmar que tal adequação era esperada e deveria ser atingida.

Em relação à LGPD, o TR-Model contemplou quase todas as necessidades de informação, mas não atendeu, de forma direta, a necessidade de informar a possibilidade de recusar a autorização de uso e ser informado sobre as consequências. Não atender de

Tabela 17: Cobertura do TR-Model em relação à GDPR. Do autor.

Requisição da LGPD	Entidade, metadado ou metaevento do TR-Model
Finalidade específica do tratamento de dados	Metadado <i>Description</i> da entidade <i>Purpose of use</i>
Forma e duração do tratamento de dados	Metaeventos <i>Start of purpose execution</i> e <i>End of purpose execution</i> da entidade <i>Purpose of use</i>
Dados de contato do controlador	Metadados da entidade <i>Actors</i>
Informações sobre compartilhamento de dados e contato do destinatário	Metadados e metaeventos da entidade <i>Transfer</i>
Responsabilidade dos envolvidos no tratamento dos dados	Metadado <i>Controller</i> na entidade <i>Purpose of use</i> ; e metadado <i>Recipient</i> na entidade <i>Transfer</i>
Direitos dos titulares dos dados	Metadados da entidade <i>Agency</i>
Possibilidade de recusar a autorização do uso dos dados e a consequência dessa escolha	Não atendido

forma direta faz referência ao fato de não ter um metadado ou metaevento no domínio específico para essa necessidade. Entretanto, o TR-Model poderia atender de forma indireta com o uso da entidade *Agency*, mas ficaria a cargo dos desenvolvedores programar suas aplicações para apresentar, em uma interface gráfica, informação de forma a prevenir o indivíduo sobre a ação. Já a informação sobre a consequência da não autorização do uso dos dados não foi contemplada na primeira versão do TR-Model e deverá ser incluída em futuras versões.

A próxima seção apresenta a validação do TR-Model com indivíduos.

5.2 Validação do TR-Model com indivíduos

A validação com indivíduos foi realizada em duas etapas: (1) Pré Teste; (2) Simulação de Transparência com cenários. Na etapa de pré testes foram coletados e analisados dados sobre conhecimento prévio dos participantes em relação ao uso dos dados pessoais; e sobre suas expectativas de Transparência a fim de verificar quanto o TR-Model atendeu as expectativas.

Na simulação de Transparência com cenários, a validação buscou verificar se as diretrizes do TR-Model foram eficazes para produzir Transparência de Dados Pessoais para uso dos indivíduos. A validação foi realizada com testes de usuários combinados com resolução de questionários. Os participantes realizaram a análise da Transparência de 05

(cinco) cenários diferentes e, para cada cenário avaliado, os mesmos responderam a um questionário de avaliação.

5.2.1 Atividades Pré Teste

Na atividade pré teste, os participantes responderam um questionário sobre conhecimentos prévios em relação ao uso de seus dados pessoais e em relação à sua preocupação com essa situação.

Quanto ao uso dos dados pessoais por aplicativos, o resultado foi quase unânime, ou seja, 97% sabiam a respeito e somente 3% (4 participantes) desconheciam. Para identificação dos participantes no texto, serão tratados os que tinham conhecimento prévio como **Perfil A** e os que não conheciam o uso dos dados como **Perfil B**.

Os participantes do Perfil B também mostraram pouca ou nenhuma preocupação com mecanismos que pudessem fornecer algum tipo de Transparência. Esses participante responderam que nunca fazem a leitura das PPS e 75% deles nunca se preocuparam com o uso dos dados pessoais; e 25% apresentaram alguma preocupação. Em razão do pequeno número de respondentes do grupo com Perfil B não serão apresentadas discussões aprofundadas sobre esses resultados, mas acredita-se que o desconhecimento acerca do assunto justifica a tranquilidade e desapego deste grupo com seus dados pessoais.

Em relação aos participantes do Perfil A, os resultados mostraram sutil mudança uma vez que 64% (75 participantes) também afirmaram nunca terem lido as PPS e 36% (42 participantes) fizeram algumas leituras. Semelhante aos participantes que não conheciam sobre o uso dos dados pessoais, ou seja, nenhum participante indicou que faz a leitura das PPS com frequência.

Os participantes do Perfil A, quando questionados sobre a preocupação com o uso dos dados pessoais, responderam conforme mostrado na Figura 7.

Outro questionamento feito aos participantes foi a frequência com que procuram informações sobre o uso de seus dados pessoais. Esse questionamento será discutido somente considerando os participantes do Perfil A, pois os demais, naturalmente responderam que nunca procuram tais informações. As respostas para essa questão foram: (1) Nunca: 34,19%; (2) Raramente: 54,70%; Frequentemente: 10,26%; e Sempre: 0,85%.

Para os participantes que responderam que *Nunca* ou *Raramente* buscam informações, foi solicitado que descrevessem o motivo pelo qual informaram tais respostas. A questão aberta foi respondida por 104 participantes os quais apresentaram respostas que contem-

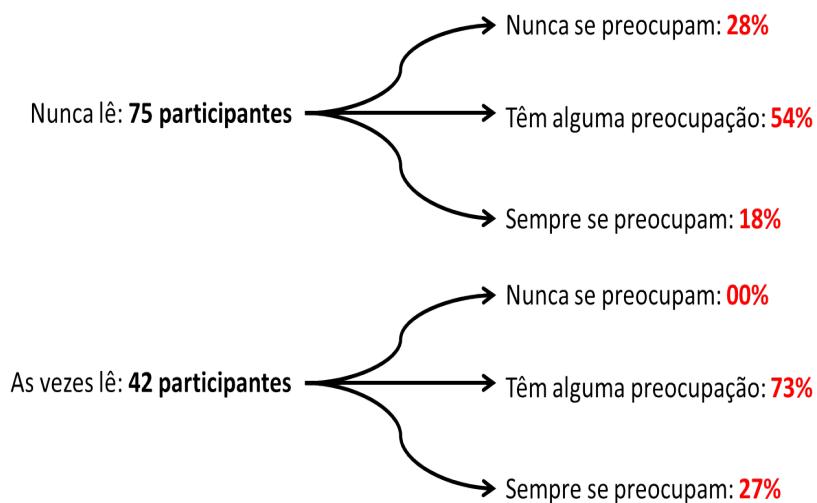


Figura 7: Relação de leituras das PPS com a preocupação com o uso dos dados pessoais - Perfil A. Do autor.

plavam elementos como:

- Falta de interesse pela informação;
- Falta de tempo, uma vez que precisa ler muita informação;
- Dificuldade em identificar e/ou encontrar a informação;
- Falta de opção, uma vez que ao negar acesso aos dados o indivíduo não pode consumir o serviço;
- Confiança na aplicação;
- Pouco conhecimento sobre o assunto;
- Textos longos, jurídicos e subjetivos.

Para aqueles que responderam que buscam informações sobre o uso dos dados pessoais, foi solicitado que descrevessem os pontos positivos e negativos das informações disponibilizadas. Treze participantes descreveram os pontos apresentados na Tabela 18.

As justificativas da não-leitura, assim como o apontamento dos pontos negativos, corroboram situações já discutidas onde a falta de uma informação apresentada de forma simples, clara e comprehensível, assim como o uso de meios pouco acessíveis, faz com que a Transparéncia não se efetive. Essas características acabam por promover um desinteresse dos indivíduos em verificar o uso dos dados pessoais, o que deixa um caminho livre para

Tabela 18: Pontos positivos e negativos das informações sobre o uso dos dados pessoais.
Do autor.

Positivo	Negativo
Direcionamento de conteúdo	Inexatidão da informação
Existência de alguma informação	Excesso de conteúdo Ambiguidade Complexidade de termos e palavras

empresas promoverem uma abordagem *black box* onde a pessoa aceita os termos de uso dos dados ou não utiliza a aplicação ou *website*.

Quando as empresas disponibilizam as informações, as mesmas têm conteúdo complexo, as informações estão enviesadas dando margem de interpretação a favor da empresa e dificultando eventuais ações dos indivíduos, além de utilizar a estratégia de “cansar” o indivíduo durante a leitura de forma que o mesmo ignore detalhes importantes.

Desta forma, mesmo com um perfil de participantes que têm preocupação com o uso dos dados pessoais e procurem informações a respeito, é notável que problemas com a qualidade de informação afetam o conteúdo da Transparência e acabam por dificultar o acesso por parte dos indivíduos. Embora o uso dos dados possa ser positivo e benéfico ao indivíduo, as dificuldades relatadas podem abrir caminhos para a utilização inadequada.

Por fim, a partir dos dados fornecidos pelos participantes do presente estudo, pode-se concluir, os mesmos não são desinteressados ou omissos em relação ao uso dos dados pessoais, mas a forma como as informações são dispostas acaba por direcionar a um comportamento no qual ignoram eventuais ações de análise devido a sua complexidade. É conhecido que as pessoas normalmente reagem melhor a análises de conteúdos rápidos, objetivos, simplificados e com linguagem simples (CYBIS; HOLTS; FAUST, 2015).

As dificuldades apresentadas corroboram a necessidade de uma abordagem que entregue informações de Transparência com foco no indivíduo e que as informações apresentadas sejam bem definidas, sem viés comercial e que permitam ao indivíduo analisar o uso de seus dados de forma simplificada e eficaz para o mesmo.

Nessa etapa também buscou-se verificar se o TR-Model atendia às expectativas dos participantes em relação às informações sobre o uso dos dados pessoais. No questionário pré-teste havia a seguinte pergunta: *Se você pudesse escolher alguma informação sobre o uso de seus dados pessoais para ter acesso, qual ou quais seriam?* Não foi utilizado o termo Transparência na pergunta para evitar que o participante se confundisse ao responder. A pergunta era aberta (textual) e obrigatória para todos os participantes.

Foram registradas 119 respostas consideradas válidas; estas passaram por uma análise a fim classificá-las. A classificação foi feita porque percebeu-se que uma única resposta poderia ter várias expectativas de informação e porque foram identificadas respostas inválidas por má interpretação ou por uso incorreto da alternativa inserindo conteúdo não relacionado com a questão.

A análise e classificação das expectativas gerou a um grupo de 13 categorias e seus resultados são mostrados na Figura 8.

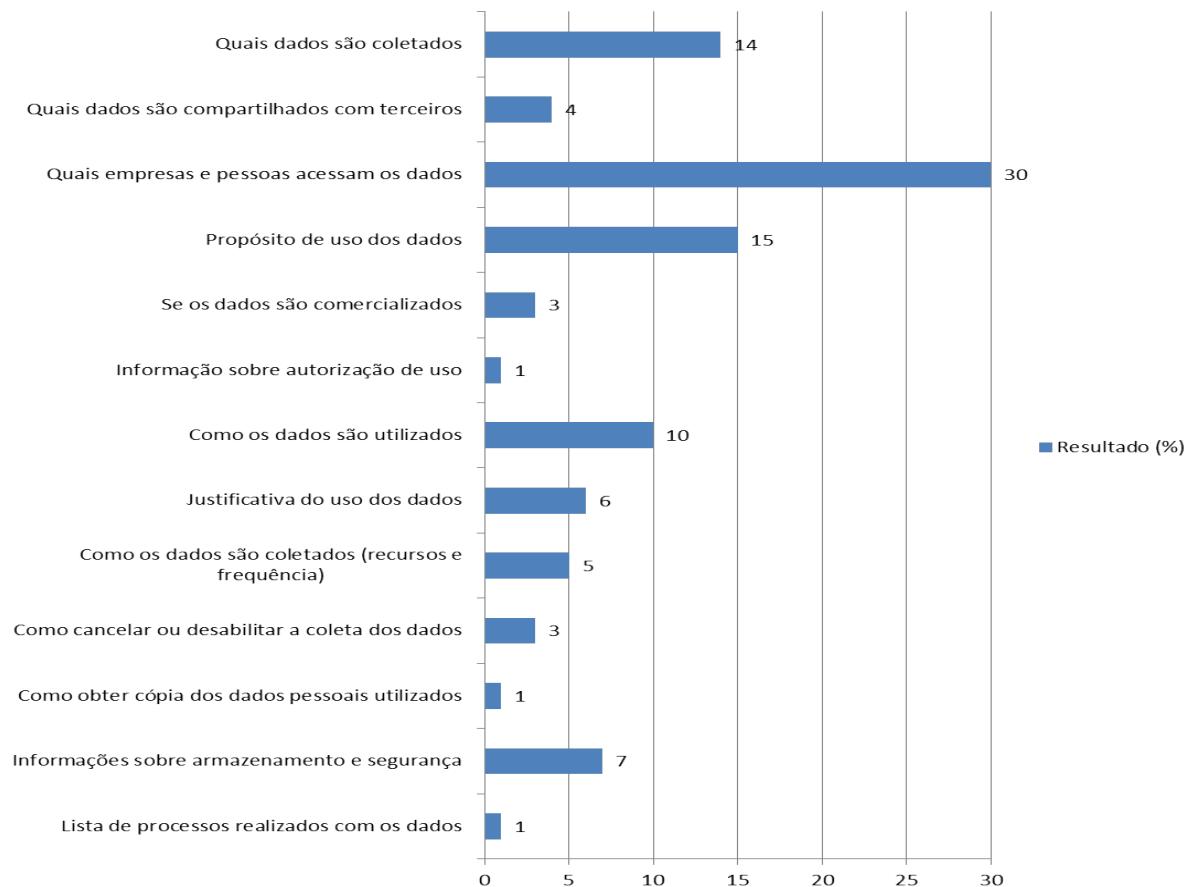


Figura 8: Expectativas dos participantes em relação à Transparência. Do autor.

Os resultados permitiram concluir que a maior preocupação dos participantes está relacionada a: *quais dados são usados, por quem são acessados e objetivo do uso dos dados*. Foi identificada também certa preocupação com o *como os dados serão utilizados*. Trata-se de uma informação mais técnica e sua Transparência dependeria de abstração dos processos realizados nos dados para uma linguagem comprehensível para os indivíduos.

Destaca-se nos resultados a expectativa por *quais pessoas e empresas terão acesso aos dados*. Essa preocupação está fortemente relacionada com a possibilidade de acesso quase irrestrito aos dados por parte de controladores e empresas terceiras. Foi possível

perceber que os participantes demonstravam certo conhecimento sobre a troca de dados pessoais entre empresas, principalmente quando discutido tal assunto nos *workshops*. A distribuição dos dados sem um controle rigoroso desperta nos participantes preocupações com segurança e privacidade, uma vez que seus dados podem ser acessados e utilizados para finalidades não desejadas pelos mesmos colocando-os em situações de risco.

Expectativas como: recursos utilizados para coletar dados, frequência de coleta, informações de segurança, armazenamento e detalhes acerca do cancelamento e da permissão de uso dos dados foram apresentadas, mas em menor número. A informação sobre segurança e armazenamento podem ser respostas de pessoas com maior familiaridade com computação, pois os mesmos têm um maior conhecimento desses fatores e sabem como problemas de segurança podem comprometer a privacidade dos dados pessoais.

Por fim, assumiu-se que o TR-Model contemplou 12 das expectativas de forma direta, ou seja, as expectativas dos participantes estão contempladas nas entidades, nos metadados ou metaventos e em suas descrições. Não foram contempladas (ou não foram contempladas de forma direta) as seguintes expectativas:

- *Lista de processos realizados nos dados*: com uma única indicação, entendeu-se que o participante esperava informações técnicas sobre o histórico de processamento dos dados pessoais, algo mais próximo do conceito de Proveniência de Dados. Por ser um aspecto que necessita de um modelo de suporte específico, tal informação não foi contemplada nessa primeira versão do TR-Model;
- *Informação sobre armazenamento e segurança*: durante o desenvolvimento do TR-Model chegou-se a cogitar acerca dessas informações, mas as mesmas foram consideradas muito técnicas (computacional) e poderiam não despertar o interesse das pessoas além de aumentar o volume de informação a ser apresentado. Entretanto, a disponibilidade dessa informação será considerada em futuras versões do TR-Model;
- *Justificativa de uso*: essa informação é contemplada na entidade *Transfer*, mas não foi inserida na entidade *Purpose of use*. De forma indireta, a justificativa de uso poderia ser inserida para complementar o metadado *Description* da entidade *Purpose of use*. Essa informação também será inserida em futuras versões.

Portanto, considerando a amostra participante desta validação, os resultados permitem assumir que o TR-Model atende 95% das expectativas das informações esperadas pelos indivíduos. Para informações que apresentaram maior frequência de expectativas,

o TR-Model foi eficaz em contemplar 100% delas em suas entidades, metadados e metaeventos.

Esse resultado deve-se à participação dos indivíduos (usuários) em reuniões, palestras, resoluções de questionários. A participação foi fundamental no desenvolvimento do TR-Model uma vez que o mesmo sempre buscou proporcionar uma *User Experience* que atendesse aos interesses dos indivíduo. Mesmo que nem todos os participantes da etapa de requisitos tenham participado da validação, assume-se que as expectativas de diferentes participantes são similares e que assim o TR-Model pode atender um amplo perfil de usuários.

A próxima subseção apresenta-se a a validação do TR-Model com a utilização de cenários de Transparência.

5.2.2 Validação do TR-Model com cenários

Os cenários foram baseados em aplicativos de monitoramento de atividades físicas tais como: corrida, caminhada e atividades em academias. Decidiu-se utilizar esse molde porque esse tipo de aplicação manipula dados pessoais como localização, peso e condicionamento físico e que reflete, sem desvios, o conceito de uso de Dados Pessoais.

As informações sobre o uso dos dados pessoais foram produzidas seguindo os metadados, metaeventos e as descrições do TR-Model. Destaca-se que os cenários foram desenvolvidos pelo próprio pesquisador, o que pode ter causado certo enviesamento na construção/*design* de alguma informação. Conforme afirmado em capítulos anteriores, a avaliação das informações de Transparência proporcionadas pelo TR-Model para o uso dos indivíduos foi a prioridade dessa validação.

Para a validação foram construídos 5 (cinco) cenários (páginas web) e cada cenário implementou uma ou várias entidades do TR-Model, conforme apresentado a seguir:

Cenário 01: teve como objetivo proporcionar informações sobre o propósito de uso dos dados pessoais e quais dados seriam utilizados, sua descrição era: “*Você está iniciando atividades de corrida, caminhada e treino em academia. Seguindo as orientações de seu Personal Trainer você faz o download de um aplicativo para registrar suas atividades. Antes da instalação você é informado que seus dados pessoais serão coletados e utilizados. Diante dessa informação, você deve acessar a interface de Transparência e verificar qual o propósito do uso dos dados pessoais e quais dados serão utilizados*”.

O Cenário 01 implementou as entidades *Purpose of use*, *Actors* e *Personal Data*. Na Figura 9 é possível identificar os metadados da entidade *Purpose of use* e, na Figura 10, a relação com a entidade *Actor* onde são apresentados os dados do controlador.

The screenshot shows a web browser window with the URL each.usp.br/cond_met_pand/transparencia/purposeone.html. The title bar says "TR-Model - Cenário 01". On the right, there are links for "Lista de Cenários" and "Avaliar este Cenário". The main content area has a blue header "Propósito de Uso". Below it, a section titled "Descrição do Objetivo" contains the text: "Acompanhar seu desempenho nos treinos para poder oferecer produtos e serviços complementares que te auxiliem na busca de suas metas." A button labeled "Controlador: NADO LIVRE ACADEMIA" is present. Another section, "Lei/Regulamentação", states: "Este propósito de uso está de acordo com a Lei Geral de Proteção de Dados, Capítulo II, Seção I Art. 7º itens I e VI, disponível [nesta página](#)". A question "Há alguma decisão tomada exclusivamente por computador?" has the answer "NÃO! O computador é responsável por sugerir o produto, mas a decisão de oferecer o mesmo ao usuário é feita por um vendedor especializado." To the left is a clock icon, and to the right is a crossed-out clock icon. Below these icons, two descriptive paragraphs are shown: "Início da execução deste Propósito de Uso: Este propósito de uso será executado a partir do momento do consentimento de uso dado por você ao aplicativo." and "Término da execução deste Propósito de Uso: Este propósito será executado enquanto o aplicativo estiver sendo utilizado por você ou até você cancelar a autorização de uso de seus dados pessoais."

Figura 9: Interface da entidade *Purpose of use*. Do autor.

The screenshot shows the same web browser window as Figure 9. An arrow points from the "Controlador: NADO LIVRE ACADEMIA" button in the main content area to a detailed view of the controller's data on the right. The detailed view is titled "Dados do Controlador: Nado Livre Academia". It lists the following information: Endereço: Rua Frei Rafael Proner, Nº: 2263; Cidade: Bandeirantes; Estado: Paraná; País: Brasil; Localização no Google Maps: [VER NO MAPA](#); and Telefone: +55 43-3145-1155.

Figura 10: Interface com informações do controlador do propósito de uso (Relação das entidades *Purpose of use* e *Actors*). Do autor.

Na Figura 11 são mostradas informações que implementadas com base nos metadados da entidade *Personal Data*, ou seja, as características dos dados pessoais utilizados para o propósito.

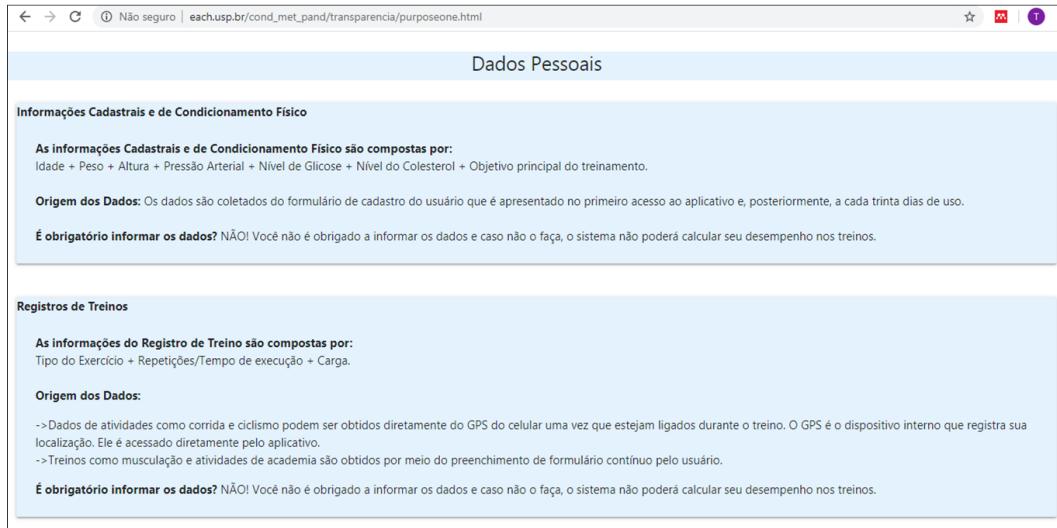


Figura 11: Interface dos metadados da entidade *Personal Data*. Do autor.

Já na Figura 12 é apresentada a implementação dos metaeventos da entidade *Personal Data* no Cenário 01.

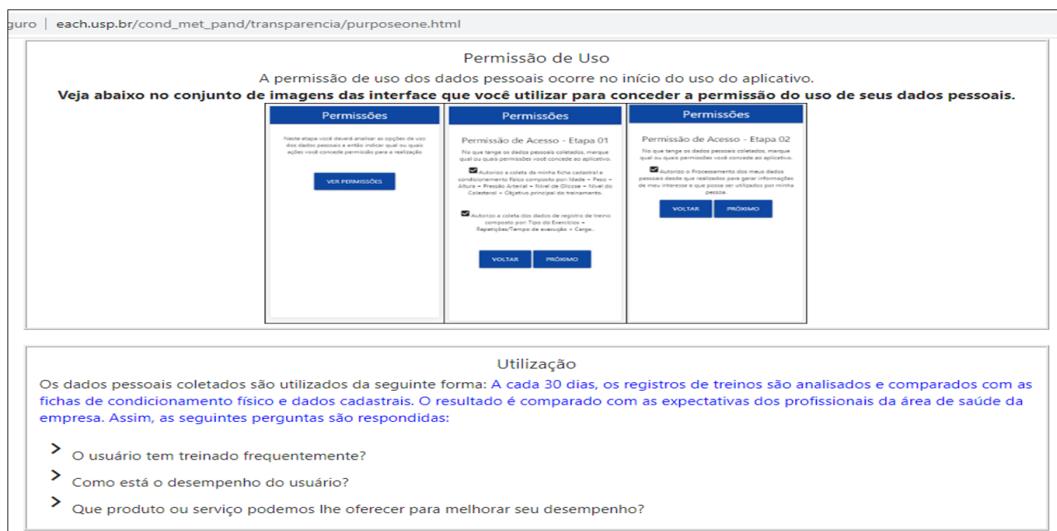


Figura 12: Interface dos metaeventos na entidade *Personal data*. Do autor

Cenário 02: implementou um cenário no qual os participantes analisaram a Transparência sobre a transferência ou compartilhamento de seus dados pessoais entre sistemas. Assim, foi implementada a entidade *Transfer*, seus metadados e metaeventos. A descrição do Cenário 02 foi: “*Você já sabe quais de seus dados serão coletados e utilizados. Entretanto, antes de começar a usar o aplicativo surge uma dúvida sobre o compartilhamento de seus dados. Assim você decide: verificar se seus dados serão compartilhados, quem vai ter acesso aos dados e como você deve fazer para denunciar eventual ação incorreta*”.

Na Figura 13 é mostrado um protótipo no qual o indivíduo visualizava os metadados *Title* e *Justification* da entidade *Transfer*, além de um botão no qual o indivíduo poderia acessar uma segunda interface contendo mais metadados da entidade *Transfer*, conforme mostrado na Figura 14.

Figura 13: Interface implementando metadados da *Transfer*. Do autor.

Figura 14: Interface implementado demais metadados da entidade *Transfer*. Do autor.

A implementação dos metaeventos da entidade *Transfer* é mostrada na Figura 15.

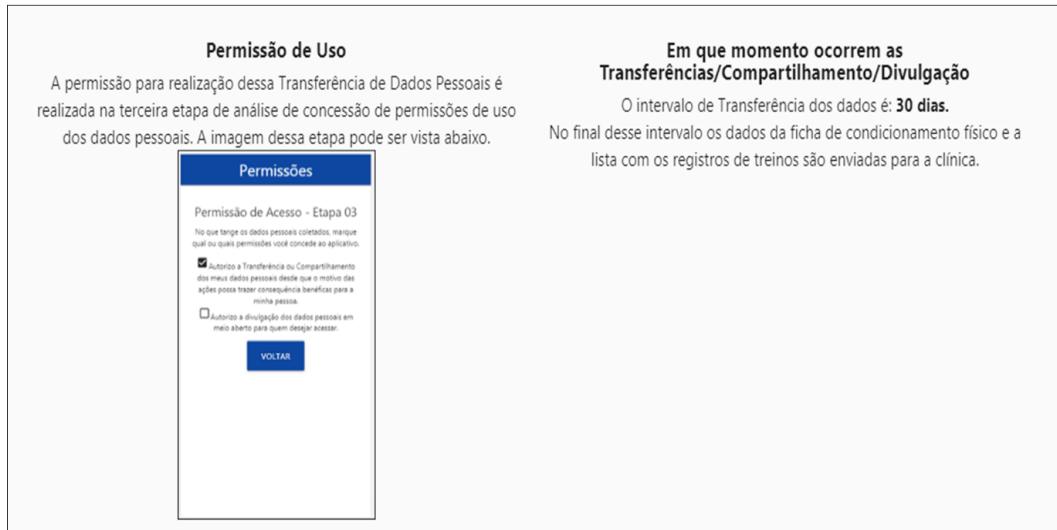


Figura 15: Interface com os metaeventos da entidade *Transfer*. Do autor.

Cenário 03: simulou uma situação em que o usuário era informado, no momento da coleta dos dados, sobre essa ação. A descrição desse cenário foi: “*Você está registrando uma atividade do seu treinamento. Durante o registro, uma mensagem informa que seus dados pessoais serão utilizados. Você então decide: verificar para qual propósito o dado pessoal lançado será utilizado e se está de acordo com o autorizado anteriormente*”.

O Cenário 03 implementou metadados da entidade *Purpose of use*. O cenário foi utilizado para verificar se o propósito de uso dos dados pessoais no momento da coleta dos dados estava de acordo com a autorização concedida na íntegra do uso da aplicação. Na Figura 16 é mostrado um protótipo em que as informações de uso dos dados eram comparados.

Item Analisado	Uso do dado pessoal autorizado por você	Uso do dado pessoal realização pelo aplicativo atualmente	Situação
Propósito de Uso	Oferecer produtos e serviços personalizados de acordo com suas expectativas de treino e seu desempenho.	Oferecer produtos e serviços personalizados de acordo com suas expectativas de treino e seu desempenho.	<input checked="" type="checkbox"/>
Este propósito de uso está de acordo com a Lei Geral de Proteção de Dados, Capítulo II, Seção I Art. 7º itens I e VI, disponível nesta página .	Este propósito de uso está de acordo com a Lei Geral de Proteção de Dados, Capítulo II, Seção I Art. 7º itens I e VI, disponível nesta página .	<input checked="" type="checkbox"/>	
Dados Pessoais	Informações Cadastrais e de Condicionamento Físico Registros de Treinos	Informações Cadastrais e de Condicionamento Físico Registros de Treinos Lista de Contatos da Agenda	<input type="checkbox"/>

Figura 16: Interface de verificação dos propósitos de uso. Do autor.

Cenário 04: implementou o metaevento *Permission of use* da entidade *Personal Data*. As informações eram apresentadas em um conjunto de imagens e descreviam (simulavam) situações em que foram atribuídas permissões de uso de dados pessoais pelos indivíduos. Sua descrição foi: “*Uma clínica médica entra em contato oferecendo alguns serviços. Você nota que os serviços oferecidos têm relação com suas atividades físicas e acredita que eles adquiriram as informações sobre você. Assim você decide: verificar se você deu permissão para compartilhar seus dados pessoais para determinada finalidade e em qual momento essa permissão ocorreu. Além disso, você verifica em que momento seus dados pessoais foram coletados*”. O protótipo mostrado na Figura 17 foi utilizado para o Cenário 04.

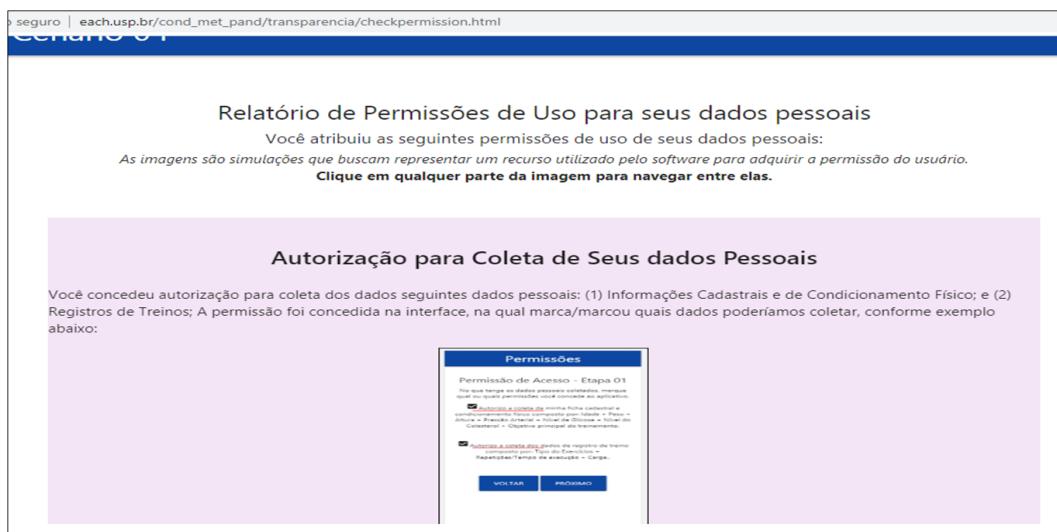


Figura 17: Interface com informações do metaevento *Permission of use* da entidade *Personal Data*. Do autor.

O **Cenário 05** implementou a entidade *Actors*, pois simulou uma situação em que o indivíduo precisava entrar em contato com o controlador e assim necessitava de dados de contato de agentes envolvidos no uso dos dados pessoais. Sua descrição foi: “*Você identificou uma utilização inadequada de seus dados. Então decide entrar em contato com o controlador responsável pelo uso dos dados e ao mesmo tempo informar a agência de proteção de dados pessoais. Para isso você deve buscar informações de como entrar em contato com o controlador e com a agência de proteção de dados*”.

As Figuras 18 e 19 apresentam as interfaces do Cenário 05.

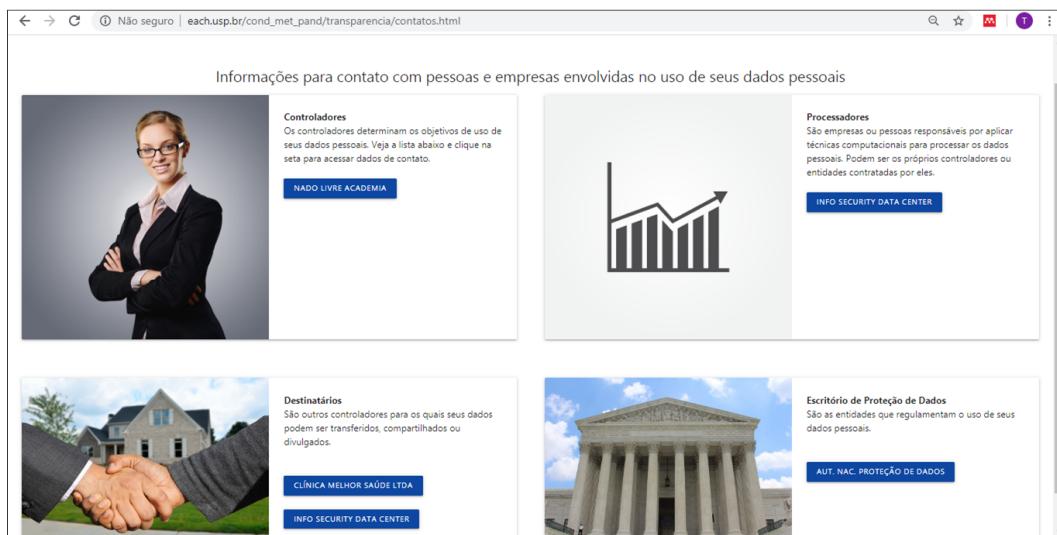


Figura 18: Interface com lista de atores envolvidos no uso do dado pessoal. Do autor.

Dados do Processador: Info Security Data Center	
Endereço:	R. Euclides Monteiro
Nº:	342
Cidade:	Ibaiti
Estado:	Paraná
Pais:	Brasil
Localização no Google Maps:	VER NO MAPA
Telefone:	55 43 3546-3166
E-mail:	atendimento@joinsoft.com.br
Procurador:	Thiago Adriano Coleti
Escritório de Proteção de Dados:	Autoridade Nacional de Proteção de Dados

Figura 19: Interface com dados de contato do ator conforme metadados da entidade *Actors*. Do autor.

Para cada cenário avaliado, foi proposto um questionário para capturar dados sobre a opinião do participante em relação à qualidade da Transparência de cada cenário. A validação com usuários foi feita com base na abordagem subjetiva descrita por Pipino, Lee e Wang (2002) e Qualitativa, apresentada por Budiu (2017) a fim de identificar como seria a experiência do indivíduo com a Transparência e também identificar pontos que poderiam ser melhorados no TR-Model.

O mesmo questionário utilizado para todos os cenários avaliados era composto pelas perguntas mostradas na Tabela 19 na página 98. Foram utilizadas duas abordagens de questões: (1) questões independentes (aqueles disponíveis independente das respostas de

outras questões) direcionavam a avaliação para questões de *Readability*, HCI e IQ discutidas nas seções anteriores; (2) questões dependentes: dependiam da resposta de outras questões, por exemplo: A pergunta *Se na questão anterior você respondeu que não conseguiu entender a Transparência ou entendeu somente parte das informações* era exibida caso o resposta da questão *Em relação à Facilidade de Entendimento e Interpretação da Transparência do cenário avaliado, qual a sua opinião?* apontasse para as opções Ruim ou Razoável.

As questões dependentes eram dissertativas a fim de obter sugestões ou justificativa dos indivíduos em relação a sua resposta para a questão anterior. O conteúdo dessas questões foram avaliados a fim de proporcionar insumos para correção ou melhorias no TR-Model.

Tabela 19: Questões para a avaliação dos cenários. Do autor.

Question	Options
Q1 - Em uma escala de 1 - Muito Ruim a 5 - Muito bom. Qual sua opinião em relação à facilidade para encontrar e identificar uma informação sobre o uso de seus dados pessoais?	Escala
Q2 - Em relação à quantidade de informação apresentada na Transparência do cenário avaliado, sua opinião é:	(a) Excessiva; (b) Pouca; ou (c) Adequada
Q2a - Se você achou a quantidade de informações sobre o Propósito de Uso POUCA ou EXCESSIVA. Por favor, descreva qual, ou quais informações poderíamos acrescentar (se você achou que há POUCAS informações) ou quais informações estão em excesso (se você achou que há MUITAS informações).	<i>Questão dissertativa dependente da questão anterior</i>
Q3 - Em relação à Facilidade de Entendimento e Interpretação da Transparência do cenário avaliado, qual a sua opinião?	(a) Ruim, (b) Razoável e (c) Adequada
Q3a - Se na questão anterior você respondeu que não conseguiu entender a Transparência ou entendeu somente parte das informações.	<i>Questão dissertativa dependente da questão anterior</i>
Q4 - Em relação à Objetividade da Transparência de Dados Pessoais no cenário avaliado, qual a sua opinião?	(a) São Objetivas e (b) Não São Objetivas

Q4a - Se você achou as informações de Transparência do cenário avaliado NÃO OBJETIVAS, por favor, explique ou exemplifique o que levou a ter essa opinião.	<i>Questão dissertativa dependente da questão anterior</i>
Q5 - Em relação à Completitude das informações de Transparência do cenário avaliado, qual a sua opinião?	(a) As informações são completas; (b) Não há informações suficientes
Q5a - Se a completitude das informações de Transparência do cenário está ruim, por favor, nos diga o que poderia ser acrescentado para deixar esta informação de Transparência mais completa?	<i>Questão dissertativa dependente da questão anterior</i>
Q6 - Algumas informações são melhores apresentadas pelo uso de texto, outras por imagens (fotos, vídeos, etc), cores, formatos etc. Em relação ao formato de exibição das informações (textos, cores, imagens etc), qual sua opinião:	(a) Adequada ou (b) Inadequada ou (c) Há formatos que precisam ser melhorados
Q6a - Se você achou o Formato de exibição de alguma informação INADEQUADO, por favor, descreva qual informação podemos melhorar a forma de exibição.	<i>Questão dissertativa dependente da questão anterior</i>
Q7 - Se você tem alguma sugestão de mudança nas informações de Transparência as quais não foram solicitadas no questionário, por favor, descreva-as aqui.	Questão dissertativa
Q8 - Em uma escala de 1 - Nada Relevante até 5 - Muito Relevante, qual sua opinião sobre a Relevância das informações mostradas sobre o uso de seus dados pessoais?	Escala
Q9 - Em uma escala de: 1 - Nenhum até 5 - Muita, qual seria seu nível de confiança em um aplicativo ou website que fornecesse informações de uso de seus dados pessoais no formato apresentado no cenário avaliado?	Escala

Participaram da validação um total de 121 (cento e vinte e um) participantes. O perfil demográfico dos participantes é mostrado na Tabela 20.

Os participantes foram convidados a participar da validação voluntariamente. O convite para participação foi distribuído via e-mail e redes sociais. Antes da realização dos

Tabela 20: Perfil dos participantes da validação. Do autor.

Perfil	Opções	Participação (%)
Idade	até 18 anos	3,31
	18 - 35	95,04
	36 - 50	1,65
	+50	0,00
Gênero	Masculino	71,07
	Feminino	28,93
	Não informado	0,00
Escolaridade	Educação básica	0,00
	Ensino Médio	4,13
	Graduação	95,04
	Pós Graduação	0,83
Área de atuação	Ciências exatas e computação	82,64
	Ciências sociais e humanas	5,79
	Comunicação e informação	1,65
	Meio ambiente	4,13
	Outros	5,79
Usa apps	Raramente	15,70
	Frequentemente	84,30

testes, os participantes eram avisados de que o objetivo da validação era o TR-Model e que eles/elas não estavam sendo avaliados em nenhum aspecto, e que poderiam desistir de participar a qualquer momento sem qualquer impacto aos mesmos. Foram disponibilizados e-mails e telefone para eventuais contatos.

Acredita-se que o perfil dos participantes representou de forma significativa uma parte dos usuários de *websites* e aplicativos móveis uma vez que não ocorreu o predomínio de um categoria específica de pessoas. As áreas de atuação em Ciências Exatas e Computação foram a de maior incidência porque muitos participantes eram alunos, professores e/ou funcionários de instituições que ofertam cursos de computação, matemática e física.

Os participantes realizaram 320 avaliações em todos os cenários disponíveis. A quantidade de avaliações em cada cenário foram:

- Cenário 01: 71 avaliações;
- Cenário 02: 68 avaliações;
- Cenário 03: 62 avaliações;
- Cenário 04: 60 avaliações;
- Cenário 05: 59 avaliações.

5.2.3 Estratégia de análise dos dados

A primeira análise refere-se aos elementos de IHC, seguida pela análise das dimensões de IQ. A apresentação dos resultados e as discussões serão apresentadas na seguintes ordem:

- Resultados das questões objetivas considerando todos os cenários;
- Os resultados por cenário avaliado e caso algum cenário tenha apresentado resultados discrepantes em relação à média geral, tal resultado será discutido;
- As questões dissertativas respondidas para justificar ou complementar alguma resposta das questões objetivas.

A fim de criar uma classificação para os resultados foi estabelecida a seguintes estratégia:

- Para questões com escala entre 1 e 5 os resultados serão classificados considerando as respostas: 1 -Nada Satisfatório, 2 - Pouco Satisfatório, 3 - Razoavelmente satisfatório; 4 - Satisfatório; 5 - Muito Satisfatório.
- Para questões com opções de respostas como Adequado, Razoável, Apropriado, foram consideradas as descrições das alternativas das respostas.

Por exemplo, em uma alternativa com respostas em escala com percentual maior de respostas para a nota 4 é considerada como Satisfatória. Já uma questão com alternativas de Apropriado, Razoável e Não Apropriado, a qual teve o maior percentual de respostas em Razoável terá tal classificação.

Para as análises de questões dissertativas será utilizado o termo **compilação das respostas**. Esse termo refere-se ao tratamento das respostas dissertativas a fim de identificar e ajustar respostas que poderiam dificultar ou impedir sua análise sem interferir ou enviesar em seu conteúdo ou, devido ao teor da pergunta, necessitaria de um tratamento. São exemplos de tratamento: identificação e soma de respostas duplicadas, descarte da resposta por uso incorreto da alternativa e/ou identificação de mais de uma resposta na sentença para questões que, por exemplo, solicitaram uma lista de sugestões para os participantes.

A próxima subseção apresenta os resultados das características de IHC.

5.2.4 Resultados das características de IHC

A validação das características de IHC teve por objetivo verificar a eficácia das descrições de metadados e metaevento no que tange à capacidade de identificação de conteúdo e o formato de exibição da Transparência para os indivíduos. Para esta análise foram consideradas duas questões previamente apresentadas na Tabela 19 na página 98.

A primeira questão avaliada foi a Q1 (**Em uma escala de 1 - Muito Ruim a 5 - Muito bom. Qual sua opinião em relação à facilidade para encontrar e identificar uma informação sobre o uso de seus dados pessoais?**); e seu resultado, considerando a análise de todos os cenários, é mostrado na Figura 20.

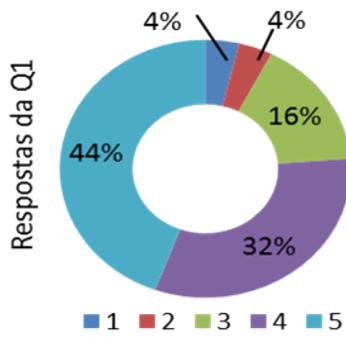


Figura 20: Resultados da Q1 considerando todos os cenários. Do autor.

O resultado da Q1 mostra que 76% das pessoas atribuíram as duas melhores avaliações para o critério de localização e identificação de uma informação de Transparência. Credita-se esse resultado a um conjunto de diretrizes simplificadas com rótulos (títulos dos metadados e metaeventos) pequenos e claros, além do fato de que houve uma preocupação para que a própria informação de uso dos dados tivesse um volume de textos pequeno e assim facilitasse sua inserção e organização no *design*.

Já a análise da Q1 por cenário apresentou os resultados mostrados na Figura 21.

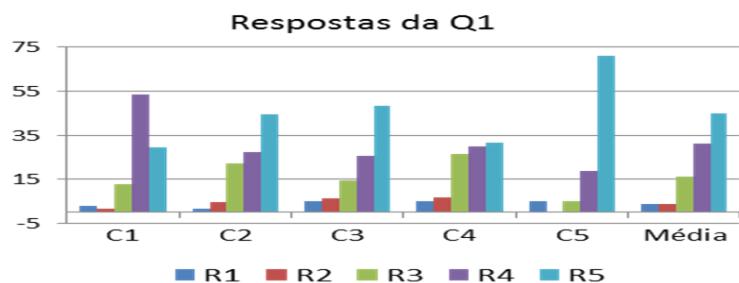


Figura 21: Resultados da Q1 organizadas por cenários. Do autor.

Particularidades nos resultados foram identificadas e são explicados a seguir:

O resultado da Q1 no cenário 1 mostrou que 53% dos participantes avaliaram a questão com nota quatro (satisfatório). Por ser a segunda melhor nota possível para a questão Q1, o Cenário 1 difere dos demais cenários que obtiveram o maior percentual de respostas na nota 5 (muito satisfatório). A Q1 questionava a facilidade para identificar e entender a informação e assumiu-se que a diferença no resultado deve-se ao seu foco, que era a Transparência em relação ao propósito de uso e aos dados pessoais. A descrição dos metadados e metaeventos dessas entidades tinham perfil textual e assim as informações foram mostradas com textos agrupados em quadros e por contexto, que pode ter sido interpretado como uma poluição de conteúdo por alguns participantes dificultando assim dificultado a identificação da informação.

Na Q1, no cenário 4 os resultados da avaliação Razoável, Satisfatória e Muito Satisfatória tiveram valores muito próximos. Os valores são superiores aos resultados das avaliações ruins, mas demonstram que não houve um consenso dos participantes em relação à facilidade de localização e identificação da informação. O cenário 4 buscava lembrar o indivíduo dos eventos envolvidos na concessão de autorização de uso de seus dados pessoais. Por ser uma informação sensível e de forte interesse dos indivíduos, assume-se que os mesmos encontraram facilmente as informações que estavam em destaque, mas que podem ter procurado por mais informações e não conseguiram identificar, uma vez que a Transparência estava limitada às imagens com os destaques para algumas informações.

Também na **Q1, mas no cenário 5** ocorreu um situação contrária a do cenário 1, pois o cenário 5 teve 72% dos participantes indicando nota 5 (cinco). Credita-se esse resultado à organização do conteúdo em forma de fichas cadastrais dos atores e esse padrão de *design* é muito conhecido por indivíduos, o que facilita a identificação da informação.

No caso do Cenário 5, o *design* contribuiu, mas não deve ser descartado o fato do conhecimento prévio do tipo da informação mostrada (ficha cadastral) ter sido determinante para o resultado. Para os outros cenários, as avaliações mostraram uma distribuição diferente com valores menores para a avaliação 5 (cinco), pois eram informações sobre o uso dos dados pessoais com as quais as pessoas não estavam tão familiarizadas e necessitaram dispensar mais ações físicas e cognitivas para analisá-las.

Outra questão avaliada como fator de IHC foi a Q6 (**Algumas informações são melhores apresentadas pelo uso de texto, outras por imagens (fotos, vídeos, etc), cores, formatos etc. Em relação ao formato de exibição das informações (textos, cores, imagens etc), qual sua opinião:**). O resultado dessa questão considerando todos os cenários são mostrados na Figura 23.

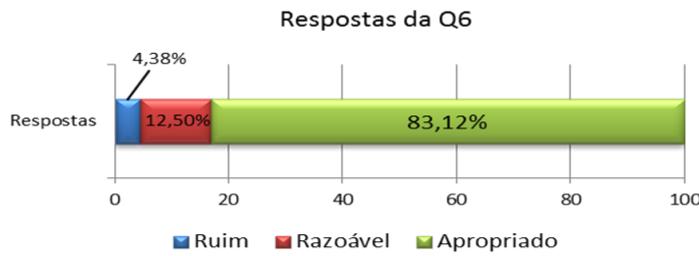


Figura 22: Resultados da Q6 considerando todos os cenários. Do autor.

A Q6 foi uma questão voltada para o formato de exibição e obteve um percentual em que 86% dos participantes indicaram que o formato de exibição foi **apropriado**. Esse resultado demonstra que os componentes de interface utilizados para construir o cenário permitiram a construção de uma interface com bons níveis de usabilidade, isso porque a forma de apresentação adequada é fundamental para atrair a atenção do indivíduo e criar um ambiente favorável para a análise das informações. Embora a decisão sobre os componentes de interface pode ser tomada pelo *designer* considerando as necessidades de projeto, nesta pesquisa, na construção dos cenários de validação, foram considerados os *design patterns* para Transparência propostos por Coleti et al. (2019).

Após a análise dos resultados de forma geral, estes foram organizados por cenários para análise de eventuais cenários discrepantes, nesse caso, resultados consideravelmente diferentes dos demais e que poderiam significar algum problema no TR-Model.

Os resultados agrupados por cenários e com a média geral são mostrados na Figura 23. Os gráficos mostram que os resultados por cenário mantiveram o mesmo padrão de respostas em relação à média geral (campo **média** dos gráficos).

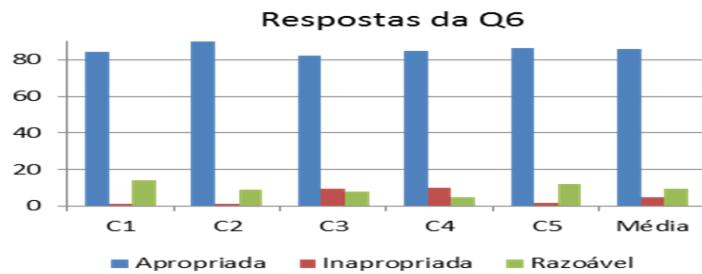


Figura 23: Resultados das Q6 organizados por cenários. Do autor.

Por existir a possibilidade de uma resposta desfavorável, aqueles que escolheram tal opção justificaram sua escolha na Q6a (**Se você achou o Formato de exibição de alguma informação inapropriado, por favor, descreva qual informação podemos melhorar a forma de exibição.**). Em relação às respostas negativas ou de ne-

cessidade de melhorias, as considerações feitas por indivíduos elencaram questões como: utilização de cores; organização de conteúdo (seguir uma sequência), melhoria na qualidade de imagens; agrupamento de informações; melhor utilização de ícones; e organização de informações em uma única página. Os problemas apresentados não foram considerados críticos para a Transparência e poderão guiar a melhoria dos *design patterns* de Transparência em trabalhos futuros.

Por fim, os resultados mostraram que os elementos de IHC aplicados para o TR-Model foram apropriados para dar suporte à visualização da informação de Transparência dos dados pessoais pelos indivíduos. Nas duas questões utilizadas para avaliação, os resultados foram classificados como satisfatório, muito satisfatório ou apropriados sendo enquadrados nas melhores avaliações possíveis.

As avaliações negativas e suas respectivas justificativas, embora em menor número, permitiram identificar pontos para os quais melhorias deverão ser consideradas, principalmente para apresentação de informações mais sensíveis para o indivíduo tais como: permissão de uso dos dados pessoais, informações sobre propósito e uso e como os dados serão utilizados.

A próxima subseção apresenta os resultados e as discussões das dimensões de IQ.

5.2.5 Resultados das dimensões de Qualidade de Informação

A análise das dimensões de IQ buscou verificar a eficácia do TR-Model em auxiliar a produção de informações de Transparência que atendessem as expectativas de qualidade para o indivíduo. Conforme já mencionado há diversos *frameworks* para apoiar a avaliação da IQ e cada um deles com foco em diferentes aplicações de qualidade, tais como semântica, estruturas dentre outros. Entretanto, considerando a particularidade da Transparência do TR-Model e a relação com o usuário final na validação, foi realizada a validação com análise semelhante à realizada por Kumar e Jakhar (2010), ou seja, por meio de questionários, seguidos por análise com base em estatística descritiva e discussão dos resultados.

As dimensões de IQ aplicadas e avaliadas para o TR-Model são apresentadas na Tabela 7 e as questões utilizadas para coletar a opinião dos participantes foram: Q2, Q3, Q4, Q5, Q8, Q2a, Q3a, Q4a e Q5a apresentadas na Tabela 19.

Os resultados da Q2 (**Em relação à quantidade de informação apresentada na Transparência do cenário avaliado, sua opinião é...**) são mostrados na Figura 24.

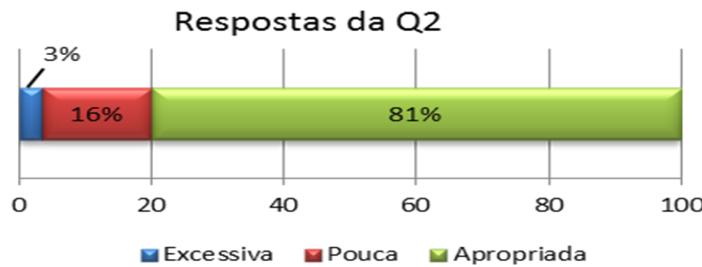


Figura 24: Resultados da Q2 considerando a soma de todos os cenários. Do autor.

Na dimensão de **quantidade de informação** (Q2), os resultados apontaram que a quantidade de informações foi considerada **apropriada** por 81% dos participantes. Esse resultado deve-se ao fato de que as descrições dos metadados e metaeventos estipularam limites claros de quantidade de palavras, sentenças e das informações de Transparência, de forma geral, para valores limitados. Essa limitação facilitou a produção de conteúdos resumidos e permitiu priorizar a Transparência necessária para o indivíduo e evitar o excesso de conteúdo desnecessários.

Já os resultados da Q2 por cenário são mostrados na Figura 25, na qual é possível concluir que os cenários tiveram comportamentos similares. Uma discrepância ocorre no cenário 4, em que 37% dos participantes indicaram que havia pouca informação de Transparência. Esse resultado é atribuído ao uso de imagens estáticas que dificultavam a navegação (mergulhar) por demais informações (efeito *drill-down* (STAIR; REYNOLDS, 2015)).

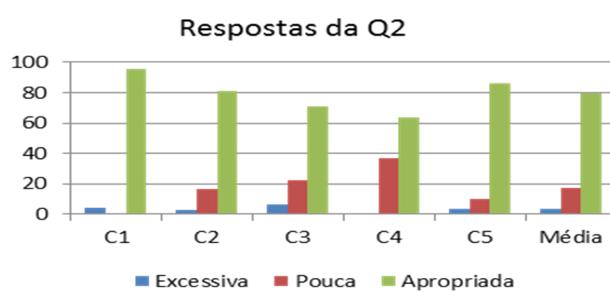


Figura 25: Resultado da Q2 por cenário. Do autor.

Embora com uma avaliação satisfatória para a Q2, 19% (64 respostas) apontaram para uma quantidade de informações Excessiva ou Pouca. Foi conduzida a compilação das justificativas apresentadas na Q2a (Se você achou a quantidade de informações sobre o Propósito de Uso POUCA ou EXCESSIVA. Por favor, descreva qual, ou quais informações poderíamos acrescentar (se você achou que há POCAS informações) ou quais informações estão em excesso (se você achou que

há MUITAS informações.) o que permitiu concluir que a alternativa de **pouca informação** foi apontada por participantes que procuraram informações adicionais àquelas mostradas pelos cenários. Assume-se que essa situação ocorreu devido às informações apresentadas estarem muito delimitada no cenário; ou ainda os participantes podem não ter compreendido essa delimitação e buscaram informações além do escopo do cenário.

Essa situação levantou uma questão a ser considerada na implementação de Transparência, qual seja, que uma informação de uma entidade sempre deve ter uma ligação na interface com as demais entidades que a complementam de forma que o indivíduo possa se aprofundar na análise.

Já em relação ao **excesso de informação**, as justificativas não foram claras e se limitaram a citar que algumas informações não eram relevantes, mas os participantes não descreveram detalhes, não sendo possível identificar um padrão em relação ao problema.

Com a dimensão da **capacidade de compreensão**, avaliada pela Q3 (**Em relação à Facilidade de Entendimento e Interpretação da Transparência do cenário avaliado, qual a sua opinião?**), o TR-Model buscou orientar a criação de informações de Transparência com o uso de linguagem adequada, principalmente nos aspectos de significado, ambiguidade, idioma e formato, a fim de garantir a compreensão do conteúdo pelo indivíduo. Os resultados desta questão são mostrados na Figura 26.

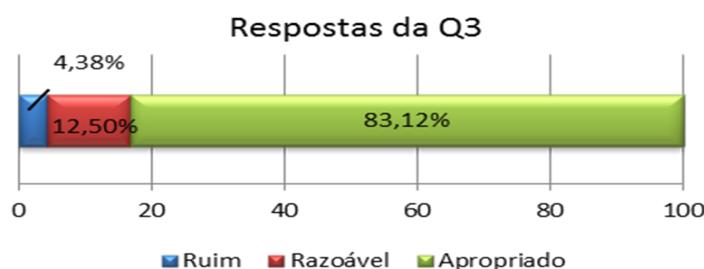


Figura 26: Resultado da Q3 considerando todos os cenário. Do autor.

Essa dimensão foi considerada apropriada por 83% dos participantes, o que indica que o TR-Model pode ser aplicado como um guia para o software fornecer a dimensão capacidade de compreensão de forma satisfatória na Transparência, a fim evitar que o indivíduo necessite realizar interpretações complexas que exijam grande carga cognitiva.

O bom resultado obtido na dimensão de capacidade de compreensão manteve-se quando analisado por cenário, conforme mostrado na Figura 27. Assim como na dimensão anterior, uma leve discrepância nos resultados ocorreu no cenário 4 e apresentou um avaliação Razoável, maior que os demais cenários, o que impactou na avaliação *Apro-*

priada com resultado menor. Esse resultado é justificado pelo perfil do cenário 4 ser um conjunto de imagens estáticas, o que pode ter dificultado o uso e entendimento para alguns usuários. Considerações sobre melhorias nesse aspecto já foram discutidas no texto.

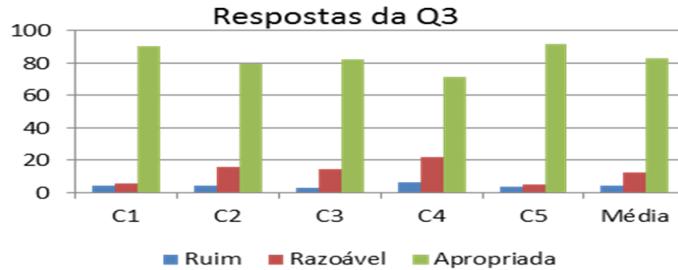


Figura 27: Resultado da Q3 por cenário. Do autor.

Atribui-se também o bom resultado dessa dimensão, aos bons resultados das avaliações dos elementos de IHC em conjunto com o bom resultado da dimensão de quantidade de informação. O vínculo da dimensão de capacidade de compreensão com os elementos de IHC e a quantidade de informação pode ser comprovado após cruzamento dos resultados dessas questões com os resultados dessa dimensão ao verificar qual o resultado da avaliação da dimensão de capacidade de compreensão para todos os participantes que responderam os valores 4 ou 5 para Q1 ou Adequado para Q2 ou Apropriado para Q6. Os resultados dessa análise são mostrados na Tabela 21.

Tabela 21: Resultado da análise da dimensão de Compreensão das informações com elementos de IHC. Do autor.

Resultado	Percentual
Ruim	2,45%
Razoável	4,90 %
Adequada	92,65 %

Os números mostram que a dimensão de capacidade de compreensão foi considerada Adequada por 92,65% dos participantes que apontaram avaliações satisfatórias/adequadas para os elementos de IHC e para a dimensão de Quantidade de Informação. Assim, a relação entre a eficácia desses parâmetros com a avaliação positiva da dimensão de capacidade de compreensão corrobora que o uso correto de componentes de interface e organização de informação coerente, além de um volume de informação apropriado, contribuem para a eficácia dessa dimensão da IQ.

Portanto, o bom resultado dessa dimensão foi atribuído a: (1) fato que a Transparência ser apresentada com formatos apropriados ao indivíduo evidenciando informações de interesse do indivíduo; e (2) as descrições dos metadados e metaeventos que orientam a

produção informações escritas de maneira simples, no idioma do indivíduo, sem palavras de que remetiam à necessidade de conhecimento de áreas específicas, sem abreviações ou símbolos. Com isso, as diretrizes evitam a sobrecarga de informação e o desperdício de tempo e esforço com a necessidade de decodificar símbolos ou interpretar conteúdos complexos.

A análise das opiniões desfavoráveis serão discutidas posteriormente, pois os resultados dessa dimensão se confundem com os resultados da dimensão de objetividade (próxima a ser discutida).

Considerou-se então que, o TR-Model foi eficaz na aplicação da dimensão de capacidade de compreensão para a Transparência de Dados Pessoais e que, aplicações que venham a utilizar esse modelo podem produzir informações sobre o uso dos dados, serão favoráveis para os indivíduos. A relação dessa dimensão com elementos de IHC permite, ainda, concluir que a melhoria ou inclusão de novas características de IHC podem elevar ainda mais a eficácia da dimensão.

Assume-se que proporcionar essa dimensão é fundamental para garantir a confiabilidade de um indivíduo no software uma vez que, ao entender o que acontece com seus dados e perceber isso no comportamento do software, poderá levá-lo a concluir que a aplicação não realiza ações indevidas.

A dimensão da **objetividade**, avaliada pela questão Q4 (**Em relação à Objetividade da Transparência de Dados Pessoais no cenário avaliado, qual a sua opinião?**), foi a melhor avaliada. Essa dimensão poderia ser avaliada por análise de especialistas com o uso de *checklists* a fim de inspecionar se a informações está livre de enviesamento, se é imparcial e não prejudicial ao indivíduo ou ao controlador. Essa checagem foi feita durante o processo de desenvolvimento do TR-Model uma vez que ao adotar essa dimensão pretendia-se garantir que a Transparência se resumisse a informar como os dados pessoais seriam usados, quem os utilizaria e como o indivíduo deveria agir para se defender, pois o foco do conteúdo deveria ser sempre a pessoa.

Entretanto, para todas as dimensões foi considerada a necessidade de avaliação dos participantes e para a dimensão de objetividade, não foi diferente. Analisar a opinião dos participantes nessa dimensão buscou garantir que os mesmos viram a Transparência como: (1) imparcial: procurava informar ao indivíduo sem tomar opinião sobre as ações em prol de algum envolvido; (2) não prejudicial: que causasse algum transtorno ao indivíduo por causa de alguma interpretação, ou causasse algum constrangimento; e (3) livre de direcionamentos: a fim de evitar favorecer ou guiar o conteúdo para algum contexto

benéfico para um, e ruim para outro.

Os resultados da dimensão de objetividade considerando todos os cenários são mostrados na Figura 28.



Figura 28: Resultado da Q4 considerando todos os cenário. Do autor.

Assim como a avaliação geral, os resultados de cada cenário também foram favoráveis, uma vez que a avaliação positiva em cada cenário ficou acima de 80% conforme mostrado na Figura 29.

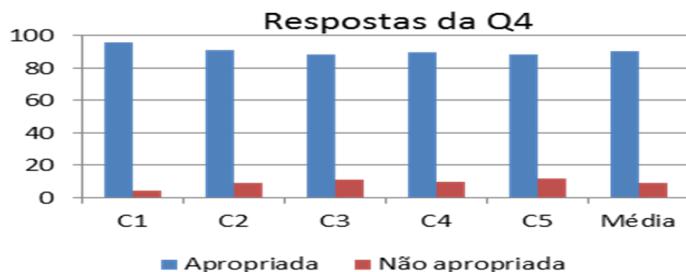


Figura 29: Resultado da Q4 por cenário. Do autor.

Com o perfil de objetividade citado e então aplicado ao TR-Model, essa dimensão obteve os melhores resultados entre todas as dimensões, com percentual superior a 90% para a avaliação **adequado**. Esse resultado, claramente, coloca a dimensão de objetividade como um ponto forte das diretrizes do TR-Model.

O julgamento dos participantes sobre essa dimensão corrobora com as intenções do TR-Model para a objetividade na Transparência e permite concluir que essa dimensão atendeu de forma muito satisfatória seu objetivo. Assim, é possível assumir que o software que fizer o uso dessa dimensão do TR-Model vai proporcionar a Transparência com informações única e exclusivamente sobre o uso dos dados, com o maior número de informações possíveis, fornecendo insumos para o indivíduo decidir sobre a corretude das ações, mas sem tomar posição sobre eventuais subjetividades deixando a decisão única e exclusivamente sob responsabilidade do indivíduo.

No que tange o pequeno percentual de avaliação negativa, as respostas das questões **Q3a (Entendimento e Interpretação)** e **Q4a (Objetividade)** tiveram certa similaridade em seu conteúdo. Assume-se que isso ocorreu por interpretação semelhante das questões por parte dos participantes, uma vez que a definição apresentada para objetividade (imparcialidade, enviesamento e não prejudicial) pode não ter sido bem assimilada por eles. Esse problema limitou uma análise efetiva das sugestões específicas para a Objetividade.

Foram registradas 54 sugestões/justificativas para a dimensão de Entendimento e Interpretação, e 29 para a dimensão da Objetividade. Seus textos apresentavam certo grau de detalhamento, mas não o suficiente, e sua análise e sumarização foi feita com inferências por parte do pesquisador. Assim, as seguintes justificativas foram consideradas: (1) Design inadequado; (2) Falta de informação ou Excesso de informação; (3) Formato de apresentação e (4) Conteúdo técnico.

O **Design** e o **Formato de Apresentação** são elementos de IHC que podem ter interferido na qualidade das informações e as considerações sobre esses fatores foram discutidas nas seção anterior. Já a **falta e/ou o excesso de informação** podem ter ocorrido por fatores como: (1) a falta de alguma informação desejada por algum indivíduo; (2) a organização das informações na tela que ter levado o indivíduo a julgar o ambiente poluído demais pelas informações em excesso; ou (3) o indivíduo procurou por informações além do escopo do TR-Model ou do escopo do cenário avaliado.

Com o propósito de verificar quais as sugestões de informações apresentadas por essas pessoas, foram cruzados os dados para as questões de pouca Objetividade e também de baixa capacidade de Entendimento e Interpretação com os dados da Quantidade de informações e as sugestões de dados que poderiam ser acrescentados ou retirados (Q2a). Foram identificadas 15 respostas (65,22% dos respondentes para a Q2a) as quais as justificativas foram analisadas por cenário e contemplavam conteúdos como:

- C1: Excesso de informação: pode ter ocorrido em razão do cenário apresentar informações sobre propósito de uso, leis e dados pessoais e essas entidades são majoritariamente textuais;
- C2: Falta de informações de quais dados serão utilizados: Não foi possível identificar o motivo da falta de informações, uma vez que todas os metadados e metaeventos da entidade *Transfer* e suas relações foram apresentadas;
- C3: Falta de detalhes sobre quais dados serão utilizados: Conclui-se que foi uma

falla na construção dos cenários, pois as informações sobre os dados pessoais mostradas no cenário limitaram-se à descrição, não havendo formas de consultar mais detalhes;

- C4: Falta de informações: Idem ao cenário C3, uma vez que eram exibidas imagens estáticas sem possibilidade de navegar em mais detalhes;
- C5: Falta de informações de como os dados serão utilizados: Não era o objetivo do cenário, embora a construção de *links* para demais informações seja necessária.

Assim, concluiu-se que os problemas por excesso de informação deram-se pelas características das entidades que proporcionavam informações para os cenários, pois utilizavam textos como principal forma de apresentação. O uso exclusivo ou majoritário de textos pode gerar a sensação de interface poluída. Já a falta da informação deu-se por duas razões: (1) construção de cenários engessados onde as informações do uso dos dados pessoais limitavam-se a *cases* muito específicos; e (2) necessidade de *links* para permitir a navegação aprofundada para determinada informação.

No TR-Model foram contemplados os relacionamentos entre as entidades, mas não estabelecidos critérios de apresentação visual dos mesmos. Entretanto, foi possível perceber a necessidade de, em futuras versão do modelo, estabelecer diretrizes voltadas para o relacionamento de informações.

O **conteúdo técnico** será uma questão a ser revisada no TR-Model a fim de identificar algum eventual descuido na descrição dos metadados e metaeventos ou na produção dos cenários de Transparência, uma vez que o modelo visa evitar o máximo possível essa situação justamente, para não dificultar a compreensão.

Já a dimensão de **completitude**, na qual a Q5 (**Em relação à Completitude das informações de Transparência do cenário avaliado, qual a sua opinião?**) registrou a opinião dos participante, foi avaliada com base no conceito apresentado por Kandari et al. (2011b) e por Wang e Strong (1996) em que afirmam que a informação não pode estar faltando e deve ser de amplitude e profundidade suficientes para realização da tarefa, nesse caso, análise como ocorre o uso dos dados pessoais.

Nessa pesquisa também assumiu-se a completitude como o fato do indivíduo não precisar procurar ou complementar a Transparência em outras fontes de dados e assim evitar a sobrecarga de trabalho e uma possível divergência de conceitos de Transparência em relação ao proposto pelo TR-Model, situação que dificultaria ainda mais o entendimento do uso dos dados. Os resultados da Q5 considerando todos os cenários são mostrados na

Figura 30.

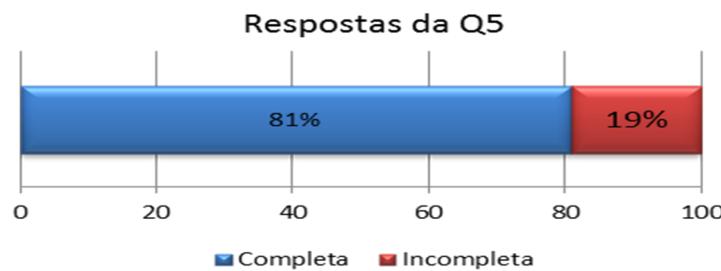


Figura 30: Resultado da Q5 considerando todos os cenário. Do autor.

O resultado de 81% de respostas indicando que a Transparência estava completa permitiu concluir que as expectativas para a dimensão foram atingidas com eficácia. Esse percentual de participantes indicou que os mesmos obtiveram todas as informações desejadas sobre o uso dos dados pessoais em detalhes necessários e inseridas no protótipo avaliado.

Já os resultados da análise da Q5 por cenário são mostrados na Figura 31.

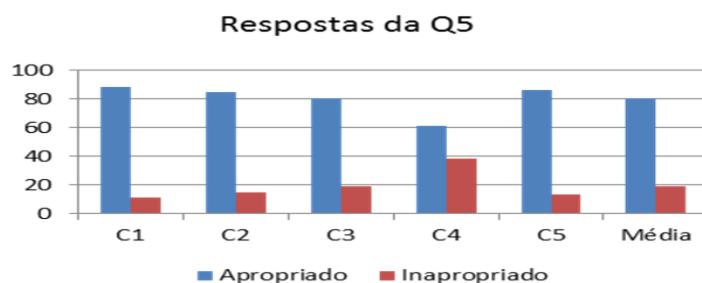


Figura 31: Resultado da Q5 por cenário. Do autor.

A análise da dimensão da **completitude** por cenário identificou que o cenário 4 apresentou um comportamento levemente desfavorável em relação aos demais cenários na questão Q5. A questão tratava a completitude das informações e o cenário informava sobre o momento/ação em que ocorre a permissão de uso dos dados com imagens estáticas. Embora o resultado das outras questões estar na média, o uso do álbum de fotos apenas com imagens sobre a permissão de acesso, pelo menos com imagens estáticas, não foi tão bem aceito quanto o esperado.

Com esse resultado, assume-se que as aplicações e sites que utilizarem o TR-Model poderão garantir que todas as informações necessárias estejam disponíveis com detalhes de amplitude e profundidade necessários para a análise do indivíduo, sem requerer a busca de informações em outras fontes. Assim, o software poderá melhorar a aceitabilidade pelo

usuário ao reduzir a carga de trabalho do indivíduo e evitar conflitos de conceitos devido ao uso de informações de fontes que não tenham uma definição confiável dos contextos de Transparência.

No que tange as avaliações desfavoráveis para a dimensão de **completitude**, as justificativas compreenderam: a falta de informação ou informação mal apresentada. Entretanto, os participantes não explicaram se a informação não se encontrava no cenário ou se era preciso buscar as informações em outras fontes. Outro fator que influenciou a completitude foi a necessidade pessoal por informações de Transparência que estavam fora do escopo do cenário.

A dimensão de **relevância**, questionada pela Q8 (**Em uma escala de 1 - Nada Relevante até 5 - Muito Relevante, qual sua opinião sobre a Relevância das informações mostradas sobre o uso de seus dados pessoais?**) foi uma das grandes preocupações no desenvolvimento do TR-Model. Essa preocupação é justificada pelo fato de que as técnicas existentes costumam apresentar informações que não despertam o interesse do indivíduo, seja por seu conteúdo, sua forma de apresentação, a falta de objetividade ou, falta completitude, de forma proposital ou não, acabam por se tornar informações irrelevantes.

De certa forma, é possível considerar que a relevância é influenciada por todas as demais dimensões já discutidas, pois no uso de qualquer uma das dimensões as informações com algum *deficit* de qualidade deixam de ser relevantes. Dentro da dimensão de relevância, o TR-Model buscou propor Transparência de interesse ao indivíduo abstraindo e fornecendo informações que pudessem ser utilizadas para entender o uso dos dados pessoais e assim atender a definição de relevância de Wang e Strong (1996).

A dimensão de relevância no TR-Model foi considerada adequada por 31% dos participantes e muito adequada por 53%. Os resultados mostrados na Figura 32 permitem considerar que o TR-Model foi eficaz em suas diretrizes para considerar a relevância da Transparência. Esses resultados mostram que as informações de Transparência são utilizáveis pelos indivíduos e permitem que os mesmos possam entender os elementos envolvidos no uso de seus dados pessoais.

Quando avaliados por cenários, os resultados apresentados na Figura 33 mostraram um padrão de comportamento semelhante a análise geral para a dimensão de relevância. Destacou-se o cenário 5 com um avaliação positiva consideravelmente melhor que os demais cenários. Atribuiu-se esse resultado a simplicidade das informações apresentadas (ficha cadastral dos envolvidos) e o fato das informações facilitarem a identificação de



Figura 32: Resultado da Q8 considerando todos os cenário. Do autor.

canais de contato com os atores.

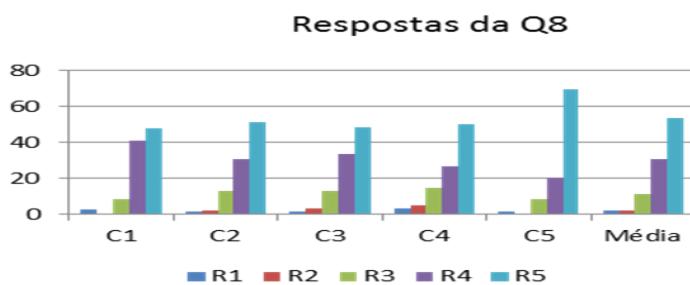


Figura 33: Resultado da Q8 por cenário. Do autor.

Portanto, com base nos resultados apresentados, concluiu-se que um software com Transparência baseada no TR-Model pode fornecer informações que sejam de interesse do indivíduo, e consequentemente, com informações relevantes, o sistema transmite ao indivíduo o máximo de informações possíveis, o que pode representar maior confiabilidade no software em relação ao uso dos dados e a continuidade de sua utilização.

A próxima seção apresenta a análise da satisfação geral dos participantes e também as sugestões de melhorias propostas para o TR-Model.

5.2.6 Análise e discussão da confiabilidade dos participantes e sugestões de mudança

Esta seção discute os resultados de duas questões aplicadas aos participantes para verificar a opinião geral dos mesmos em relação à Transparência criada com base no TR-Model.

A questão Q9 (**Em uma escala de: 1 - Nenhum até 5 - Muita, qual seria seu nível de confiança em um aplicativo ou website que fornecesse informações de uso de seus dados pessoais no formato apresentado no cenário avaliado?**) questionava sobre uma eventual confiabilidade dos indivíduos em aplicações de software

que apresentassem as informações de modo similar ao mostrado nos cenários. Essa questão justificativa-se nas ponderações de Murmann e Fischer-Hübner (2017a) em que destacam a Transparência como elemento fundamental para garantir a confiança do indivíduo no software.

Na Figura 34 são mostrados os resultados dessa questão para os cenários avaliado e a média geral da avaliação.

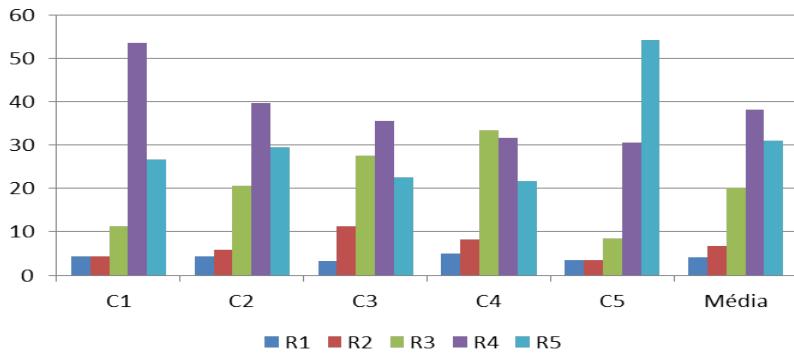


Figura 34: Confiança dos participantes na utilização hipotética do TR-Model para Transparência. Do autor.

Da Figura 34 é importante destacar os extremos negativos e positivos dos resultados. No extremo negativo, o resultado relacionados à nenhuma ou pouca confiança nas informações foi pequeno para todos os cenários. Embora não deve ser ignorado, esse resultado mostra que a Transparência do TR-Model conseguiu produzir informações para proporcionar alguma confiança ao indivíduo nas aplicações.

No que tange ao resultado positivo, prevaleu a nota 4 (de 1 a 5) em três dos quatro cenários avaliados e a nota 5 em um dos cenários. Esse resultado mostra que a Transparência proposta pelo TR-Model pode influenciar positivamente na confiança dos usuários. Assumiu-se que o cenário positivo deve-se a um conjunto de diretrizes objetivas, relevantes e simples e que obteve uma aceitabilidade geral (média dos resultados das questões) de mais de 80% .

Uma avaliação com comportamento diferente foi feita para o cenário quatro. Esse cenário obteve a maior avaliação como R3 (média) e, como já discutido, pode ser atribuído ao *design* estático e limitado, o que não permitiu um flexibilidade na busca por informações. De certa forma, a falta de confiança em aplicações com esse perfil de Transparência corrobora a necessidade por informações flexíveis e com maior poder de navegação a fim de permitir ao indivíduo formar um modelo mental mais consistente sobre o uso de seus dados.

Já a questão Q7 (**Se você tem alguma sugestão de mudança nas informações de Transparência as quais não foram solicitadas no questionário, por favor, descreva-as aqui**) solicitou aos participantes sugestões gerais de como melhorar o TR-Model. A questão era dissertativa não obrigatória e, ainda assim, foram obtidas 29 respostas. As respostas foram compiladas a fim de ajustar duplicidades, múltiplas sugestões por respostas ou o uso incorreto da questão para informar dados não solicitados. As sugestões apresentadas são mostradas na Tabela 22.

Tabela 22: Sugestões de Melhorias no TR-Model indicadas na Q7. Do autor.

Descrição	Tipos
IHC Maior utilização de figuras e animações para atingir um público maior e com pouca habilidade de leitura; Uso de tabelas para organização do conteúdo; Uso de letras maiores; Seleção de ícones mais apropriados para Transparência; Informações mais sucintas e apresentadas em listas.	
IQ Apresentar mais detalhes de onde os dados foram utilizados; Apresentar informações indicando se o uso dos dados de forma benéfica teria algum custo financeiro; Mostrar detalhes da mudança da autorização de uso da informação quando a utilização da mesma for diferente daquela autorizada no início; Destacar se há outras finalidades de uso além daquela apresentada.	

Das sugestões de IHC, as sugestões de maior uso de imagens e Informações mais sucintas contemplam diretamente especificações propostas no TR-Model. As demais solicitações estão relacionadas mais ao *design* da informação e deverão contemplar um futuro *design pattern* para Transparência baseado no referido Perfil de Aplicação de Metadados.

As sugestões de IQ contemplam futuras mudanças importantes como o maior detalhamento relacionado à permissão de uso de dados e em relação aos propósitos de uso dos dados pessoais. Essas melhorias poderão ser aplicadas em futuras versões do TR-Model.

O número relativamente pequeno de sugestões está entre os fatores que permitem considerar o TR-Model como eficaz para proporcionar Transparência ao indivíduo e que o mesmo pode apoiar aplicações de software em fornecer detalhes sobre o uso de seus dados pessoais de forma relevante, concisa e adequada ao indivíduo.

A próxima subseção apresenta as considerações finais sobre a validação do TR-Model.

5.2.7 Considerações sobre a Validação do TR-Model

A validação do TR-Model buscou verificar a eficácia do modelo nos contexto de **abrangência/cobertura necessidades de Transparência e experiência do usuário com as informações de Transparência**. O TR-Model foi desenvolvido com o objetivo de atender os indivíduos por meio de Transparência com formato adequado, relevante, concisa e compreensiva ao indivíduo de forma que o mesmo pudesse analisar como seus dados pessoais são utilizados e decidir sobre uma eventual intervenção.

Os resultados da validação de cobertura de informação permitiu concluir que o TR-Model:

- atendeu às diretrizes de Transparência estabelecidas pela GDPR, de forma que uma aplicação de software que venha a utilizar o modelo estará de acordo com essa regulamentação, assim como os indivíduos poderão dispor de mais confiança na aplicação, uma vez que a mesma atenderá a regulamentação supracitada;
- atendeu a 90% das expectativas de Transparência dos participantes diretamente em suas entidades, metadados e metaeventos e respectivas descrições. Demais expectativas não foram atendidas diretamente pelo TR-Model, mas poderiam utilizar de outros metadados e metaeventos similares. Já outras expectativas poderão ser adicionadas ao modelo em versões futuras.

O resultado satisfatório dessa validação, de certa forma, já era esperado porque a observação das regulamentações foi uma das principais ações durante a etapa de análise de requisitos. A validação da cobertura buscou verificar se o objetivo de atender as demandas da GDPR foi, de fato, atingido pelo TR-Model garantindo um *status* de modelo em *compliance* com a GDPR.

Em relação às expectativas dos participantes, o bom resultado deve-se à participação de pessoas durante todo o processo de desenvolvimento do TR-Model. Os *workshops*, entrevistas e a própria validação permitiram obter uma série de insumos que, trabalhados juntamente com as demandas da GDPR, e os conceitos técnicos e científicos, permitiram proporcionar um modelo com um conjunto de informações muito próximo ao esperado pelos indivíduos. As expectativas não contempladas podem ser consideradas pontuais, ou que chegaram a ser discutidas nos projetos, mas que, a princípio, acreditava-se não representar o interesse dos indivíduos.

Na avaliação da Experiência do Usuário foram utilizados cenários baseados em um

sistema de acompanhamento de atividades físicas que permitiu simular situações realistas de uso e também situações de necessidade de Transparência pelos usuários. As informações de Transparência foram produzidas com base nas entidades, metadados, metaeventos e descrições propostos pelo TR-Model.

Os resultados mostraram que a Transparência apresentada foi considerada satisfatória uma vez que 80% das respostas dos participantes se enquadram em critérios de satisfatório ou muito satisfatório nos quesitos de IHC e de IQ. Dos cinco cenários avaliados, os cenários C1, C2, C3 e C4 foram considerados muito satisfatórios e o cenário quatro (C4) teve uma avaliação satisfatória, mas com diversas considerações dos participantes devido à técnica estática de exibição de informação que limitou o acesso as informações complementares e também de interesse dos mesmos.

Sugestões e críticas eram esperadas, uma vez que ao realizar testes com usuários, uma série de situações não previstas podem ocorrer. As sugestões e críticas foram considerados importantes *inputs* que serão utilizados para auxiliar nas melhorias nas próximas versões do TR-Model.

O próximo capítulo apresenta as Considerações Finais dessa pesquisa.

6 CONSIDERAÇÕES FINAIS

Essa tese apresentou o TR-Model, um modelo baseado em MAP para apoiar a produção de informações sobre o uso dos dados pessoais a fim de prover aplicações com Transparência de Dados Pessoais com formato de apresentação adequado, informações relevantes e compreensíveis para os titulares dos dados. O TR-Model foi composto por entidades, metadados, metaeventos e descrições de como aplicá-los em aplicações de software.

A proposta do TR-Model surgiu da necessidade de melhorar a Transparência sobre o uso dos dados pessoais para os indivíduos. Embora já existissem recursos que procuravam informar os titulares dos dados, os mesmos não observavam determinados aspectos inerentes aos indivíduos o que dificultava a utilização das informações para analisar o uso dos dados pessoais. Assim, foram identificadas necessidades de melhorias na forma de apresentação, no conjunto de informações considerados como Transparência e em elementos de qualidade de informação como relevância, capacidade de compreensão e objetividade.

O TR-Model foi desenvolvido com base na abordagem de um Perfil de Aplicação de Metadados (MAP). Embora um MAP seja comumente utilizado para estabelecer metadados para interoperabilidade de dados, suas características tais como definição de modelo de domínio, definição de metadados e consequente descrição desses metadados para o domínio se mostrou promissora, uma vez que poderia dar suporte a definição e descrição de um conjunto de informações de Transparência.

Considerando que, para condução dessa tese buscou-se obter o máximo de conhecimento sobre os requisitos do domínio de Transparência de Dados Pessoais, pode-se afirmar que a utilização do MAP foi eficaz e ao mesmo tempo objetiva para a proposta do TR-Model, uma vez que sua estratégia de construção de um perfil de metadados é relativamente simples quando se tem bons insumos sobre o requisito, permitindo assim a construção de um modelo de domínio que contemple as necessidades dos envolvidos.

Com isso, assumiu-se o objetivo principal da tese foi atingido com eficácia, uma vez que o uso da abordagem de Perfil de Aplicação de Metadados permitiu a construção de um

modelo de domínio, um conjunto de entidades, metadados, metaeventos. As descrições de uso dos metadados e metaeventos foram criadas considerando elementos como *Readability*, *Infovis* e Qualidade de Informação e principalmente pela participação de indivíduos em atividades como *workshops*, palestras e entrevistas em que foram discutidas questões relacionadas à Transparência de seus dados.

As validações conduzidas por meio de avaliação de cenários e resolução de questionários por usuários permitiu a simulação da utilização do TR-Model para uma situações real de uso de dados pessoais, assim como obter ricas *insights* dos participantes sobre a eficácia das informações do TR-Model.

A construção dos cenários, embora feita pelo próprio pesquisador para fins de obter a opinião dos participantes em relação ao uso dos dados pessoais, permitiu confirmar que todos os metadados, metaeventos e descrições podem ser implementadas em interfaces de software por meio de componentes de interfaces já existentes e comuns para desenvolvedores web ou de aplicações móveis.

Logo, constatou-se que: (1) o TR-Model atendeu a 90% das expectativas de Transparência apresentadas pelo indivíduo no questionário; (2) as avaliações relacionadas à forma de apresentação e a qualidade da informação do conteúdo foram vistas de forma positiva por 85% dos participantes; (3) o TR-Model pode dar suporte à Transparência de Dados Pessoais para atender as regulamentações da União Europeia (GDPR) e Brasileira (LGPD) na época da defesa dessa tese.

De fato, deve-se considerar que regulamentações/leis sofrem modificações com o tempo e por diversas razões o que pode deixar essa versão do TR-Model obsoleta. Assim, acompanhar a evolução das regulamentações e assim evoluir o TR-Model para atende-las pode ser considerada uma tarefa a fim de garantir que o modelo possa garantir informações de qualidade para os titulares dos dados.

Destaca-se também que o TR-Model orienta a forma de apresentação das informações, mas não vincula ou obriga a utilização de alguma padrão de interface específico. Esse aspecto abre possibilidades para os mais diversos meios de apresentação da informação de forma que possa atingir o maior público possível e facilitar a análise da interface, além de levantar possibilidades de novos componentes de interface para dar suporte à Transparência de Dados Pessoais.

Com base em todo o processo construção e validação do TR-Model, assumiu-se que o mesmo pode ser um modelo ou diretriz para empresas construírem aplicações de software transparentes para seus usuários.

No que tange as empresas desenvolvedoras de software, o TR-Model não especifica técnicas ou padrões de desenvolvimento específicos, mas visa orientar as empresas em como ser transparente para os titulares dos dados. Considerando o fato de que o TR-Model foi validado para ser bom para o indivíduo sem considerar eventuais dificuldades de implementação, o desenvolvimento de meios para as empresas utilizarem o TR-Model de forma eficaz, produtiva, e reutilizável em diversos projetos integra uma vasta gama de possibilidades de trabalhos futuros. Trabalhos futuros nesse sentido podem ter como objetivo tornar o TR-Model bom para as empresas (assim como já seria bom para os indivíduos) permitindo que controladores e processadores não rejeitem o TR-Model por dificuldades ou incapacidade de utilizá-lo em seus projetos.

Já os indivíduos titulares dos dados, acredita-se que, ao utilizarem um software com Transparência feita com base no TR-Model poderão acessar um conjunto de informações com qualidade, além de formato de apresentação e conteúdos adequados. Com essas características acredita-se que será facilitada a compreensão dos indivíduos e evitada a sobrecarga cognitiva e física para analisar os conteúdos, entender termos técnicos, identificar a importância e relevância da informações ou procurar informações em outras fontes. Assume-se também que os indivíduos apresentarão mais confiança nas aplicações que divulgarem as informações sobre como utilizam seus dados e, inseridos no fluxo do uso dos dados, poderão agir para garantir o *fair use* de seus dados pessoais.

Destaca-se também que o TR-Model é a primeira etapa de uma pesquisa que visa buscar a melhoraria a Transparência de Dados Pessoais, no que tange facilitar ou viabilizar que empresas desenvolvedoras de software sejam transparentes ao mesmo tempo que os indivíduos possam garantir que seus direitos ao acesso às informações a as possibilidade de monitorar e negociar o uso de seu dados pessoais seja uma realidade.

A inclusão dos indivíduos no centro do uso dos dados pessoais, como um agente atuante e não como um mero espectador, levanta uma série de questões voltadas aos aspectos humanos tais como ética, liberdade de expressão, liberdade de escolha, integridade e veracidade das informações dentre outras, que poderão fomentar futuras pesquisas.

Os resultados dessa pesquisa são limitados ao fato de que os cenários da validação foram construídos pelo próprio pesquisador. Mesmo seguindo todas as diretrizes do TR-Model, a escolha dos componentes de interface e a organização das informações na interface podem ter interferido positiva ou negativamente na avaliação. Pode ocorrer que a construção do mesmo cenário por outros desenvolvedores, com outros *user interface design patterns* e outros componentes de interface possam produzir resultados diferentes

em uma eventual avaliação.

Além das possibilidades de trabalhos futuros já apresentadas nesse capítulo, também pode-se considerar:

- Analisar a aplicabilidade das sugestões dos participantes da validação e, se aplicáveis, melhorar os metadados, metaeventos e descrições no TR-Model;
- Avaliar a aplicabilidade do TR-Model para regulamentações de dados pessoais diferentes da GDPR e da LGPD e aplicar ajustes e melhorias a afim de que o modelo possa atender mais países e organizações;
- Publicar o TR-Model na Internet a fim de disponibilizar o modelo para uso público;
- Pesquisar e desenvolver *User Interface Design Patterns* para dar suporte ao TR-Model para que o modelo possa também direcionar, de forma mais efetiva, a apresentação de conteúdo;
- Desenvolver um *framework* para reuso do TR-Model em aplicações e *website* de modo a facilitar a utilização do modelo por empresas de desenvolvimento de sistemas.

REFERÊNCIAS

- ABDULLAH, K.; CONTI, G.; BEYAH, R. A visualization framework for self-monitoring of web-based information disclosure. *IEEE International Conference on Communications*, n. June 2008, p. 1700–1707, 2008. ISSN 05361486.
- ABIB, G. A qualidade da informação para a tomada de decisão sob a perspectiva do sensemaking: Uma ampliação do campo. *Ciencia da Informacao*, v. 39, n. 3, p. 73–82, 2010. ISSN 01001965.
- ACKERMAN, M. S.; MAINWARING, S. D. Privacy Issues and Human-Computer Interaction. *O'Reilly & Associates*, p. 1–19, 2005.
- ALLARD, S. DataONE: Facilitating eScience through Collaboration. *Journal of eScience Librarianship*, v. 1, n. 1, p. 4–17, 2012. ISSN 2161-3974.
- AMAR, Y.; HADDADI, H.; MORTIER, R. Privacy-Aware Infrastructure for Managing Personal Data Personal Data Arbitering within the Databox Framework. *Proceedings of SIGCOMM 16*, p. 571–572, 2016.
- AMARAL, F. *Introdução à Ciência dos Dados*. Rio de Janeiro: Alta Books, 2016. 320 p. ISBN 9788576089346.
- APPS, A.; MACINTYRE, R. Dublin Core Metadata for Electronic Journals. *Lecture Notes in Computer Science*, v. 1923, p. 93–102, 2000. ISSN 03029743. Disponível em: <<http://eprints.rclis.org/12183/2/appsmacecdl2000.pdf>>.
- BARRETO, P.; SALGADO, L.; VITERBO, J. Assessing the Communicability of Human-Data Interaction Mechanisms in Transparency Enhancing Tools. *Proceedings of the 2018 Federated Conference on Computer Science and Information Systems*, v. 15, p. 897–906, 2018.
- BELLAMY, B.; ALONSO, C. Reframing data transparency .. *Centre for Information Policy Leadership and Telefónica Senior Roundtable*, v. 1, n. June, p. 1–20, 2016.
- BENYON, D. *Interação Humano Computador*. São Paulo: Pearson Education, 2011. ISBN 9788579361098.
- BIER, C.; KUHNE, K.; BEYERER, J. PrivacyInsight: The next generation privacy dashboard. *Lecture Notes in Computer Science (9857, LNCS)*, 2016.
- BIOLCHINI, J.; MIAN, P. G.; NATALI, A. C. C. *Systematic Review in Software Engineering*. [S.l.], 2005. 1–30 p.
- BONATTI, P. et al. Transparent Personal Data Processing : The Road Ahead. *Proceedings of International Conference on Computer Safety, Reliability and Security - Lecture Notes on Computer Science*, v. 10489, n. September, p. 337–349, 2017.

- BUDIU, R. *Quantitative vs. Qualitative Usability Testing*. 2017. Disponível em: <<https://www.nngroup.com/articles/quant-vs-qual/>>.
- CAFARO, F. Using embodied allegories to design gesture suites for human-data interaction. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*, p. 560, 2012. Disponível em: <<http://dl.acm.org/citation.cfm?id=2370216.2370309>>.
- CANADA, O. o. P. C. of. Metadata and Privacy: A Technical and Legal Overview. n. October, 2014. Disponível em: <[https://www.priv.gc.ca/information/research-recherche/2014/md{_}201410{_}.>](https://www.priv.gc.ca/information/research-recherche/2014/md{_}201410{_}.)
- CAVANILLAS, J. M.; CURY, E.; WAHLSTER, W. The Big Data Value Opportunity. *New Horizons for a Data-Driven Economy*, p. 3–11, 2016.
- CHRISTL, W. How Companies Use Personal Data Against People. Automated Disadvantage, Personalized Persuasion, and the Societal Ramifications of the Commercial Use of Personal Information. *CrackedLabs*, n. October, 2017. Disponível em: <<http://crackedlabs.org/en/data-against-people>>.
- COLETI, T. A. et al. Análise da Transparência de Dados Pessoais em Políticas de Privacidade de Dados. *Anais do IX Workshop sobre Aspectos da Interação Humano-Computador para a Web Social*, v. 9, p. 025–036, 2018. Disponível em: <<http://portaldeconteudo.sbc.org.br/index.php/waihcws/article/view/3893>>.
- COLETI, T. A. et al. Design Patterns to Support Personal Data Transparency Visualization in Mobile Applications. *Proceedings of HCII - Design Patterns to Support Personal Data Transparency Visualization in Mobile Applications*, v. 1, p. 46–62, 2019.
- CORDIS. *Virtual Open Access Agriculture & Aquaculture Repository*. 2017. Disponível em: <<https://cordis.europa.eu/project/rcn/204632/en>>.
- COYLE, K.; BAKER, T. *Guidelines for Dublin Core Application Profiles*. 2009. Disponível em: <<http://www.dublincore.org/specifications/dublin-core/profile-guidelines/>>.
- Crabtree, A. et al. Enabling the new economic actor: data protection, the digital economy, and the Databox. *Personal and Ubiquitous Computing*, Springer London, v. 20, n. 6, p. 947–957, 2016. ISSN 16174909.
- CRABTREE, A.; MORTIER, R. Human Data Interaction: Historical Lessons from Social Studies and CSCW. *Proceedings of 14 European conference on Computer Supported Cooperative Work*, p. 19–23, 2015.
- CRADOCK, E.; STALLA-BOURDILLON, S.; MILLARD, D. Nobody puts data in a corner? Why a new approach to categorising personal data is required for the obligation to inform. *Computer Law and Security Review*, v. 33, n. 2, p. 142–158, 2017. ISSN 02673649.
- CYBIS, W. d. A.; HOLTS, A. B.; FAUST, R. *Ergonomia e Usabilidade: Conhecimentos, Métodos e Aplicações*. São Paulo: Novatec Editora, 2015. 487 p.

- De Luca, C. *Dois na Web*. 2019. Disponível em: <<https://cbn.globoradio.globo.com/comentaristas/dois-na-web/CRISTINA-DE-LUCA-DOIS-NA-WEB.htm>>.
- DIALLO, S. Y. et al. Understanding interoperability. *Emerging M and S Applications in Industry and Academia Symposium 2011, EAIA 2011 - 2011 Spring Simulation Multiconference*, n. May 2014, p. 84–91, 2011.
- DROZD, O. Privacy pattern catalogue: A tool for integrating privacy principles of ISO/IEC 29100 into the software development process. *IFIP Advances in Information and Communication Technology*, v. 476, p. 129–140, 2016. ISSN 18684238.
- EARP, J. B. et al. Examining Internet privacy policies within the context of user privacy values. *IEEE Transactions on Engineering Management*, v. 52, n. 2, p. 227–237, 2005. ISSN 00189391.
- ELMQVIST, N. Embodied Human-Data Interaction. In: *CHI 2011 Extended Abstracts on Human Factors in Computing Systems*. [S.l.: s.n.], 2011. p. 1–4. ISBN 9781450302685.
- EPSTEIN, D. A. Personal Informatics in Everyday Life. n. Figure 1, p. 429–434, 2015.
- FAZLIOGLU, M. *Transparency and the GDPR: Practical guidance and interpretive assistance from the Article 29 Working Party*. 2017. Disponível em: <<https://iapp.org/news/a/transparency-and-the-gdpr-practical-guidance-and-interpretive-assistance-from-the-article\29-working-party/>>.
- FERREIRA, A.; LENZINI, G. Can Transparency Enhancing Tools Support Patient's Accessing Electronic Health Records? *New Contributions in Information Systems and Technologies. Advances in Intelligent Systems and Computing*, v. 353, p. 1121–1132, 2015. Disponível em: <http://link.springer.com/10.1007/978-3-319-16486-1{_}>.
- FEW, S. Data Visualization for Human Perception. *The Encyclopedia of Human-Computer Interaction*, v. 2, p. <https://www.interaction-design.org/literature/book>, 2016.
- FILGUEIRAS, L. V. L. et al. Keep System Status Visible: Impact of Notifications on the Perception of Personal Data Transparency. *Human-Computer Interaction. Perspectives on Design*, v. 1, p. 513–530, 2019.
- FISCHER-HÜBNER, S. et al. Transparency, privacy and trust – Technology for tracking and controlling my data disclosures: Does this work? *IFIP Advances in Information and Communication Technology*, v. 473, p. 3–14, 2016. ISSN 18684238.
- FISCHER-HÜBNER, S.; ANGULO, J.; PULLS, T. How can cloud users be supported in deciding on, tracking and controlling how their data are used? *IFIP Advances in Information and Communication Technology*, v. 421, p. 77–92, 2014. ISSN 18684238.
- GOMES, P.; GAMA, S.; GONÇALVES, D. Designing a personal information visualization tool. *Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries - NordiCHI '10*, p. 663, 2010. Disponível em: <<http://doi.acm.org/10.1145/1868914.1868999{\\\%}5Cnhttp://portal.acm.org/citation.cfm?doid=1868914.1868>>.

GOOGLE. *Como o Waze funciona?* 2017. Disponível em: <<https://support.google.com/waze/answer/6078702?hl=pt-BR>>.

GUARDA, P.; RANISE, S.; SISWANTORO, H. Security analysis and legal compliance checking for the design of privacy-friendly information systems. *Proceedings of ACM Symposium on Access Control Models and Technologies, SACMAT*, Part F128644, p. 247–254, 2017.

GURRIN, C.; SMEATON, A. F.; DOHERTY, A. R. LifeLogging: Personal big data. *Foundations and Trends in Information Retrieval*, v. 8, n. 1, p. 1–125, 2014. ISSN 15540677.

HADDADI, H. et al. Human-data interaction. *University of Cambridge*, n. 837, p. 1–9, 2013. Disponível em: <<http://128.232.0.20/techreports/UCAM-CL-TR-837.pdf>> 5Cnpapers2://publication/uuid/20A8C3B6-3820-4ECB-A8C2-402BC7EB7899 5Cnhttp://www-ipv4.cl.cam.ac.uk/techreports/UCAM-CL-TR-8>.

HEDBOM, H. A Survey on Transparency Tools for Enhancing Privacy. p. 67–82, 2009. ISSN 1868-4238. Disponível em: <<http://link.springer.com/10.1007/978-3-642-03315-5>>.

HOLTZ, L. E.; NOCUN, K.; HANSEN, M. Towards displaying privacy information with icons. *IFIP Advances in Information and Communication Technology*, v. 352 AICT, p. 338–348, 2011. ISSN 18684238.

HOSSEINI, M. et al. Foundations for Transparency Requirements Engineering. In: . [s.n.], 2016. p. 225–231. Disponível em: <<http://link.springer.com/10.1007/978-3-319-30282-9>>.

HSIEH, O. Human computer interaction and data visualization. *Advanced Writing: Pop Culture Intersections*, p. 1–24, 2016. Disponível em: <<http://scholarcommons.scu.edu/engl/176>>.

HUANG, D. et al. Personal visualization and personal visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, v. 21, n. 3, p. 420–433, 2015. ISSN 10772626.

IACHELLO, G.; HONG, J. End-user Privacy in Human-Computer Interaction. *Now. The essence of knowledge, Foundations and Trends in Human Computer Interaction*, v. 1, p. 1 – 137, 2007.

ILLINSKY, N.; STEELE, J. *Design Data Visualizations*. [S.l.]: OReilly Media Inc., 2011. ISBN 9781449312282.

ISMAIL, S.; SHAIKH, T. A Literature Review on Semantic Web - Understanding the Pioneers' Perspective. *ICCSEA, SPPR, UBIC*, n. October, p. 15–28, 2016.

JANIC, M.; WIJBENGA, J. P.; VEUGEN, T. Transparency enhancing tools (TETs): An overview. *Workshop on Socio-Technical Aspects in Security and Trust, STAST*, n. October 2014, p. 18–25, 2013. ISSN 23251689.

- KANDARI, J. et al. Information quality on the World Wide Web: Development of a framework. *International Journal of Information Quality*, v. 2, n. 4, p. 324–343, 2011. ISSN 17510457.
- KANDARI, J. et al. Information quality on the World Wide Web: Development of a framework. *International Journal of Information Quality*, v. 2, n. 4, p. 324–343, 2011. ISSN 17510457.
- KIRRANE, S. et al. A scalable consent, transparency and compliance architecture. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 11155 LNCS, n. i, p. 131–136, 2018. ISSN 16113349.
- KITCHENHAM, B. *Procedures for Performing Systematic Reviews*. [S.l.], 2004. 1–28 p.
- KUMAR, S.; JAKHAR, M. Understanding user evaluation of Information Quality Dimensions in a digitized world. *Proceedings of Production and Operations Management Society*, v. 1, p. 1–10, 2010.
- KUNUNKA, S. et al. End User Comprehension of Privacy Policy Representations. In: . [s.n.], 2017. p. 135–149. Disponível em: <<http://link.springer.com/10.1007/978-3-319-58735-6{\}-}>>
- LAW, D. *Proteção De Dados Do Brasil Lgpd Conhecendo a Lei De Dados Do Brasil*. [S.l.: s.n.], 2020. 20 p.
- LEBO, M. S.; SUTTI, S.; GREEN, R. C. "Big data" gets personal. *Science Translational Medicine*, v. 8, n. 322, p. 322fs3–322fs3, 2016. ISSN 1946-6234. Disponível em: <<http://stm.sciencemag.org/cgi/doi/10.1126/scitranslmed.aad9460>>.
- LEE, Y. W. et al. AIMQ: A methodology for information quality assessment. *Information and Management*, v. 40, n. 2, p. 133–146, 2002. ISSN 03787206.
- MALTA, M. C.; BAPTISTA, A. A. A panoramic view on metadata application profiles of the last decade. *International Journal of Metadata, Semantics and Ontologies*, v. 9, n. 1, p. 58–73, 2014.
- MASHHADI, A.; KAWSAR, F.; ACER, U. G. Human Data Interaction in IoT: The ownership aspect. In: *2014 IEEE World Forum on Internet of Things, WF-IoT 2014*. [S.l.: s.n.], 2014. ISBN VO -.
- MAUS, G. Decoding, hacking, and optimizing societies: Exploring potential applications of human data analytics in sociological engineering, both internally and as offensive weapons. *Proceedings of the 2015 Science and Information Conference, SAI 2015*, p. 538–547, 2015.
- MCAULEY, D.; MORTIER, R.; GOULDING, J. The Dataware manifesto. *2011 3rd International Conference on Communication Systems and Networks, COMSNETS 2011*, 2011.
- MCKILLIP, J. L.; JAYKUS, L. A.; DRAKE, M. rRNA stability in heat-killed and UV-irradiated enterotoxigenic *Staphylococcus aureus* and *Escherichia coli* O157:H7. *Applied and Environmental Microbiology*, v. 64, n. 11, p. 4264–4268, 1998. ISSN 00992240.

MEW, K. *Aprendendo Material Design: Domine o Material Design e crie interfaces bonitas e animadas para aplicativos móveis e web.* São Paulo: Novatec Editora, 2016. 200 p. ISBN 9788575225127.

MICHAEL, K. et al. Monitoring People using Location-Based Social Networking and its Negative Impact on Trust : An Exploratory Contextual Analysis of Five Types of “Friend ” Relationships Monitoring People using Location-Based Social Networking and its Negative Impact on Tr. n. May, 2011.

MILNE, G. R.; CULNAN, M. J. Strategies for reducing online privacy risks: Why consumers read (or don’t read) online privacy notices. *Journal of Interactive Marketing*, v. 18, n. 3, p. 15 – 29, 2004. ISSN 1094-9968. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1094996804701085>>.

MONTI, M. Projeto de Lei 4060/12 - Tratamento de dados pessoais. *Câmara dos Deputados*, v. 1, p. 1–46, 2014.

MORAN, G. Understanding Metadata Concepts and Properties. *wiki.pentaho.com Metadata*, p. 11–13, 2008.

MORGANA, A.; BAPTISTA, A. A. The Use of Application Profiles and Metadata Schemas by Digital Repositories : Findings from a Survey. *International Conference on Dublin Core and Metadata Applications*, n. Dcmi, p. 146–157, 2015. ISSN 19391366.

MORI, J. et al. Keyword extraction from the Web for personal metadata annotation. *CEUR Workshop Proceedings*, v. 184, n. May, p. 51–60, 2004. ISSN 16130073.

MORTIER, R. et al. Human-Data Interaction: The Encyclopedia of Human-Computer Interaction. *The Encyclopedia of Human-Computer Interaction*, p. 1–48, 2016. Disponível em: <<https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed>>.

MURMANN, P.; FISCHER-HÜBNER, S. Tools for Achieving Usable Ex Post Transparency: A Survey. *IEEE Access*, v. 5, p. 22965–22991, 2017. ISSN 21693536.

MURMANN, P.; FISCHER-HÜBNER, S. *Usable Transparency Enhancing Tools*. [S.l.], 2017. 43 p. Disponível em: <<http://www.diva-portal.org/smash/get/diva2:1119515/FULLTEXT02.pdf>>.

MURPHY, J. F. The General Data Protection Regulation (GDPR). *Irish Medical Journal*, v. 111, n. 5, p. 747, 2018. ISSN 03323102.

NEIL, T. *Mobile Design Pattern Gallery*. 2. ed. [S.l.]: OReilly, 2014. 299 p. ISBN 0636920029311.

PATRICK, A. S. From Privacy Legislation to Interface Design : Implementing Information Privacy in Human- Computer Interactions Conference Paper in Lecture Notes in Computer Science · March 2003 NRC Publications Archive (NPArc) From Privacy Legislation to Interface Des. n. January, 2015.

PATRICK, A. S.; KENNY, S. From Privacy Legislation to Interface Design: Implementing Information Privacy in Human-Computer Interactions. p. 107–124, 2003. ISSN 03029743. Disponível em: <<http://link.springer.com/10.1007/978-3-540-40956-4{\}>>

- Pioneiro Neto, A. *Manual ABA para adequação à LGPD Orientações e boas práticas.* [S.l.: s.n.], 2020. 24 p.
- PIPINO, L. L.; LEE, Y. W.; WANG, R. Y. Data Quality Assessment. *Communications of the ACM*, v. 45, n. 4, p. 211–218, 2002. ISSN 00320838.
- PRESSMAN, R. S. *Software Engineering – A Practitional Approach.* 8. ed. [S.l.]: Mc-Graw Hill,, 2014.
- ROBOL, M.; SALNITRI, M.; GIORGINI, P. Toward GDPR-Compliant Socio-Technical Systems: Modeling Language and Reasoning Framework. *Lecture Notes in Business Information Processing*, v. 10, p. 236–250, 2017.
- ROMANSKY, R. Social media and personal data protection. *International Journal on Information Technologies & Security*, v. 4, n. September, p. 65–79, 2014.
- SALIH, D. Natural User Interfaces — Breakling.de. *LM Research Topics in HCI*, p. 1–7, 2010. Disponível em: <<http://www.braekling.de/usability/2906-natural-user-interfaces.html>>.
- SAMPSON, D. G.; ZERVAS, P.; CHLOROS, G. Supporting the process of developing and managing LOM application profiles: The ASK-LOM-AP tool. *IEEE Transactions on Learning Technologies*, IEEE, v. 5, n. 3, p. 238–250, 2012. ISSN 19391382.
- SCHNEIER, B. *Data and Goliath. The Hidden Battles to Collect Your Data and Control Your World.* [S.l.]: W. W. Norton & Company, 2015. 320 p. ISBN 978-0393244816.
- SPAGNUELLO, D.; BARTOLINI, C.; LENZINI, G. Modelling Metrics for Transparency in Medical Systems. In: . [s.n.], 2017. p. 81–95. Disponível em: <<http://link.springer.com/10.1007/978-3-319-64483-7{\}>>.
- STAIR, R. M.; REYNOLDS, G. W. *Princípios de Sistemas de Informação.* 3. ed. [S.l.]: Cengage Learningl, 2015. 752 p. ISBN 9788522118625.
- STVILIA, B. et al. A framework for information quality assessment. *Journal of the American Society for Information Science and Technology*, v. 58, n. 12, p. 1720–1733, 2007. ISSN 15322882.
- TENNIS, J. T. Metadata Application Profiles. In: *Encyclopedia of Archival Concepts, Principles, and Practices.* [S.l.]: Rowman & Littlefield, 2015. p. 2–5.
- TOLEDO, M. D. E. *Lei Geral de Proteção de Dados. um guia completo.* [S.l.]: Emprendedorismo Legal, 2020. 32 p.
- TOM, J.; SING, E.; MATULEVIČIUS, R. Conceptual representation of the GDPR: Model and application directions. *Lecture Notes in Business Information Processing*, v. 330, n. January, p. 18–28, 2018. ISSN 18651348.
- TURILLI, M.; FLORIDI, L. The ethics of information transparency. *Ethics and Information Technology*, v. 11, n. 2, p. 105–112, 2009. ISSN 13881957.
- UKLON. *Scholarly Works Application Profile.* 2009. Disponível em: <<http://www.ukoln.ac.uk/repositories/digirep/index/Scholarly{\}Works{\}Application{\}>>.

UNION, E. *General Data Protection Regulation*. 2018. Disponível em: <<https://gdpr-info.eu/>>.

VOIGT, P.; BUSSCHE, A. v. d. *The EU General Data Protection Regulation (GDPR): A Practical Guide*. 1st. ed. [S.l.]: Springer Publishing Company, Incorporated, 2017. ISBN 3319579584, 9783319579580.

WANG, R. Y.; STRONG, D. M. Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, v. 12, n. 4, p. 5–33, 1996. ISSN 11217588.

WATSON, T. *Getting GDPR Consent & Opt-in*. 2016. Disponível em: <<https://www.zettasphere.com/gdpr-consent-opt-in-examples/>>.

APÊNDICE A – PROTOCOLO DE REVISÃO SISTEMÁTICA

A.1 Objetivo

Identificar trabalhos que especifiquem, de forma direta ou indireta, quais informações sobre o uso dos dados pessoais compõem a Transparência de Dados Pessoais em aplicações de software, e como a Transparência pode ser apresentada (*design* de interface) ao indivíduo. Entende-se por forma direta, artigos que propuseram um conjunto de informações para Transparência; por forma indireta, deve-se compreender artigos que não propuseram dados para Transparência, mas consideraram, e elencaram, algum conjunto de informações.

A.2 Pergunta de Pesquisa

A Revisão Sistemática (RS) foi composta pelas questões: (1) quais informações sobre o uso dos dados pessoais são utilizados para compor a Transparência de Dados Pessoais? e (2) quais técnicas são utilizadas para apresentar as informações de Transparência para os indivíduos?

A.3 PICOC

População (P): Trabalhos que apresentavam quais informações devem contemplar a Transparência de Dados Pessoais; e/ou trabalhos sobre técnicas de apresentação de Transparência para o indivíduo.

Intervenção (I): Foram observados os Metadados de Transparência, a descrições dos metadados e as técnicas de apresentação de informações.

Controle (C): Não aplicado.

Outcomes (O): Listagem das informações de Transparência e técnicas de apresentação das informações.

Contexto (C): Transparência de Dados Pessoais e Privacidade de Dados.

A.4 Período de publicação considerado para os artigos

Foram considerados artigos publicados entre 2008 e 2019. Esse período foi selecionado com base na pesquisa exploratória que mostrou que possíveis resultados para a questão de pesquisa estariam em trabalhos mais recentes em relação ao período da execução da RS.

A.5 Keywords

As palavras chave utilizadas foram:

- Data Legibility;
- Data Transparency;
- Personal data disclose;
- Personal data privacy;
- Personal data processing;
- Personal data Transparency;
- Personal data trust;
- Personal data use;
- Personal data using;
- Peronsal Transparency processing;
- Use of personal data;
- Using of personal data;

- Model;
- Metadata;
- Strategy; e
- Approach.

A.6 Critérios de seleção de base de dados

Bases de dados indexadas e *Snowball*. A técnica de *snowball* foi utilizada quando identificadas referências bibliográficas relevantes nos trabalhos selecionados e que deveriam integrar a RS. As referências identificadas no *snowball* que atenderam os critérios de inclusão foram inseridas manualmente na RS.

A.7 Idiomas

Inglês e Português.

A.8 Fontes de Dados e Métodos de busca e fontes de dados

A Revisão Sistemática foi conduzida nas bases de dados indexadas:

Pesquisa com *String* de busca utilizando pesquisa avançada.

- IEEE;
- ACM;
- Web of Science;
- Springer.

O método de busca utilizado foi o de *String* de busca com o uso da *Advanced Search* disponível nos *websites* das bases. As *Strings* utilizadas nas bases de dados são mostradas na Tabela 23.

Tabela 23: *Strings* de busca para as bases de dados. Do autor.

Base de Dados	String
IEEE	((("personal data use"OR "personal data using"OR "use of personal data"OR "using personal data"OR "data legibility"OR "data transparency"OR "personal data privacy"OR "personal data disclose"OR "personal data trust"OR "personal transparency processing"OR "personal data processing") AND (model OR strategy OR approach OR metadata)))
ACM	+("personal data usepersonal data usinguse of personal datausing personal datapersonal data transparencydata legibilitydata transparency personal data privacypersonal data disclosepersonal data trustpersonal transparency processingpersonal data processing") +("modelstrategyapproachmetadata")
Web of Science	ALL=((("personal data use"OR "personal data using"OR "use of personal data"OR "using personal data"OR "data legibility"OR "data transparency"OR "personal data privacy"OR "personal data disclose"OR "personal data trust"OR "personal transparency processing"OR "personal data processing") AND (model OR strategy OR approach OR metadata)))
Springer	((("personal data use"OR "personal data using"OR "use of personal data"OR "using personal data"OR "data legibility"OR "data transparency"OR "personal data privacy"OR "personal data disclose"OR "personal data trust"OR "personal transparency processing"OR "personal data processing") AND (model OR strategy OR approach OR metadata)))

A.9 Critérios de Inclusão e Exclusão

Os critérios de inclusão e exclusão são mostrados na Tabela 24.

A.10 Forma de Análise

Tabulação de Dados e Análise Qualitativa.

A.11 Critério de qualidade

- Estudos publicados em eventos ou revistas científicas;
- Relatórios técnicos publicados por instituições de pesquisa.

Tabela 24: Critérios de Inclusão e Exclusão da Revisão Sistemática. Do autor.

Critério	Descrição
Inclusão	Trabalhos que apresentavam, direta ou indiretamente, informações de como compor a Transparência de Dados Pessoais
Inclusão	Trabalhos que discutiam técnicas de apresentação da Transparência de Dados Pessoais para os indivíduos
Exclusão	Artigos que tratavam de Transparência, mas sem considerar dados pessoais
Exclusão	Artigos que tratavam de visualização de dados por indivíduo, mas fora do contexto de Transparência.;

A.12 Extração de Conteúdos

A extração de conteúdos dos trabalhos foi executada com as atividades:

- Extração direta de trabalhos que especificavam claramente as informações de Transparência;
- Extração indireta de trabalhos que não especificavam informações de Transparência, mas que adotaram um conjunto das mesmas para amparar a pesquisa;
- Extração das formas de apresentação das informações de Transparência.

A.13 Síntese dos Dados

Os resultados foram tabulados descrevendo quais informações de Transparência cada trabalho abordava; também foram descritas as formas de apresentação da Transparência; foi realizada, também, uma discussão relacionada ao conhecimento obtido pelo pesquisador desse trabalho. A discussão buscou amparar a resolução das perguntas dessa RS.

APÊNDICE B – TABELA DE EXTRAÇÃO DE DADOS DA REVISÃO SISTEMÁTICA

Tabela 25: Tabulação da Extração de Dados da RS. Do autor.

Nº	Informações de Transparência	Forma de Apresentação
1	(1) Dados Pessoais; (2) Processos realizados com os dados pessoais; (3) Propósito de Uso; (4) Descrição de como e onde os dados são armazenados; (5) Destinatário(s) dos dados pessoais.	Os autores não apresentam meios de apresentação voltados para os indivíduos. O trabalho é voltado para a proposta de um vocabulário em RDF para implementação de níveis de granularidade para os dados pessoais; e para melhorar comunicação entre serviços financeiros que utilizam dados pessoais.
2	(1) Informações sobre pessoas autorizadas a acessar aos dados pessoais sobre saúde.	Não são discutidos padrões para apresentação das informações.

3	<p>O trabalho apresenta uma modelagem UML contendo classes, atributos e relacionamentos para modelar a GDPR. Embora a Transparência em si, não tenha sido modelada, suas entidades e atributos poderiam ser utilizados para dar suporte à mesma. A entidades e atributos são mostrados na Figura 35 no Anexo B.</p>	<p>O artigo aborda, especificamente, a modelagem dos elementos da GDPR e não apresenta qualquer padrão de <i>design</i> para Transparência.</p>
4	<p>(1) Informações sobre os agentes envolvidos no uso dos dados pessoais; (2) Informações sobre quais dados pessoais são utilizados e seus respectivos conteúdos; (3) Informações relacionadas aos processos e interações são conduzidas sobre os dados pessoais; e (4) Informações sobre objetivos de uso, políticas de privacidade e decisões tomadas ou informações produzidas com os dados pessoais.</p>	<p>O artigo propõe quatro conjuntos de dados, chamados de facetas, de requisitos de Transparência e não apresenta nenhum <i>design pattern</i>.</p>
5	<p>(1) Objetivo de uso; (2) Cadeia de processamento dos dados que envolvam vários processadores; (3) Informações de transferência e compartilhamento dos dados; (4) Contatos e obrigações dos utilizadores dos dados; (5) Localização dos agentes que não estejam na União Europeia; e (6) Leis e direitos dos indivíduos.</p>	<p>Ícones utilizados para destacar informações em PPS podem ser utilizados para a <i>Ex ante Transparency</i>. Já para a <i>Ex post Transparency</i>, os autores propõem o uso da ferramenta <i>Data Track</i>, que disponibiliza uma representação visual para os diversos caminhos os quais os dados podem ser conduzidos. Cada nó do <i>Data Track</i> armazena informações sobre o uso ou compartilhamento dos dados pessoais, tais como agentes, objetivos de uso e informações de compartilhamento.</p>

6	<p>Apresenta um conjunto de métricas de avaliação de Transparência. Entretanto, no que tange aos dados pessoais apresenta os seguintes: (1) Informações sobre quem manipula os dados pessoais; (2) Informações sobre como os dados são protegidos; (3) Informações sobre uso incorreto dos dados pessoais.</p>	<p>O artigo é focado em métricas de auditoria e controle de uso de dados e não apresenta contribuições relacionadas ao formato de apresentação das informações.</p>
7	<p>(1) Bases de dados compartilhadas; (2) Pseudônimos de transações; (3) Informações de consentimento de uso dos dados; (4) Localização atual dos dados pessoais.</p>	<p>Forma tabular, para apresentação de dados simples; <i>Tracer View</i> e <i>Timeline</i>, para auxiliar no rastreamento do uso dos dados pessoais; e tutoriais para esclarecer dúvidas específicas.</p>
8	<p>(1) Dados de interação com aplicações e <i>websites</i> como: <i>strings</i> de busca, endereços de <i>websites</i>, <i>sistema operacional utilizado</i> e tempo das atividades nas aplicações; (2) Dados sobre o indivíduo como: idade, gênero, gosto musical, preferências, orientação religiosa, telefone e endereço; e (3) Dados de objetos e ações relacionados ao indivíduo, tais como: número do cartão de crédito, E-mail, Afiliações em grupos, rotinas, percursos, localização e número de passaporte.</p>	<p>O artigo não apresenta padrões de <i>design</i> para implementação de Transparência, pois o mesmo tem forte aspecto jurídico e foca-se em apresentar maneiras de classificar os dados pessoais de acordo com determinadas regulamentações existentes.</p>
9	<p>(1) Informações sobre quais dados pessoais são armazenados; (2) A origem dos dados e quem vai utilizá-los; (3) Informações sobre compartilhamento dos dados; (4) Lógica envolvida no processamento dos dados pessoais; e (5) O propósito de uso.</p>	<p>Os autores apresentaram um conjunto de requisitos para um <i>dashboard</i> de Transparência que foram baseados na ISO 9241-11 com um Modelagem Formal para Proveniência dos Dados. A implementação prática foi feita com o uso de <i>web design</i> com HTML.</p>

10	<p>Os autores apresentam um conjunto de categorias que organizam os dados pessoais em: (1) Classes: dados sobre os indivíduos, tais como etnia, gênero, religião e preferências políticas; (2) Funções legais, que referem-se às ações conduzidas nos dados pessoais; e (3) Notações auxiliares, que referem-se às informações sobre propósito de uso, consentimento e uso e qualidade de informação</p>	<p>O artigo trata as categorias de Transparência em notações formais e não aborda nenhuma forma de apresentação para indivíduos de pouco letramento.</p>
11	<p>Os autores não propõem metadados de Transparência, mas analisam TETs que disponibilizam um conjunto informações como: (1) Localização de agentes envolvidos no uso dos dados; (2) Rastreamento do compartilhamento dos dados; (3) Informações explícitas dos dados pessoais; (4) Informações implícitas tais como endereço IP ou tempo de uma transação; (5) Dados derivados, tais como perfil de usuários ou preferências; e (5) Dados de conclusão, que representam algum conhecimento obtido sobre o indivíduo.</p>	<p>As TETs analisadas apresentaram formatos de apresentação da Transparência tais como: codificação de cores, para destacar e/ou dar ênfase para algum evento específico no uso dos dados pessoais; árvores de dados, para mostrar a hierarquia de dados pessoais e elementos envolvidos na utilização; marcadores de localização, para exibir locais de agentes que utilizam os dados pessoais; e mapas de informações e linhas de tempo, para mostrar a relação e temporização do uso dos dados pessoais.</p>
12	<p>As informações de Transparência são apresentadas como parte dos Princípios de Privacidade e contemplam: (1) Quem processa os dados pessoais; (2) Qual o propósito do uso dos dados pessoais; (3) Quais processos são conduzidos nos dados pessoais.</p>	<p>Para o princípio de Transparência, os autores destacam que a Transparência deve ser clara e proporcionar condições para o indivíduo entender e controlar o uso de seu dados pessoais. A apresentação das informações poderiam ser realizada por meio exemplos e tutoriais.</p>

13	<p>O artigo apresenta uma RS na qual os artigos analisados contemplam informações de Transparência como: (1) Localização de agentes envolvidos no uso dos dados pessoais; (2) <i>Strings</i> utilizadas em <i>websites</i> de busca; (3) Dados da rede do indivíduo; (4) Históricos de Navegação; (5) <i>Logs</i> de navegação e tarefas dos indivíduos; (6) Dados pessoais compartilhados; (8) quais dados foram obtidos de sensores como celulares e câmeras.</p>	<p>Não são apresentadas propostas de padrões de <i>design</i>, mas trabalhos analisados sugerem o uso de modos de visualização como mapas (informações de localização), gráficos (para auditoria); e árvores (históricos e logs de navegação).</p>
14	<p>Os autores apresentam informações de Transparência como características necessárias das TETs analisadas. As informações disponibilizadas pelas TETs contemplaram: (1) Dados pessoais utilizados; (2) Como os dados pessoais são armazenados; (3) Como os dados pessoais são processados; (4) Informações sobre o compartilhamento/divulgação dos dados pessoais; (5) Quem acesso o navegador; (6) <i>Compliance</i> do <i>website</i> com as políticas de privacidade; (7) Dados coletados do navegador; (8) Fluxo dos dados pessoais na rede; e (9) Informações demográficas dos agentes que acessam os dados pessoais.</p>	<p>As TETs analisadas disponibilizam mecanismos distintos de apresentação de suas informações de Transparência. Foram apontados como meios de apresentação da informação: (1) Uso de ícones desenvolvidos, para dar apoio à análise de Privacidade dos Dados; (2) Uso de cores e formatos, para destacar ou intensificar a importância de uma informação; (3) <i>Dashboard</i> com textos e gráficos; (4) Listas de itens, para informações diversas; e (5) Mapas com informações de localização.</p>
15	<p>Não foi objetivo do trabalho propor metadados de Transparência, mas com a apresentação dos ícones, os autores relacionaram informações de Transparência tais como: (1) Dados pessoais utilizados; (2) Propósito de uso; (3) Obrigações legais; e (4) Indivíduos envolvidos no uso dos dados pessoais.</p>	<p>Os autores propõem a apresentação da Transparência com base em ícones desenvolvidos especificamente para questões de privacidade de dados pessoais.</p>

ANEXO A – DIRETRIZES DE TRANSPARÊNCIA DE DADOS PESSOAIS DA GDPR TRADUZIDAS PARA O PORTUGUÊS.

Tabela 26: Diretrizes da GDPR para Transparência de Dados Pessoais. Fonte: <https://gdpr-info.eu/>.

Artigo/Secão/ Alínea	Descrição
13/1/a e 14/1/a	A identidade e os detalhes de contato do controlador e, quando aplicável, do representante do controlador.
13/1/b e 14/1/b	Os dados de contato da entidade responsável pela proteção de dados, se aplicável.
13/1/c e 14/1/c	Os objetivos do tratamento a que os dados pessoais se destinam, bem como a base jurídica do tratamento.
14/1/d.	As categorias de dados pessoais utilizados.
13/1/d e 14/2/b	Se o tratamento se basear no artigo 6(1), alínea f, os interesses legítimos prosseguidos pelo responsável pelo tratamento ou por terceiros.
13/1/f e 14/1/f	Quando aplicável, que o responsável pelo tratamento tem por objetivo transferir dados pessoais para um destinatário de um país terceiro ou organização internacional verificar a existência ou ausência de uma regulamentação compatível.

13/2/a e 14/2/a	O período durante o qual os dados pessoais serão armazenados ou, se isso não for possível, os critérios utilizados para determinar esse período.
13/2/b e 14/2/c	A existência do direito de solicitar ao controlador acesso e retificação ou exclusão de dados pessoais ou restrição de processamento referente ao titular dos dados ou de objeções ao processamento, bem como o direito à portabilidade dos dados.
13/2/e	Se o fornecimento de dados pessoais é um requisito estatutário ou contratual ou um requisito necessário para celebrar um contrato, bem como se o titular dos dados é obrigado a fornecer os dados pessoais e as possíveis consequências da falta de fornecimento desses dados.
13/2/f e 14/2/g	Se existir uma tomada de decisão automatizada, incluindo a criação de perfil, referida nas seções 1 e 4 do artigo 22 e, pelo menos nesses casos, deve ser fornecidas informações significativas sobre a lógica envolvida, bem como a importância e as consequências previstas desse tratamento para o titular dos dados.
14/2/f	De qual fonte os dados pessoais se originam e, se aplicável, se vieram de fontes acessíveis ao público.

ANEXO B – MODELAGEM DE CLASSES E ATRIBUTOS PARA A GDPR

Este Apêndice apresenta a modelagem de classes, atributos e relacionamentos apresentados por Tom, Sing e Matulevičius (2018) a fim de prover o estado da arte das entidades e atributos da GDPR para implementação de uso dos dados pessoais de acordo com a regulamentação. A modelagem é mostrada na Figura 35.

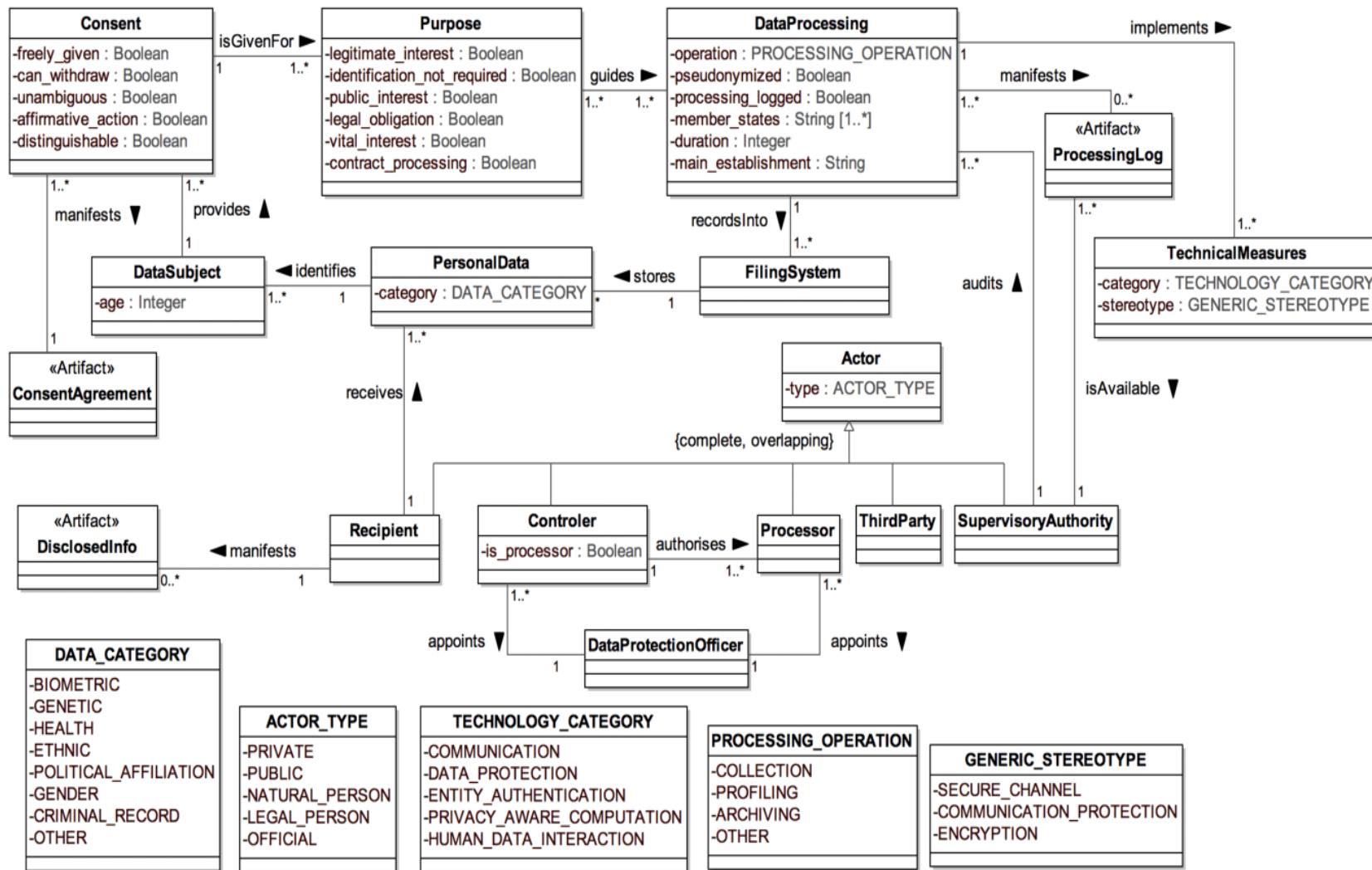


Figura 35: Modelagem UML para a GDPR. Extraído de Tom, Sing e Matulevičius (2018).