



UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO  
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Transparência de Informações Prosopográficas: modelo de dimensões para apoio às  
necessidades informacionais do cidadão sobre os políticos.

José Soares Barroso Júnior

**Orientadores**

Até 15 de Agosto de 2018: Dra. Claudia Cappelli

A partir de 15 de Agosto de 2018: Dr. Mariano Gomes Pimentel

**Co-orientadora**

Dra. Vanessa Tavares Nunes

RIO DE JANEIRO, RJ - BRASIL

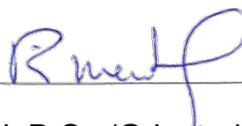
FEVEREIRO DE 2020

Transparência de Informações Prosopográficas: modelo de dimensões para apoio às necessidades informacionais do cidadão sobre os políticos.

JOSÉ SOARES BARROSO JÚNIOR

DISSERTAÇÃO APRESENTADA COMO REQUISITO PARCIAL PARA OBTENÇÃO DO TÍTULO DE MESTRE PELO PROGRAMA DE PÓS-GRADUAÇÃO EM INFOR-MÁTICA DA UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO (UNI-RIO). APROVADA PELA COMISSÃO EXAMINADORA ABAIXO ASSINADA.

Aprovada por:



Mariano Gomes Pimentel, D.Sc (Orientador) — (UNIRIO)



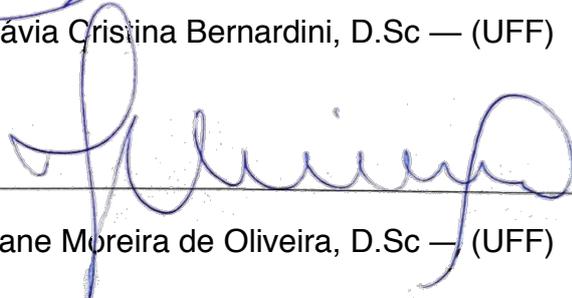
Vanessa Tavares Nunes, D.Sc (Co-orientadora) — (UnB)



Bruna Diirr Gonçalves da Silva, D.Sc — (UNIRIO)



Flávia Cristina Bernardini, D.Sc — (UFF)



Thaianne Moreira de Oliveira, D.Sc — (UFF)

RIO DE JANEIRO, RJ - BRASIL

FEVEREIRO DE 2020

Catálogo informatizada pelo(a) autor(a)

B95	<p>Barroso Júnior, José Soares</p> <p>Transparência de informações prosopográficas: modelo de dimensões para apoio às necessidades informacionais do cidadão sobre os políticos. / José Soares Barroso Júnior. -- Rio de Janeiro, 2020. 208 f.</p> <p>Orientador: Mariano Gomes Pimentel. Coorientadora: Vanessa Tavares Nunes. Dissertação (Mestrado) - Universidade Federal do Estado do Rio de Janeiro, Programa de Pós-Graduação em Informática, 2020.</p> <p>1. Informação prosopográfica. 2. Transparência. 3. Modelagem dimensional. 4. Data warehouse. I. Pimentel, Mariano Gomes, orient. II. Nunes, Vanessa Tavares, coorient. III. Título.</p>
-----	---

*Dedico este trabalho à Deus, meus pais e  
minhas irmãs.*

## **Agradecimentos**

Chegar até aqui foi realmente desafiador. Tão importante quanto a conclusão da pesquisa, foi a trajetória. Durante esse tempo eu descobri coisas que não enxergava em mim. Encarei alguns desafios e os superei, aprendi muito.

Lembro-me como se fosse hoje, a primeira aula foi de Fundamentos de Sistemas de Informação, e lá estava ela, a Profª Renata Araújo, que nos recebeu com muita elegância e atenção. Em sua fala, ela dizia que estávamos iniciando uma jornada de muita leitura, muita fala, muita escrita, exposição e muitas conexões. Muito grato por esta disciplina, muito grato por me fazer pensar "fora da caixa", desconstruir meus pensamentos tecnicistas e ser provocado a todo instante.

Foram momentos de muitas inquietações. Cada aula, cada revisão, cada submissão era uma comemoração, sim, porque um dos meus desejos ao entrar no mestrado foi aprender a escrever e publicar um artigo científico. A cada reunião, a cada travessia da ponte voltando pra casa, eram momentos de ressignificações. Ao recordar de onde eu vim, pelas restrições que eu tive em morar na zona rural, eu me senti grato por cada conhecimento compartilhado. Coisa que me move. Internamente me sinto a todo momento forjado a desafiar, a observar e a compartilhar, e pra mim, isso é transformador.

Por isso eu quero agradecer todas as minhas inspirações enquanto aprendiz-pesquisador, em especial aos meus orientadores: Claudia Cappeli, Vanessa Nunes, Mariano Pimentel, à vocês minha admiração. Obrigado pela paciência. Foram muitos momentos de conquistas, momentos de tristeza, de tensão na entrega dos artigos, vocês foram verdadeiros parceiros.

Gratidão aos demais professores com quem tive a oportunidade de conhecer e aprender: Kate Revoredo, Cristina Bicharra, Sean Siqueira, Gleison Santos, Fernanda Baião e Cristina Cerdeiral. Agradeço também aos meus colegas de turma, à minha amiga-irmã Tatiana Barboza e aos colegas de trabalho, em especial à Rozana Moreira.

Um abraço forte de agradecimento aos meus amigos que entenderam as minhas constantes ausências e a minha família, que é meu porto seguro. A todos vocês muito obrigado, por fazer parte dessa caminhada.

BARROSO JUNIOR, JOSÉ SOARES **Transparência de Informações Prosopográficas: modelo de dimensões para apoio às necessidades informacionais do cidadão sobre os políticos.** UNIRIO, 2020. 208 páginas. Dissertação de Mestrado. Departamento de Informática Aplicada, UNIRIO.

## RESUMO

A crescente demanda por informações sobre políticos e a velocidade com que as notícias são divulgadas pela mídia e pelas redes sociais revelam uma visão mais participativa do cidadão nas ações políticas e governamentais, principalmente nos períodos eleitorais. Desse modo, a transparência da informação passa a ser primordial para um regime democrático, através da organização e consolidação de dados provenientes de fontes oficiais e da sua materialização em ferramentas para consulta e disseminação de forma confiável. Durante o período das eleições presidenciais no Brasil em 2018, este trabalho procurou entender alguns modelos cognitivos que explicam o comportamento eleitoral, o que levou à investigação de fatores que influenciam a escolha do voto, aprovação ou desaprovação de alguns objetivos. Tais fatores foram fundamentais para identificar os dados de maior interesse a serem investigados e posteriormente visualizados pelo cidadão. Desta forma, esta pesquisa é direcionada para a organização de dados de natureza prosopográfica, pois eles tratam da biografia dos políticos e de suas relações com o coletivo. Os dados aqui tratados são oriundos de 3 fontes distintas. A primeira foi o "Portal Excelências", mantido pela organização não governamental "Transparência Brasil", que concentra as informações contidas nas bases de dados oficiais das casas legislativas brasileiras. A segunda foi a base de dados do Departamento de Ciências da Computação da Universidade Federal de Sergipe (UFS) e a terceira, o portal de dados abertos do Tribunal Superior Eleitoral (TSE). No entanto, os dados contidos nessas plataformas não são adequadamente modelados e dificultam a decisão analítica, comprometendo a transparência das informações. Nesse sentido e utilizando a abordagem DSR (Design Science Research), depois de entender o que é informação de interesse prosopográfico, esta pesquisa propôs a construção de um modelo multidimensional de dados e a consolidação e disponibilização dos mesmos em um sistema de consulta de informações sobre os políticos. A técnica de modelagem di-

mensional foi utilizada para processar o grande volume de dados coletados de forma mais rápida e permitir o detalhamento das informações para melhorar a compreensão do cidadão. O sistema de consulta foi desenvolvido de modo a levar informações sobre os fatos relacionados ao político. O objetivo é fomentar a transparência de informações para que, de maneira mais confiável, permita que o cidadão possa acompanhar a vida de um candidato, sua história e suas experiências como político.

**Palavras-chave:** Informação Prosopográfica, Transparência, Modelagem Dimensional, Data warehouse

## ABSTRACT

The increase in demand for information about politicians and the speed with which news are spread by the media and social networks reveal a more participative view of the citizen in political and governmental actions, especially in the election periods. So, the transparency of information becomes essential for a democratic regime, through the organization and consolidation of data from official sources and its materialization into tools for reliable consultation and dissemination. During the period of the presidential elections in Brazil in 2018, this work sought to understand some cognitive models that explain electoral behavior, which led to the investigation of factors that influence the choice of vote, approval or disapproval of some objectives. These factors were fundamental to identify the data of greatest interest to be investigated and subsequently viewed by the citizen. Thus, this research is directed to the organization of data of a prosopographic nature, because they deal with the biography of politicians and their relations with the collective. The data treated here comes from 3 different sources. The first was the website "Portal Excelências", maintained by the non-governmental organization "Transparência Brasil", which concentrates the information contained in official databases of the Brazilian legislature houses. The second was the database of the Department of Computer Sciences of the Universidade Federal de Sergipe (UFS) and the third, the open data portal of the Superior Electoral Court "Tribunal Superior Eleitoral"(TSE). However, the data contained in these platforms is not properly modeled and makes the analytical decision difficult, compromising the transparency of the information. In this sense, and using the Design Science Research (DSR) approach, after understanding what is an information of prosopographic interest, this research proposed the construction of a multidimensional data model and its consolidation and availability through a system of information consultation about politicians. The dimensional modeling technique was used to process the large volume of data collected in a faster way and to allow the detailing of the information to improve citizen understanding. The consultation system was designed to provide information about the facts related to the politicians. The goal is to foment the transparency of information so that, more reliably, it allows citizens to follow a candidate's life, history and experiences as a politician.

**Keywords:** Prosopographic Information, Transparency, Dimensional Modeling, Data warehouse

## Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Motivação e Contexto . . . . .	1
1.2	Objetivos Gerais e Específicos . . . . .	3
1.3	Problema, Questão de Pesquisa e Proposta de Solução . . . . .	4
1.4	Ambiente . . . . .	5
1.5	Organização da Dissertação . . . . .	8
<b>2</b>	<b>Abordagem de Pesquisa: <i>Design Science Research</i> (DSR)</b>	<b>9</b>
2.1	Critérios de Avaliação do Artefato . . . . .	15
2.2	Técnica de Codificação da <i>Grounded Theory</i> . . . . .	19
<b>3</b>	<b>Fundamentação Teórica e Técnica</b>	<b>22</b>
3.1	Quadro Teórico . . . . .	22
3.1.1	Informação Prosopográfica . . . . .	22
3.1.2	Transparência de Informações e Transparência de Software . . . . .	24
3.1.3	Transparência de Informações no Contexto Governamental Brasileiro . . . . .	26
3.1.4	Teoria do Voto - Modelos Cognitivos que Explicam o Comportamento Eleitoral . . . . .	32

3.2	Quadro Técnico . . . . .	36
3.2.1	Modelagem Conceitual . . . . .	36
3.2.2	Data Warehouse . . . . .	38
3.2.3	Modelagem Multidimensional . . . . .	42
3.3	Conjecturas Teóricas . . . . .	46
<b>4</b>	<b>Conjecturas comportamentais: Fatores que Influenciam o Voto</b>	<b>48</b>
4.1	Primeira Etapa: levantamento dos fatores de influência do voto . . . . .	50
4.2	Segunda Etapa: avaliação dos fatores . . . . .	53
4.3	Terceira Etapa: avaliação da influência do agrupamento de fatores . . . . .	54
4.4	Relação entre o FIV.Br2018 e os modelos cognitivos . . . . .	60
<b>5</b>	<b>Proposta da Solução: modelo dimensional e o sistema de informações pro-</b> <b>sopográficas</b>	<b>63</b>
5.1	Etapas da Solução Proposta . . . . .	63
5.2	Dicionário de Dados — Base de Dados do Transparência Brasil, Tribu- nal Superior Eleitoral e Universidade Federal de Sergipe . . . . .	64
5.3	Pesquisa sobre os Modelos Cognitivos — Comportamento Eleitoral . . . . .	64
5.4	O Modelo Conceitual de Informações Prosopográficas . . . . .	65
5.5	O Artefato: o modelo multidimensional e o sistema de informação de apoio ao cidadão . . . . .	68
5.5.1	Definição do Escopo . . . . .	69
5.5.2	Prototipação . . . . .	69
5.5.3	Extração e tratamento dos Dados . . . . .	70
5.5.4	Sistema Gerenciador de BD . . . . .	70
<b>6</b>	<b>O Artefato: PROSOPOLÍTICOS</b>	<b>71</b>

6.1	O Modelo Multidimensional: organização dos dados . . . . .	71
6.1.1	Projeção dos Modelos Dimensionais . . . . .	73
6.1.2	Origem dos Dados e Limitações . . . . .	78
6.1.2.1	Base Transparência Brasil: dados da câmara dos deputados e do senado brasileiro . . . . .	78
6.1.2.2	Base UFS: dados sobre os processos judiciais e posicionamento político . . . . .	82
6.1.2.3	Base TSE: repositório de dados abertos . . . . .	85
6.1.3	Processos de Extração e Transformação e Carregamentos dos Dados . . . . .	86
6.2	O Sistema de Informações Prosopolíticos . . . . .	92
<b>7</b>	<b>Avaliação</b>	<b>95</b>
7.1	Os Sujeitos Entrevistados . . . . .	95
7.1.1	Seleção do Contexto . . . . .	95
7.1.2	Observação do uso do sistema e realização das entrevistas . . . . .	96
7.2	Interpretação dos Resultados . . . . .	99
7.2.1	Rigor da Interpretação dos Resultados . . . . .	113
7.2.2	Discutindo os resultados obtidos em comparação com os trabalhos relacionados . . . . .	114
<b>8</b>	<b>Conclusão</b>	<b>116</b>
8.1	Principais Contribuições Científicas e Tecnológicas . . . . .	118
8.2	Limitações da Pesquisa . . . . .	119
8.3	Trabalhos Futuros e Oportunidades . . . . .	120
<b>A</b>	<b>Estudo Exploratório - Questionário Fase 2</b>	<b>134</b>

<b>B</b>	<b>Estudo Exploratório - Questionário Fase 3</b>	<b>148</b>
<b>C</b>	<b>Dicionário de Dados do Esquema Transparência Brasil</b>	<b>158</b>
<b>D</b>	<b>Dicionário de Dados do Esquema da UFS</b>	<b>179</b>
<b>E</b>	<b>Plataforma de Dados Abertos - Layout Arquivo Candidatura</b>	<b>183</b>
<b>F</b>	<b>Relatório de Codificação - Categorias e Grupos</b>	<b>186</b>
<b>G</b>	<b>Implementação da Extração dos Dados do Mandato</b>	<b>189</b>

## Lista de Figuras

1.1	Tela principal do Portal Excelências, (TRANSPARENCIA, 2015) . . . . .	6
2.1	<i>Design Science Research Methodology</i> (DSRM). Fonte: (PIMENTEL; FILIPPE; SANTORO, 2018) . . . . .	13
2.2	Modelo-DSR (PIMENTEL; FILIPPE; SANTORO, 2018) e (PIMENTEL, 2019) . . . . .	16
2.3	Modelo-DSR instanciado de (PIMENTEL; FILIPPE; SANTORO, 2018) e (PIMENTEL, 2019) . . . . .	17
2.4	Modelo TAM traduzido de (DAVIS, 1989) . . . . .	18
3.1	SIG de Transparência [Cappelli and Leite 2009] . . . . .	24
3.2	Operacionalização da Característica de Entendimento, (CAPPELLI, 2009) . . . . .	25
3.3	Ranking de cumprimento à Lei de Acesso a Informação (UNIÃO, 2017) . . . . .	31
3.4	Modelo de Michigan (PERFEITO, 2007) . . . . .	34
3.5	Triângulo de Ullmann (ULLMANN, 1979) . . . . .	37
3.6	Elementos de um <i>Data warehouse</i> (KIMBALL; ROSS, 2011) . . . . .	40
3.7	Exemplo de um cubo dimensional. (ORACLE, 2018) . . . . .	43
3.8	Modelo <i>Star-schema</i> Adaptado de (POE; BROBST; KLAUER, 1997) . . . . .	45
3.9	Modelo <i>Snow-flake</i> . Adaptado de (POE; BROBST; KLAUER, 1997) . . . . .	45
3.10	Modelo Multi-estrela. Adaptado de (POE; BROBST; KLAUER, 1997) . . . . .	45

4.1	Exemplo da interpretação de um depoimento . . . . .	51
4.2	3ª Etapa: questões sobre o agrupamento de fatores . . . . .	55
4.3	Terceira Etapa - Estatísticas dos Fatores de Influência . . . . .	57
4.4	Identificação dos fatores que influenciam a decisão de voto . . . . .	59
5.1	Visão das etapas da solução proposta. . . . .	64
5.2	Modelo Conceitual (representado como um Diagrama UML) . . . . .	66
6.1	Diagrama de Entidade e Relacionamento da Base Transparência Brasil . . . . .	72
6.2	Diagrama de Entidade e Relacionamento da Base DCC-UFS . . . . .	73
6.3	Diagrama do Modelo Dimensional Mandato Político . . . . .	74
6.4	Diagrama do Modelo Dimensional Atividade Econômica . . . . .	74
6.5	Diagrama do Modelo Dimensional Candidatura . . . . .	75
6.6	Diagrama do Modelo Dimensional Verbas . . . . .	76
6.7	Diagrama do Modelo Dimensional Viagem . . . . .	77
6.8	Diagrama do Modelo Dimensional Emenda . . . . .	77
6.9	Diagrama do Modelo Dimensional Candidatura Doações e Voto . . . . .	78
6.10	Mandato eletivo antes: forma original como campo texto . . . . .	79
6.11	Mandatos eletivos depois: visualização dos dados transformados . . . . .	79
6.12	Bancadas política antes: tabela original . . . . .	79
6.13	Bancada política depois: visualização dos dados transformados . . . . .	80
6.14	Emendas parlamentares antes: forma original . . . . .	81
6.15	Emendas parlamentares depois: dados transformados . . . . .	81
6.16	Organograma do Poder Judiciário. (MATOS, jfal) . . . . .	83
6.17	Processo ETL para a dimensão DIM_POLITICO . . . . .	87
6.18	Processo ETL para a dimensão DIM_PROCESSO . . . . .	87

6.19	Recuperação dos mandatos a partir das descrição do cargo . . . . .	88
6.20	Processo ETL para a FATO_MANDATO . . . . .	88
6.21	Transformação dos dados de Candidaturas . . . . .	89
6.22	Processo ETL para a FATO_CANDIDATURA . . . . .	89
6.23	Processo ETL para a FATO_EMENDA . . . . .	90
6.24	Processo ETL para a FATO_ATIVIDADE_ECONOMICA . . . . .	90
6.25	Processo ETL para a FATO_VERBA . . . . .	90
6.26	Processo ETL para a FATO_VIAGEM . . . . .	91
6.27	Transformação dos dados de matérias apresentadas ao plenário . . . . .	91
6.28	Processo ETL para a FATO_MATERIA . . . . .	91
6.29	Interface de Consulta do Sistema <a href="http://www.prosopoliticos.uff.br">www.prosopoliticos.uff.br</a> . . . . .	93
6.30	Resultado da busca: registro ampliado . . . . .	94
7.1	Interface do <i>software</i> ATLAS.Ti apresentando as etapas de codificação inicial e axial. . . . .	98
7.2	Relatório de códigos e criação de uma rede de categorias a partir de um documento . . . . .	99
7.3	Planejamento da análise dos dados . . . . .	100
7.4	Rede sobre as fontes de informações sobre políticos. Fonte: do próprio autor, . . . . .	101
7.5	Rede sobre as redes sociais e <i>fake news</i> . Fonte: do próprio autor, . . . . .	103
7.6	Rede sobre educação e engajamento político. Fonte: do próprio autor, . . . . .	104
7.7	Rede sobre as características do sistema e percepção de utilidade. Fonte: do próprio autor, . . . . .	106
7.8	Percepções sobre a facilidade de uso do sistema. Fonte: do próprio autor. . . . .	109
7.9	Indícios das intenções de uso do sistema. Fonte: do próprio autor, . . . . .	110

7.10 Contribuições para sociedade e sugestão de melhorias. Fonte: do próprio autor, . . . . .	111
---	-----

## Lista de Tabelas

2.1	Principais autores e suas proposições sobre a <i>Design Science Research</i> . Fonte (DRESCH, 2013), p.82 . . . . .	11
2.2	Produtos do DSR. Traduzido de (VAISHNAVI; KUECHLER, 2004) . . . . .	12
4.1	10 Fatores mais influentes e menos influentes . . . . .	54
4.2	Distribuição Sócio Demográfica: gênero, idade, orientação sexual, concordância entre gênero e sexo, cor/etnia, estado, localização, grau escolaridade, profissão, renda bruta e religião. . . . .	56
4.3	Ranking das avaliações sobre os grupos de fatores (3ª Etapa) . . . . .	58
4.4	Comparativo dos modelos cognitivos com o modelo FIV.Br2018 . . . . .	60
5.1	Seleção das regras para compor o modelo dimensional . . . . .	69
7.1	Perfil dos Entrevistados . . . . .	95

## Lista de Nomenclaturas

<b>API</b>	Application Programming Interface (Interface de Programação de Aplicações)
<b>BD</b>	Banco de Dados
<b>CSV</b>	<i>Comma Separated Value</i> (Valor separado por vírgula)
<b>DM</b>	<i>Data Mart</i> (Armazém de Dados Departamental)
<b>DW</b>	<i>Data Warehouse</i> (Armazém de Dados)
<b>DSR</b>	<i>Design Science Research</i> (Pesquisa em Ciência do Projeto)
<b>DSRM</b>	<i>Design Science Research Methodology</i>
<b>XML</b>	<i>Extensible Markup Language</i> (Linguagem de Marcação Extensível)
<b>ETL</b>	<i>Extract Transform Load</i> (Extração, Transformação e Carregamento)
<b>FASIV</b>	<i>Fatores que Suportamente Influenciam o Voto</i>
<b>GT</b>	<i>Grounded Theory</i> (Teoria Fundamentada em Dados)
<b>MDX</b>	<i>Multidimensional Expressions</i> (Expressões Multidimensionais)
<b>NFR</b>	<i>Non-Functional Requirement</i> (Requisitos não funcionais)
<b>OLAP</b>	<i>On-Line Analytical Processing</i> (Processamento Analítico On-Line)
<b>PDI</b>	Pentaho Data Integration
<b>PDA</b>	Plano de Dados Abertos
<b>SI</b>	Sistemas de Informações
<b>SIG</b>	<i>Softgoal Interdependency Graph</i> (Gráfico de Interdependência de MetaObjetivos)
<b>TAM</b>	<i>Technology Acceptance Model</i> (Modelo de Aceitação da Tecnologia)
<b>TSE</b>	Tribunal Superior Eleitoral
<b>UFS</b>	Universidade Federal de Sergipe

# 1. Introdução

Neste capítulo é apresentada a motivação e o problema investigado nesta pesquisa, sua relevância, seu objetivo, as questões de investigação, as delimitações e, por fim, a organização do conteúdo desta pesquisa nos demais capítulos desta dissertação.

## 1.1 Motivação e Contexto

A transparência de informações vem sendo discutida em diversos setores da sociedade e tornou-se amplamente difundida depois que algumas iniciativas foram criadas, tais como: Lei Sarbanes-Oxley (SOX, 2002), que visa garantir a criação de mecanismos de auditoria e segurança confiáveis nas empresas de modo a mitigar riscos aos negócios, evitar a ocorrência de fraudes ou assegurar que haja meios de identificá-las quando ocorrerem, garantindo a transparência na gestão das empresas; *Open Government Partnership* (OGP, 2011), que serviu de mecanismo de luta contra a corrupção por meio da participação social e o desenvolvimento de novas tecnologias; e a Lei de Acesso à Informação (LAI, 2011), que regulamenta o direito constitucional de acesso às informações públicas no Brasil.

Ao implementar o conceito de transparência num contexto político, pode-se assumir que, quanto maior o grau de transparência das informações, maiores serão as possibilidades de torná-las mais confiáveis aos olhos do cidadão e melhorar o entendimento sobre as informações históricas de um político ou de um grupo político. Desse modo, o cidadão terá suas análises e sínteses beneficiadas.

Das motivações para desenvolver esta pesquisa, a primeira delas parte de um dos mais importantes eventos no contexto político da história do país: as eleições para os principais cargos de mandato eletivo do Brasil. A eleição de 2018 ocorreu de maneira atípica e provocaram no cidadão um certo desconforto e temor, tanto no papel de receptor quanto

emissor da informação e tomador de decisão.

Diferentemente das eleições anteriores, uma indústria de desinformação se instalou e, através do uso e abuso das *fake news*, que segundo (ALLCOTT; GENTZKOW, 2017), consiste em artigos de notícias que são intencionais e verificadamente falsos e podem enganar os leitores, fez do ciberespaço um universo de incerteza e de pós-verdade. O que vivenciamos foi um cenário contagiante de descontentamento, falta de diálogo e repressão. A ocorrência dos memes, que segundo (BLACKMORE, 2000), são informações que influenciam o comportamento e que são transmitidos por imitação, foram disparados via aplicativos de mensagens instantâneas e se tornaram meios de provocação de violência exibicionista. Em razão disso, percebeu-se que as relações sociais e familiares foram e estão sendo comprometidas por causa dessa comunhão de desafetos. Nota-se que há um abismo entre conversar e convencer, e em meio a essa situação, uma transposição de indignação em xingamentos, ódio e brutalidade.

A decisão de voto, que antes pensávamos ser baseada nos fatos, apresentações e defesa de propostas, hoje percebemos como sendo baseada principalmente em sentimentos de afeto, amor e ódio, que reúnem os eleitores em torcidas. Não importa tanto o debate, não importa tanto os meios, o importante é torcer para o seu candidato ganhar. As eleições parecem uma espécie de videogame em que os eleitores estão tentando fazer o seu candidato passar de nível a qualquer custo. Diferente das eleições presidenciais de 2014, a ocorrência das *fake news* foi bastante expressiva. Percebeu-se uma campanha com base em notícias falsas (como a afirmação de que o armamento da população diminui a violência ou como a simplificação de fatos relativos à corrupção), preconceitos dos mais diversos, apologia à violência, desqualificação do oponente, entre outras coisas, pareceram estratégias que se tornaram aceitáveis para os eleitores. Em se tratando de tomada de decisão, (LIMEIRA; MAIA, 2010) explica que a construção da decisão de voto do eleitorado, ao longo do tempo, é fortemente influenciada pelas mensagens midiáticas, pelas estratégias de propaganda dos partidos e candidatos, bem como pelas interações sociais vivenciadas pelos eleitores ao longo da campanha eleitoral. Considerando esse cenário – em que as redes sociais proliferam e estimulam a troca de informação, onde atuam sistemas inteligentes, e que a publicidade pode ser feita com base em buscas e comportamentos –, torna-se imperativo tentarmos compreender o que está influenciando a tomada de decisão do cidadão brasileiro na contemporaneidade, e de que forma podemos minimizar essa pulverização de informações aparentemente duvidosas, oriundas de várias fontes de informações.

Esta pesquisa teve como ponto de partida o Portal Excelências do Grupo Transparência Brasil, que em sua versão restrita, integra as informações contidas nas bases de dados

oficiais das casas legislativas brasileiras. Sobre esse Portal, surgiram alguns questionamentos em relação ao uso, a relevância das informações para o cidadão e sobretudo a forma como estas poderiam estar organizadas, considerando que as fontes de dados oficiais são de origens diversas e, além disso, podem apresentar estruturas e modelagens distintas.

Nesta pesquisa, estamos interessados em conteúdos prosopográficos. Sobre a natureza da informação prosopográfica, em (STONE, 1971), o autor explica que a prosopografia consiste numa verificação de características de uma pessoa, comuns a um grupo de pessoas na história por meio de um estudo coletivo de suas vidas. Consideramos importante a análise das informações disponíveis nos diversos portais de transparência que detalham ou exploram a vida de um político. No entanto, os dados contidos nessas plataformas não estão organizados de modo a tornar a informação mais inteligível e mais confiável ao cidadão. Acreditamos que, se os dados prosopográficos dos políticos estiverem organizados segundo um modelo multidimensional de dados, baseado num modelo conceitual de conhecimento, então esses dados serão melhor compreendidos pelos cidadãos. Essa é a hipótese desta pesquisa.

Na literatura, foram encontradas algumas contribuições sobre prosopografia com foco na identificação de perfis, baseados em modelos estatísticos. Enquanto o trabalho de (BRAGA; NICOLÁS, 2008) examinou os perfis de candidatos e criou indicadores para medir o grau de disponibilidade de informações contidas nos sites das casas legislativas e analisou o comportamento desse índice em cada unidade federativa, esta corrente pesquisa buscou compreender o comportamento do cidadão em função de fatores que influenciam a decisão de cunho eleitoral, bem como a relevância da informação. Além disso, está inclinada a explorar a organização dos dados através de técnicas de modelagem multidimensional de dados, bem como orientar a extração por meio de fontes confiáveis. Ademais, visa dar condições que apoiem o cidadão em suas pesquisas por informações de natureza política, sobre alguém detentor de um mandato eletivo, seja para tomada de decisão ou algum trabalho de cunho investigativo.

## **1.2 Objetivos Gerais e Específicos**

O objetivo desta pesquisa é construir conhecimento sobre como organizar e apresentar dados prosopográficos sobre políticos ou candidatos a cargos eletivos em apoio às necessidades informacionais dos cidadãos. Dentre as “apostas” que fizemos, consideramos desenvolver modelos multidimensionais de dados e apresentar os dados por meio de um

sistema de consulta interativo, o que nos levou a desenvolver um sistema computacional para implementar essa proposta de solução, e assim podermos avaliar a sua adequação.

O sistema desenvolvido nessa pesquisa visa compreender a organização de dados que são de interesse do cidadão e, para que isso seja possível, os seguintes procedimentos foram seguidos:

- Examinar as fontes de dados existentes;
- Verificar a autenticidade das bases, se são bases oficiais ou não;
- Estabelecer acesso às fontes de dados e documentá-las;
- Pesquisar sobre o comportamento do eleitor e identificar quais informações são relevantes para o cidadão ao analisar um perfil político;
- Especificar e documentar um modelo conceitual de conhecimento;
- Definir e documentar os requisitos de negócios para composição do modelo multi-dimensional, o que inclui a seleção dos processos de negócios, a granularidade do dado, as tabelas de dimensões e as tabelas de fatos;
- Decidir sobre as ferramentas utilizadas e executar as atividades de extração, transformação e carregamento dos dados: filtragem, combinação e padronização das dimensões;
- Criar os meta-dados e publicação das estruturas dimensionais para apresentação dos dados; e
- Produzir o sistema de informação para consulta dos dados prosopográficos.

### **1.3 Problema, Questão de Pesquisa e Proposta de Solução**

Um dos pontos de relevância para a interação entre política e as plataformas digitais é destacado por (EISENBERG, 2003) como sendo o poder de fixação da informação pela internet, onde o receptor tem a capacidade de acessar e rever de maneira sistemática todas as suas informações, mantendo-se bem informado e inteirado ao meio político a que pertence. Isso ajuda a compreender o porquê da internet exercer influência significativa quando se trata de um processo eleitoral. Contudo, (PERISSINOTO; COSTA; TRIBESS, 2009) relata que, por vezes, ficamos limitados ao material que está disponível na internet, tendo, em muitos casos, falta de informações sobre determinados aspectos.

Nessa perspectiva, os portais de transparência carecem de maior atenção, pois percebe-se uma baixa rastreabilidade de informações e dificuldade do cidadão em relação ao entendimento das informações prosopográficas disponíveis nas diversas plataformas, o que torna a tarefa de tomada de decisão bastante dificultosa.

O problema abordado nesta pesquisa consiste na dificuldade de entendimento sobre as informações prosopográficas dos políticos brasileiros dada a falta de organização, falta de estruturação e incompletude dos dados. Na sociedade contemporânea, as informações são disponibilizadas através de vários sistemas de comunicação e informação disponíveis na internet. Contudo esses dados não seguem uma organização ou estrutura que atendam às necessidades informacionais e analíticas de quem busca pela informação.

**Diante do exposto, a questão de pesquisa a ser respondida é: como possibilitar a melhoria do entendimento do cidadão e trazer informações relevantes sobre a vida dos políticos, por meio de um modelo multidimensional de dados prosopográficos?**

Propomos neste trabalho um modelo multidimensional de dados sobre os políticos, suas histórias, suas candidaturas, etc. Quando pensamos nesse modelo, estamos considerando os fenômenos contemporâneos que estimulam a busca e compartilhamento de informações sobre políticos e que têm sido considerados para a construção de um artefato que consolide essa proposta. Através de um modelo multidimensional, foi criado um sistema de informações como uma plataforma de consulta, capaz de concentrar as informações relevantes sobre a vida do políticos, organizadas de modo a facilitar o entendimento sobre elas. A abordagem epistemológica-metodológica utilizada nesta pesquisa, como forma de produção de conhecimento, a partir da construção do artefato, foi a *Design Science Research (DSR)*.

Discutimos nossa abordagem à luz do Portal Excelências mantido pela organização não-governamental Transparência Brasil, que em sua versão restrita concentra as informações contidas nas bases de dados oficiais das casas legislativas brasileiras. A organização “Transparência Brasil” concentra sua atuação na busca pela integridade do poder público, principalmente através do aumento do acesso à informação para combater a corrupção.

#### **1.4 Ambiente**

O uso da Internet para divulgação de informações em campanhas e para a promoção da transparência na política aumentou substancialmente. A quantidade de informações disponíveis sobre as atividades parlamentares de cada deputado, sobretudo as criminais,

nunca foi tão grande e, dada a rápida velocidade de difusão da informações através de vários meios, tornou-se facilmente acessível ao cidadão comum.

Para fomentar a disseminação da informação, o Portal Excelências, projetado pela organização Transparência Brasil, foi o cenário que originou a discussão e a identificação do problemas trabalhados nesta pesquisa.

De acordo com (TRANSPARENCIA, 2015), o Transparência Brasil é uma associação sem fins econômicos ou lucrativos, destinada a promover a defesa do interesse público por meio da edificação da integridade do Estado brasileiro e o combate à corrupção, contribuindo para o aperfeiçoamento das instituições e do processo democrático.

O Portal Excelências é uma das iniciativas criadas pelo Grupo Transparência Brasil. Segundo (ABRAMO et al., 2015), o projeto tem o objetivo de trazer informações sobre os parlamentares em exercício nas casas legislativas das esferas federal e estadual. Além disso, os dados informados são extraídos de fontes públicas e governamentais, tal como as próprias casas legislativas, o Tribunal Superior Eleitoral, tribunais estaduais e superiores, tribunais de Contas e de outros projetos mantidos pela Transparência Brasil. A Figura 1.1 apresenta a tela principal do projeto.

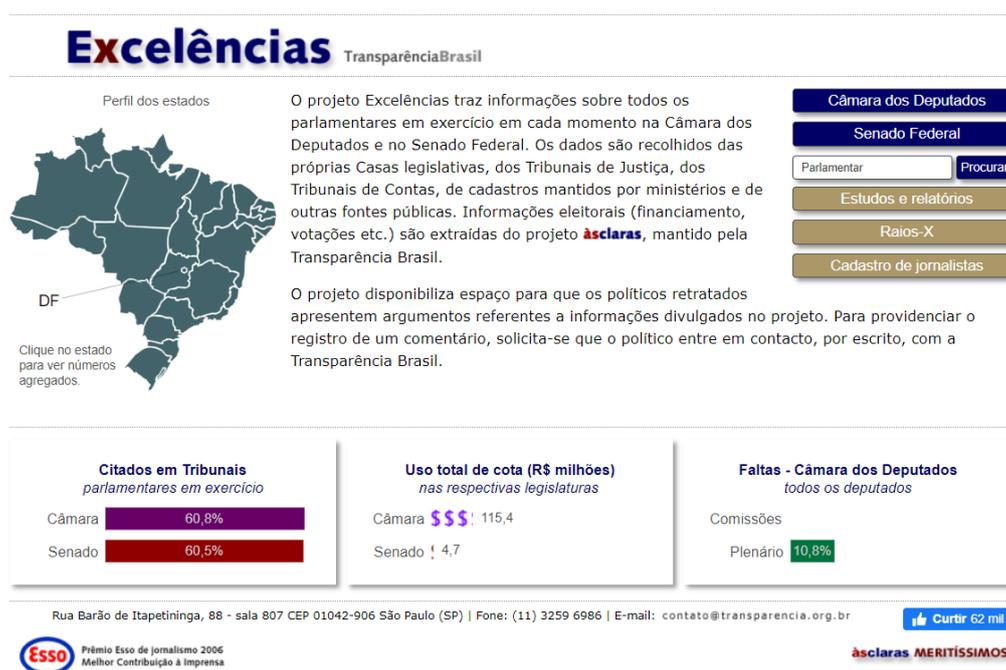


Figura 1.1: Tela principal do Portal Excelências, (TRANSPARENCIA, 2015)

Para sustentação dos projetos, o Grupo Transparência Brasil tem como colaboradores: profissionais de ciências sociais, ciências de dados, tecnologias da informação e analistas em transparência. Por falta de recursos, o projeto não pôde dar prosseguimento às suas atividades e permanece acessível somente para usuários autorizados pelo Grupo.

As informações que estão disponíveis no Portal Excelências consistem em:

- Dados pessoais, cargos e funções públicas exercidas pelos parlamentares;
- Informações sobre o desempenho legislativo dos parlamentares, tais como faltas, votações, utilização de verbas de gabinete;
- Histórico das ementas ao orçamento apresentadas;
- Registros de doações de campanhas eleitorais;
- Notas publicadas à imprensa sobre envolvimento do candidato em casos de corrupção;
- Informações sobre processos em que o parlamentar é indiciado como réu na Justiça Federal, nos Tribunais Superiores, na Justiça Eleitoral e nos Tribunais de Contas; e
- Recurso de busca por parlamentar.

Em contato com esse grupo e ao analisar o conteúdo do Portal Excelências, a fim de colaborar com a continuidade de algumas ações, surgiram alguns questionamentos que nortearam a delimitação do problema da presente pesquisa, visto na seção 1.3. As discussões foram em torno das seguintes indagações:

- O Portal Excelências possui todas as informações relevantes do político e seus relacionamentos sociais?
- O cidadão consegue entender com clareza e concisão as informações expostas no Portal?
- Há uma ideia de compartimentação ou agrupamento de informações?
- As informações estão categorizadas de modo a ajudar o(a) eleitor(a), em tempos de eleição, a organizar melhor suas ideias e fazer suas comparações?
- Que tipos de informações poderiam ser exibidas no ambiente?

Essas foram as principais indagações sobre a organização dos dados em prol da transparência de informações aos cidadãos que recebem a todo tempo dados oriundos de várias fontes e plataformas de informações espalhadas pelas redes e ecossistemas digitais.

## **1.5 Organização da Dissertação**

O conteúdo desta dissertação está estruturado da seguinte forma: o capítulo 2 descreve a abordagem epistemológico-metodológica; o capítulo 3 apresenta a fundamentação teórica da pesquisa; o capítulo 4 apresenta as conjecturas comportamentais da pesquisa; o capítulo 5 apresenta a solução proposta nesta pesquisa, descrevendo a organização da informação através do modelo de dimensões para informações prosopográficas; o capítulo 6 apresenta o desenvolvimento do artefato desenvolvido a partir da solução proposta; o capítulo 7 relata a avaliação do artefato e da solução com base em uma pesquisa empírica e discutindo os resultados obtidos em comparação com trabalhos relacionados; e o capítulo 8 é a conclusão desta dissertação.

## 2. Abordagem de Pesquisa: *Design Science Research* (DSR)

Adotamos *Design Science Research* (DSR) como abordagem epistemológico-metodológica da pesquisa. A DSR, de acordo com (BAX, 2017), é o tipo de metateoria que auxilia o pesquisador a criar conhecimento teórico juntamente com o desenvolvimento de artefatos, juntos numa pesquisa de caráter científico.

Há um série de discussões acerca do desenvolvimento tecnológico em contraste com a construção do conhecimento científico. No trabalho de (PIMENTEL; FILIPPE; SANTORO, 2018), é apontada a falta de rigor teórico-metodológico em muitas pesquisas da Área de Computação. Esses autores reconhecem que DSR legitima o desenvolvimento de artefatos como um meio para a produção de conhecimentos científicos do ponto de vista epistemológico, principalmente a partir da obra “As Ciências do Artificial”, de Herbert Simon (SIMON, 1996).

Herbet Simon popularizou *Design Science* (SIMON, 1996). Contudo, no século XV, Leonardo da Vinci (1452-1519) já tinha uma percepção da relevância e influência das ciências da engenharia. Reconhecido e memorável pela qualidade e encantamento dos projetos realizados e pelo seu interesse por várias áreas de conhecimento, Da Vinci foi um inventor de soluções. De acordo com (FONSECA, 1997), Da Vinci se interessou pela engenharia, designadamente pela engenharia militar, campo no qual inventa uma enorme quantidade de maquinaria, desde as famosas máquinas voadoras, cujos desenhos todos conhecemos, dos carros de assalto aos submersíveis. Além disso, a pesquisa de (DRESCH, 2013), reúne alguns autores que foram importantes para construção do conceito da *Design Science*. Dentre eles Giovanni Vico, crítico da visão analítica das ciências tradicionais, é considerado um dos mais importantes críticos com obras publicadas entre 1702 e 1725, cuja ideias são conhecidas como inovação epistemológica.

Em (TAKEDA; VEERKAMP; YOSHIKAWA, 1990), o primeiro método é formalizado para o desenvolvimento de trabalhos voltados para a ciência do design. Em seguida,

as ideias de (MARCH; SMITH, 1995) e (VAISHNAVI; KUECHLER, 2004), com base em Hebert Simon, impulsionaram as pesquisas com características práticas e aplicadas, e propuseram o *Design Cicle*, que inspirou a concepção do método *Design Science Research*, instrumentalizando as investigações *Design Science*.

Em (JR; CHEN; PURDIN, 1990), os autores apresentam suas contribuições para a disciplina de *Design Science*, embora não tenham recorrido aos conceitos propostos por (SIMON, 1996) e sequer utilizado a expressão *Design Science* em seu artigo – optaram pela expressão “*engineering approach*” -, mas a relação entre ambas é clara. Esses autores também reforçam a sugestão de construção de teorias, “o desenvolvimento de novas ideias e conceitos, e construção de *frameworks* conceituais, novos métodos, ou modelos”.

Conforme (DRESCH, 2013), a partir de 2000, *Design Science* passa a permear a Área de Gestão Organizacional. Nesse sentido, alguns autores começam a dialogar sobre a necessidade de uma ciência que respalde a projeção inovadora de artefatos organizacionais. No campo da gestão, (ROMME, 2003) recomenda que a DSR seja aplicada em pesquisas nessa área, além disso ele afirma que rigor é imprescindível para garantir que os produtos das pesquisas sejam facilmente ensináveis nas academias e aceitos em publicações, garantindo maior interação entre o mundo prático e o teórico. Já (AKEN, 2004) traz publicações acerca da motivação e relevância do uso da DSR na área de Gestão, onde os resultados dariam suporte aos desafios do mundo real, tendo como contribuição o conhecimento que também poderia ser aplicado em outras iniciativas.

De acordo com (PRZEYBILOVICZ, 2014), os autores indicam que as teorias elaboradas em DSR devem ter as seguintes características:

- Trabalharem com objetivos contingenciais;
- Não envolverem explicação pura ou previsão do fenômeno estudado;
- Serem prescritivas;
- Serem compostas de teorias computacionais (Kernel), Ciências Naturais, Sociais e Matemáticas;
- Enquanto teorias explicativas preocupam-se em “o que é”, as preditivas “o que será”, as normativas “o que deve ser”, as teorias do Design Science dizem “como/porque” (HUFF; TRANFIELD; AKEN, 2006);
- *Design Science* mostra como as teorias explicativas, preditivas, ou normativas podem ser colocadas em prática; e

- Em *Design Science* o objetivo é tanto de prescrever as propriedades que um artefato deveria ter quanto alcançar determinados objetivos e os métodos no processo de construção do artefato.

Um resumo dos principais autores e suas proposições acerca da DSR é apresentado na Tabela 2.1.

Tabela 2.1: Principais autores e suas proposições sobre a *Design Science Research*. Fonte (DRESCH, 2013), p.82

Autor	Proposição
Leonardo da Vinci	Utiliza-se das ciências da engenharia para solucionar problemas que, até então, as ciências tradicionais não tinham conseguido resolver.
Giovanni Battista Vico	Contesta a “análise reducionista cartesiana” e propõe que o conhecimento científico seja fundamentado nas “ciências do gênio (l’ingenium)”.
Herbert Simon	Crítica o uso exclusivo do método analítico ou reducionista. Defende que o projeto do conhecimento é mais importante do que o objeto de conhecimento. Propõe o uso das ciências da concepção – Design Science.
Takeda et al.	Discute e faz uma primeira tentativa de formalização de um método de pesquisa fundamentado nos conceitos de Design.
Nunamaker et al.	Busca formalizar um método para a pesquisa fundamentada em Design Science. Expõe alguns produtos da pesquisa amparada pela Design Science.
Walls et al.	Defende a utilização dos conceitos da Design Science para condução de pesquisas. Aborda o conceito de teorias prescritivas e sua importância para o desenvolvimento de soluções práticas e efetivas para problemas existentes.
Gibbons et al.	Aborda um novo modo para produção do conhecimento: Tipo 2. Este conhecimento estaria mais voltado à construção de conhecimentos relevantes produzidos no contexto de aplicação e não somente na academia.
Le Moigne	Versa sobre as Novas Ciências: voltadas à concepção e não somente na análise do objeto de pesquisa.
March e Smith	Defendem a integração entre a Design Science e as ciências tradicionais a fim de conduzir pesquisas preocupadas em desenvolver soluções.
Romme	Aborda o uso da Design Science na área de gestão. Afirma que é necessário uma ciência que auxilie na criação de novos artefatos organizacionais. Discute ainda a questão de rigor e relevância das pesquisas em gestão.
Van Aken	Preocupado com a relevância das pesquisas na área de gestão e nas organizações de forma geral, sugere a aplicação da Design Science para condução de pesquisas mais relevantes. Afirma que as pesquisas realizadas devem ser prescritivas, facilitando sua utilização pelas organizações, e também, generalizáveis – não servir para resolver somente um problema em dada situação, mas sim para resolver problemas de uma certa Classe de Problemas.

O método DSR envolve um processo rigoroso para projetar artefatos a fim de resolver problemas observados, para fazer contribuições de pesquisa, para avaliar os projetos e para comunicar os resultados ao público apropriado (HEVNER et al., 2004). No trabalho de (ALAN et al., 2004), os autores afirmam que o conhecimento em DSR se manifesta na forma de artefatos, que podem ser constructos, modelos, *frameworks*, arquitetura, princípios de *design*, métodos, e/ou instanciações e teorias do *design*, conforme é apresentado na tabela 2.2. Ainda em (ALAN et al., 2004), os artefatos podem ser classificados como materiais e abstratos, por exemplo as instanciações são um tipo de artefato material, enquanto os outros tipos são considerados artefatos abstratos. A Teoria do *Design* geral-

mente inclui artefatos abstratos e instanciações. Em um artigo muito citado, (MARCH; SMITH, 1995) comparam a pesquisa DSR com as pesquisas em ciências naturais e produzem produtos gerais para DSR: constructos, modelos, métodos e instanciações.

Tabela 2.2: Produtos do DSR. Traduzido de (VAISHNAVI; KUECHLER, 2004)

Tipo de Artefato	Descrição
Constructo	Vocabulário conceitual de um domínio
Modelo	Conjuntos de proposições ou declarações que expressam relações entre constructos
<i>Framework</i>	Guia, conceitual ou real, que serve como suporte ou guia
Arquitetura	Sistemas de estrutura de alto nível
Princípio de Projeto	Princípios fundamentais e conceitos para orientar o projeto
Método	Conjunto de etapas para executar tarefas, como fazer
Instanciação	Implementação em ambientes que operacionalizam ou não, construtos, modelos, métodos e outros artefatos abstratos, sendo que no último caso, esse conhecimento permanece tácito.
Teorias de Projeto	Conjunto prescritivo de instruções sobre como fazer algo para atingir determinado objetivo. Uma teoria geralmente inclui outros artefatos abstratos tais como constructos, modelos, <i>frameworks</i> , arquiteturas, princípios de <i>design</i> e métodos.

(AKEN, 2004) ressalta que os artefatos também podem incluir inovações sociais. (JÄRVINEN, 2007) complementa que os artefatos apresentam novas propriedades de aspectos técnicos, sociais ou recursos informativos. Sendo assim, essa interpretação inclui qualquer objeto concebido com uma solução incorporada para um problema de pesquisa entendido. De acordo com (HEVNER et al., 2004), numa abordagem pragmática, a DSR não anseia alcançar verdades últimas, grandes teorias ou leis gerais, mas procura identificar e compreender os problemas do mundo real e propor soluções apropriadas, úteis, fazendo avançar o conhecimento teórico da área. Considerando a definição de Bunge (BUNGE, 2012), a tecnologia pode ser vista como o campo do conhecimento relativo ao projeto de artefatos e ao planejamento de sua realização, operação, ajuste, manutenção e monitoramento à luz do conhecimento científico. (PEFFERS et al., 2007) definem a *Design Science Research Methodology* (DSRM) como um método para o desenvolvimento de pesquisas de natureza tecnológica, em que se deseja desenvolver artefatos. Nesse mesmo artigo, os autores apresentam as etapas da DSRM e os objetivos pretendidos em cada fase. Na Figura 2.1, é apresentada a metodologia proposta pelos autores. Na presente pesquisa, buscamos seguir esses passos e fundamentos. A primeira fase do método é a identificação do problema e a motivação. Essa fase requer a definição do problema específico e a justificativa para a relevância da solução. A justificativa do valor da solução incentiva o pesquisador e o público da pesquisa a buscar e aceitar os resultados, além disso ajuda a entender a compreensão do pesquisador sobre o problema.

A segunda fase compreende a definição dos objetivos da pesquisa. A partir da definição do problema e da delimitação do que é possível e viável, pode-se inferir os objetivos

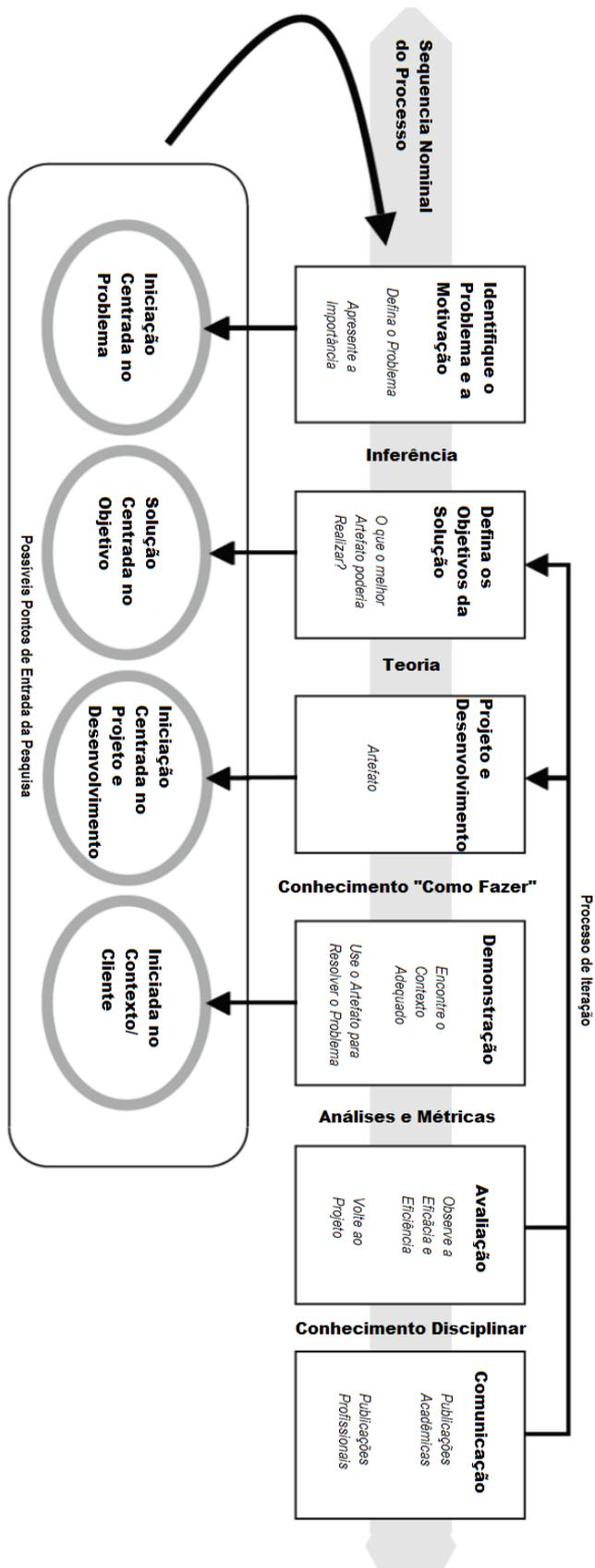


Figura 2.1: *Design Science Research Methodology* (DSRM). Fonte: (PIMENTEL; FILIPPE; SANTORO, 2018)

da solução, que podem ser quantitativos, por exemplo, em que medida uma solução seria melhor que as atuais, ou qualitativa, tal como uma descrição de como será um novo artefato (PEFFERS et al., 2007). Os recursos necessários nessa etapa incluem o conhecimento do problema, as soluções atuais e, se possível, deve-se explicitar a eficácia desejada da solução.

A terceira fase do método é dedicada ao projeto e ao desenvolvimento do artefato. Os autores afirmam que, conceitualmente, um artefato de pesquisa de design pode ser qualquer objeto projetado. Nessa etapa, deve-se especificar as funcionalidades que serão implementadas, a definição da arquitetura e a criação do artefato em si. O pesquisador deve utilizar conhecimento teórico para construir o artefato adequado à solução do problema.(PEFFERS et al., 2007).

A quarta fase é para realizar uma Demonstração do artefato. É nessa fase que o artefato é usado para demonstrar que ele resolve uma ou mais instâncias do problema. Segundo (PEFFERS et al., 2007), isso poderia envolver seu uso em experimentação, simulação, estudo de caso, prova de conceito ou outra atividade apropriada, e os recursos necessários para a demonstração inclui conhecimento efetivo de como usar o artefato para resolver o problema.

A quinta fase é para a Avaliação. Os autores recomendam que o pesquisador busque observar e medir o quanto o artefato resolve o problema. Nessa fase de avaliação, devem ser realizadas atividades que comparam os objetivos da solução com os resultados obtidos. Essa etapa requer, do pesquisador, conhecimentos em métricas e técnicas de análise de dados. Caso o resultado obtido não seja o desejado, o pesquisador deverá recorrer à etapa do Projeto e desenvolvimento afim de ajustar o artefato desenvolvido ou então desenvolver um novo artefato. Os autores afirmam que a Avaliação pode ser entendida como a verificação da utilidade, qualidade e eficácia do artefato, rigorosamente demonstradas via métodos científicos. Dependendo da natureza do problema, (PEFFERS et al., 2007) indicam que a avaliação pode se dar de várias formas: comparação das funcionalidades do artefato com os objetivos da solução (descritos na fase 2), desempenho das medidas, exemplo, valor do orçamento, ou itens produzidos, resultados de pesquisas de satisfação, *feedback* dos participantes ou usuários finais e simulações.

A sexta e última fase compreende a Comunicação. Deve-se comunicar o problema e sua importância, o artefato, sua utilidade e novidade, o rigor de seu design e sua eficácia para pesquisadores e outros públicos relevantes, como profissionais praticantes, quando apropriado. Os autores recomendam que, para realizar a Comunicação, o pesquisador divulgue os trabalhos através de publicações e eventos científicos.

Outra abordagem para apoiar os pesquisadores é proposta por (PIMENTEL; FILIPPE; SANTORO, 2018) e (PIMENTEL, 2019). Esses autores apresentam o Modelo DSR que serve de instrumento para o pesquisador explicitar os elementos de uma pesquisa em DSR, visando assim apoiar o planejamento e a comunicação da pesquisa. Esse modelo foi utilizado como referência nessa pesquisa, e é uma forma de checagem do planejamento da mesma. A Figura 2.2, apresenta os componentes do Modelo DSR.

Os elementos da instância do Modelo DSR da presente pesquisa encontram-se apresentados na Figura 2.3. Cabe destacar que instanciar o Modelo DSR foi importante para “organizar as ideias” em relação aos elementos e como eles precisam estar inter-relacionados na pesquisa. Serviu, também, como instrumento de comunicação entre os pesquisadores envolvidos (mestrando e orientadores), e também como uma forma de verificar o que já havia sido feito e do que ainda estava faltando na pesquisa.

Ainda em (PIMENTEL; FILIPPE; SANTORO, 2018), sobre o Modelo DSR, os autores relatam que um artefato é desenvolvido para tentar resolver um problema em um dado contexto. O artefato deve ser fundamentado pelo estado da técnica (conhecimentos técnicos relacionados ao desenvolvimento do artefato) e por conjecturas teóricas (que são as suposições de como resolver o problema, com base em teorias que compõem o quadro teórico). O contexto é, frequentemente, o ponto de partida para a compreensão de uma pesquisa em DSR. A presente pesquisa partiu do site “Portal Excelências”, que é o contexto que optamos investigar. A partir desse portal, buscamos estudar os fenômenos contemporâneos relacionados a esse portal, compreender um problema e desenvolver uma solução, que seria o artefato desenvolvido nesta pesquisa.

## **2.1 Critérios de Avaliação do Artefato**

Como a abordagem metodológica segue os preceitos do DSR, e ele apoia a projeção de artefatos para resolver um problema de pesquisa, é primordial a avaliação do comportamento do usuário ao interagir com o artefato. Para essa avaliação, adotamos o modelo *Technology Acceptance Model* (TAM) (DAVIS, 1989) como base para a definição dos tópicos de entrevistas sobre o comportamento do usuário ao utilizar o artefato desenvolvido nesta pesquisa.

De acordo com (DAVIS, 1989), o *Technology Acceptance Model* (TAM) representa os motivos pelos quais os usuários aceitam ou rejeitam a tecnologia da informação e, dessa forma, serve para indicar como melhorar a aceitação, pois oferece um suporte para prever e explicar a aceitação de uma nova tecnologia. No modelo TAM, estão represen-

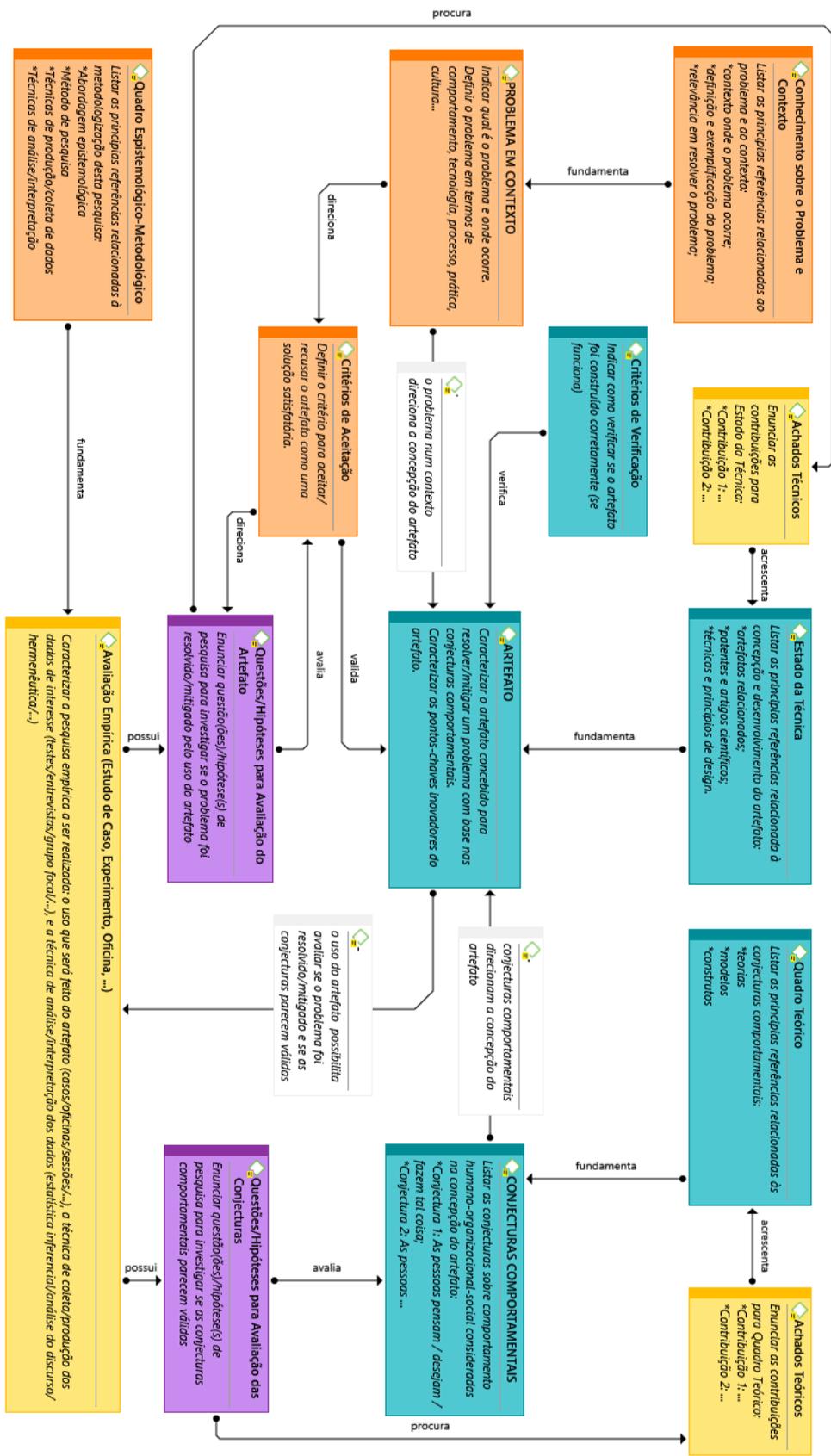


Figura 2.2: Modelo-DSR (PIMENTEL; FILIPPE; SANTORO, 2018) e (PIMENTEL, 2019)

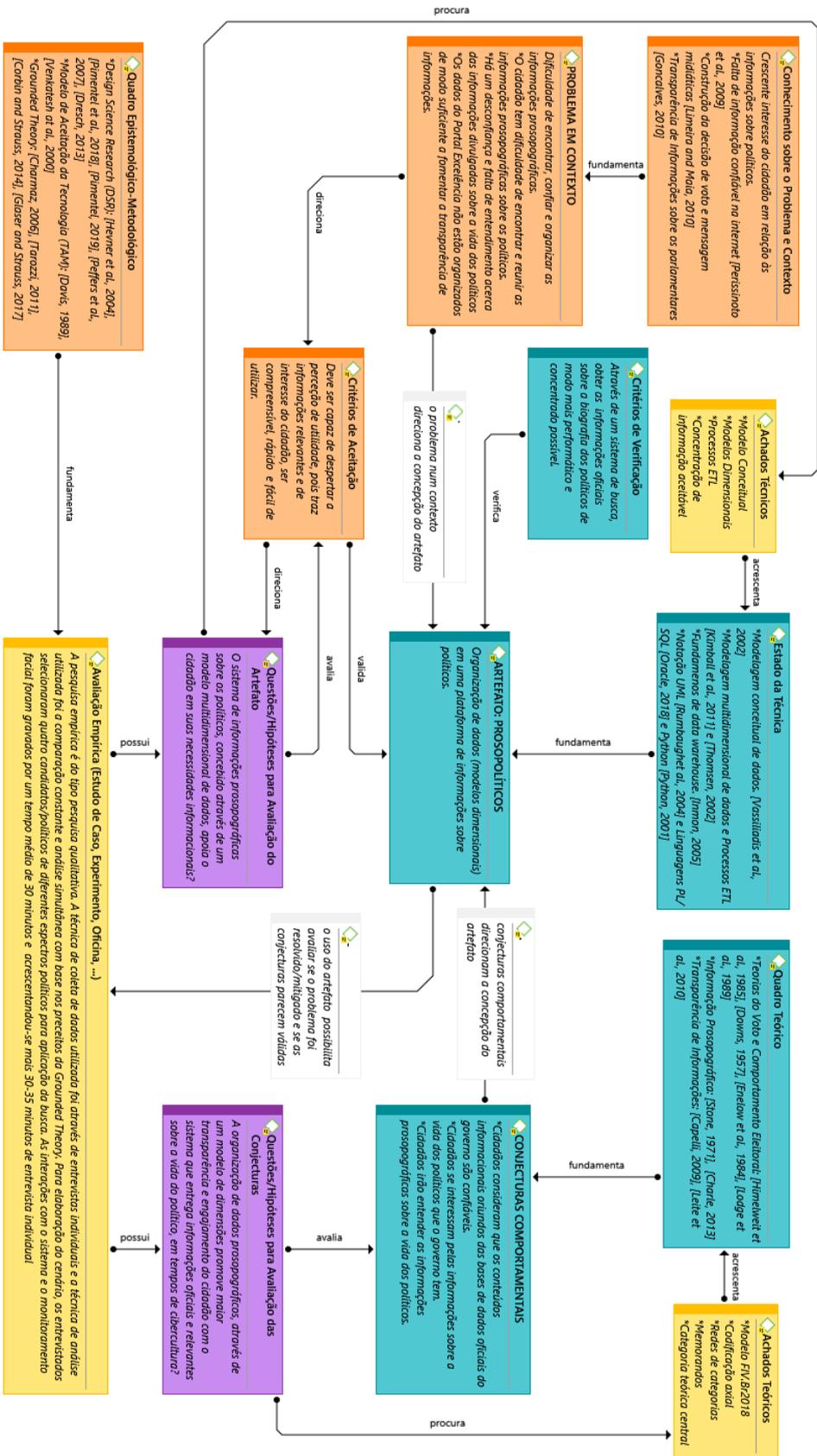


Figura 2.3: Modelo-DSR instanciado de (PIMENTEL; FILIPPE; SANTORO, 2018) e (PIMENTEL, 2019)

tados vários fatores que influenciam a decisão de um usuário sobre utilizar ou não uma nova tecnologia. Conforme apresenta a Figura 2.4, o modelo TAM se baseia nos seguintes constructos:

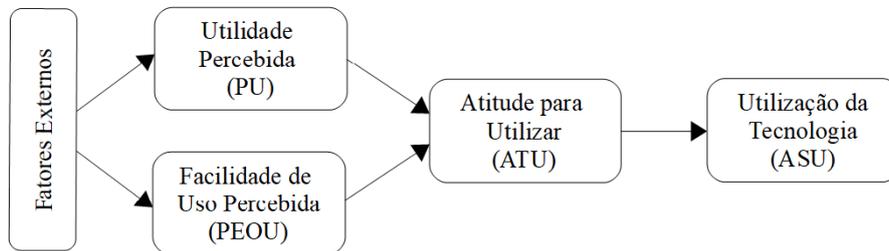


Figura 2.4: Modelo TAM traduzido de (DAVIS, 1989)

1. Utilidade Percebida (*Perceived Usefulness*) = grau em que um indivíduo acredita que a utilização do sistema contribui para aumentar o desempenho do seu trabalho.
2. Facilidade de Uso Percebida (*Perceived Ease of Use*) = grau em que um indivíduo acredita que o uso de um determinado sistema é intuitivo, não exigindo grande esforço.
3. Atitude Para Utilizar (*Attitude Toward Using*) = é definida como um sentimento positivo ou negativo de um indivíduo sobre a utilização do sistema. A utilidade e facilidade influenciam na motivação e atitude para o seu uso.
4. Utilização da Tecnologia (*Actual System Usage*) = é o resultado da motivação do indivíduo para utilizar a ferramenta através de sua atitude.

Porém, como este modelo é característico de pesquisas quantitativas, diferente dessa pesquisa, tais constructos serviram de base para elaboração dos “Critérios de Aceitação” que aparecem no Modelo-DSR. Por isso os constructos “Utilidade Percebida”, “Facilidade de Uso” e a “Intensão de Uso” (atitude para usar) serviram de inspiração para a definição do roteiro a ser seguido durante as entrevistas, em prol da avaliação do sistema de informações para consulta aos políticos.

Com o suporte da TAM, elaboramos o roteiro das entrevistas com base nos 4 grupos de características. Novamente, consideramos que tais características classificam como os critério para aceitação ou recusa do artefato de acordo com o Modelo-DSR. Sendo o primeiro grupo representado pelas fontes de informação, é o grupo introdutório, onde a conversação iniciará e o entrevistado se aproximará do tema.

A - Fontes de informação sobre candidatos

- Quais as fontes de informação sobre candidatos você costuma consultar/confiar para tomar uma decisão em quem votar?
- Você busca informações em sites? Quais?

#### B - Utilidade Percebida

- Como foi a sua experiência com o sistema de informações?
- Que informações foram úteis e quais foram inúteis pra você?
- Para que essas consultas poderiam ser úteis a você?
- O que você mudaria no sistema? Por quê?
- Você encontrou as informações que queria sobre cada candidato?

#### C - Facilidade de Uso

- Como foi sua navegação pelo sistema?
- O que você achou dos textos, das imagens e demais recursos da interface ?
- Como as informações foram encontradas?

#### D - Intenção de uso

- Em que situações você voltaria a usar esse sistema?
- No final das contas, você gostou ou não-gostou do sistema?
- Você usaria esse sistema uma outra vez? Por quê?

## 2.2 Técnica de Codificação da *Grounded Theory*

Nesta pesquisa, a metodologia empregada para coleta de dados sob o paradigma qualitativo foi a *Grounded Theory* (GT), traduzida no Brasil para Teoria Fundamentada em Dados.

A Teoria Fundamentada em Dados foi concebida na década de 60 nos Estados Unidos e, diferente dos métodos qualitativos tradicionais baseados em teste de hipóteses,

(GLASER; STRAUSS, 2017) explica que essa metodologia consiste na descoberta e no desenvolvimento de uma teoria a partir das informações obtidas e analisadas sistematicamente e comparativamente. (URQUHART; LEHMANN; MYERS, 2010) afirmam que a Teoria Fundamentada em Dados possui 4 características principais, a saber:

1. O principal objetivo é a construção de teoria e não o teste.
2. O conhecimento prévio do domínio não deve levar a hipóteses pré-concebidas ou conjecturas sobre a pesquisa que a própria procura falsear ou verificar.
3. O processo de pesquisa envolve o constante esforço para coletar e comparar dados e constantemente contrastar novos dados com quaisquer conceitos emergentes e construções da teoria que está sendo desenvolvida.
4. Todos os tipos de dados são aplicáveis e são selecionados pelo pesquisador através de amostragem teórica.

(URQUART, 2010, pg. 359, tradução nossa)

De acordo com (CRESWELL; POTH, 2016) os passos executados na Teoria Fundamentada em Dados constituem a organização dos dados, a leitura preliminar, a codificação, a categorização, a representação e interpretação dos dados. Ainda em (CRESWELL; POTH, 2016), o autor sustenta que, inquestionavelmente a espinha dorsal de uma pesquisa qualitativa, é a ampla coleta de dados, proveniente e múltiplas fontes de informação.

Seguimos alguns princípios da metodologia de coleta de dados indicada na *Grounded Theory* Construtivista, na qual (CHARMAZ, 2006) reforça que, na GT baseada no construtivismo, o pesquisador é o co-construtor dos significados, ele é um co-construtor nas interações. A autora explica as principais práticas:

“O envolvimento simultâneo na coleta e na análise dos dados; A construção de códigos e categorias analíticas a partir dos dados e não de hipóteses preconcebidas e logicamente deduzidas; A utilização de método comparativo constante, que compreende a elaboração de comparações durante cada etapa de análise; O avanço no desenvolvimento da teoria em cada passo da coleta e da análise dos dados; A redação de memorandos para elaborar categorias, especificar as suas propriedades, determinar relações entre as categorias e identificar lacunas; A amostragem dirigida à construção da teoria, e não visando a representação populacional.” (CHARMAZ, 2019, pg. 19)

Entendemos que a GT se propõe em fazer emergir uma teoria a partir dos dados, e por isso eles coletam um pouco e analisam e teorizam, depois voltam a campo e produzem

novos dados que serão novamente interpretados e servirão para refinar a teoria, e assim sucessivamente. Em DSR, partimos de conjecturas e o que nos interessa nesta pesquisa, é a técnica de codificação da GT como suporte para a interpretação dos dados.

## 3. Fundamentação Teórica e Técnica

Este capítulo tem por objetivo apresentar a base de conhecimento teórico e técnico relacionado a esta pesquisa, que enfoca o entendimento de informações prosopográficas sobre políticos. A Seção 3.1 apresenta o quadro teórico desta pesquisa: o conceito de informação prosopográfica e os principais elementos que caracterizam esse tipo de informação (Subseção 3.1.1); transparência de informações e transparência de software (Subseção 3.1.2); transparência de informações no contexto do governo brasileiro (Subseção 3.1.3); e modelos cognitivos que explicam o comportamento do eleitor brasileiro (Subseção 3.1.4). Na Seção 3.2 é apresentado o quadro técnico da pesquisa: modelagem conceitual (Subseção 3.2.1), *Data Warehouse* (3.2.2) e modelagem multidimensional (Subseção 3.2.3). Por fim, na Seção 3.3 são apresentadas as conjecturas da pesquisa.

### 3.1 Quadro Teórico

De acordo com (BAX, 2017), no ciclo de rigor da DSR, deve-se trabalhar o conhecimento científico relacionado ao projeto de pesquisa. O pesquisador deve consultar a base de conhecimento, que consiste no conjunto de referências científicas encontradas na literatura. Além disso, a própria pesquisa também corroborará com a base de conhecimento, a medida que corrobora com as teorias e produz novos conhecimentos teóricos.

A seguir é apresentado o quadro teórico que fundamenta as conjecturas teóricas indicadas no mapa da Figura 2.2

#### 3.1.1 Informação Prosopográfica

A prosopografia é um método bastante utilizado pelos historiadores e, de acordo com (STONE, 2011), consiste numa investigação das características comuns de um grupo de atores na história por meio de um estudo coletivo de suas vidas. Este método visa esta-

belecer um universo definido e então investigar um conjunto de questões uniformes, tais como, informes de nascimento e morte, casamento e família, origens sociais e posição econômica herdada, lugar de residência, educação, tamanho e origem da riqueza pessoal, ocupação, religião, experiência em cargos e assim por diante.

Sobre as definições de prosopografia, (RIBEIRO, 2018) afirma que não há uma única definição do conceito de prosopografia, há uma multiplicidade de compreensões, porém existe algo que é comum em quase todas as definições, a saber, o estudo do indivíduo a partir da sua totalidade em que está imerso. Dessa forma, (STONE, 2011) procura, em suas definições, interpretar a relação do indivíduo com a sua universalidade sociológica.

Evidências sobre a utilização do método prosopográfico pode ser encontrada em (SIRINELLI, 1996), quando identificou que haviam muitas oportunidades na observação das coincidências de itinerário, geração e sociabilidade que configuram o campo de ação dos indivíduos que se aproximam ou se afastam em razão dos seu conflitos de ideias.

(STONE, 2011) explica que a pesquisa e desenvolvimento da prosopografia foi instituída por duas escolas distintas. A primeira, denominada Escola Elitista, preocupa-se com a dinâmica dos pequenos grupos, com a interação em termos de família, casamento e laços econômicos, de um número restrito de indivíduos. Segundo (LASSWELL; LERNER, 1965), os temas de estudo usualmente são as elites do poder, tais como os senadores da Roma antiga, dos Estados Unidos, os membros do parlamento ou dos gabinetes ingleses, mas os mesmos processos e modelos podem ser e são aplicados aos líderes revolucionários. As pesquisas tem um inclinação maior para os aspectos genealógicos, as atividades políticas e comerciais através de estudos de casos e pouco orientada estatisticamente.

A segunda escola, de acordo com (STONE, 2011), considera fortemente a análise de dados mais qualitativos ou correlações estatísticas por meio de estudos de casos. Inspirada nos estudos das massas, não há um preocupação com os detalhes ou aspectos mais íntimos da elite. Consideram a movimentação e opinião popular em vez do posicionamento dos “Grandes Homens”.

Na presente pesquisa, ambas as escolas contribuem para a formalização das características que serão investigadas sobre os políticos. Além disso, a prosopografia tem um papel muito importante, pois ela serve de suporte informacional para que sejam disponibilizados dados relevantes ao cidadão a fim de averiguar, entender e correlacionar alguns aspectos do comportamento político.

### 3.1.2 Transparência de Informações e Transparência de Software

A transparência surge no decorrer dos últimos anos nos contextos social, empresarial e político como relata (FUNG; GRAHAM; WEIL, 2007). Uma justificativa, segundo (SOUSA; LEAL; LEITE, 2016), é a ampliação do acesso à informação como forma de possibilitar um maior nível de confiabilidade e auditabilidade em serviços prestados.

Buscando entender o termo transparência, (CAPPELLI; LEITE, 2008) definiu transparência organizacional como sendo a existência de políticas, padrões e procedimentos que visam a fornecer, aos interessados, informações sobre a organização segundo características gerais de acesso, uso, qualidade de conteúdo, entendimento e auditabilidade. Sob essa ótica, os autores afirmam que a transparência do processo é a característica que possibilita ao cidadão acesso, facilidade de uso, qualidade de conteúdo, entendimento e auditoria aos/dos processos que tratam de informações de seu interesse, sob a tutela de centros de autoridade.

Nesse sentido, (CAPPELLI, 2009) trata do problema de entender, implementar e avaliar a transparência nos processos e informações organizacionais, construindo uma definição e sistematização da transparência por meio de um catálogo de transparência. Cappelli apresentou o *Transparency SIG (Softgoal Interdependency Graph)*, que define princípios de transparência como características representadas em um catálogo de requisitos não-funcionais (NFR) (Chung et al., 2000), e operacionalizações para sistematizar a incorporação da transparência em processos organizacionais e informações. O SIG é formado por 3 níveis, sendo que o primeiro nível contém os requisitos não-funcionais que estabelecem uma relação de contribuição com a transparência e os elementos de cada nível são representados por *softgoals*, resultando num grafo conforme explica (CHUNG; LEITE, 2009) e que se encontra na Figura 3.1

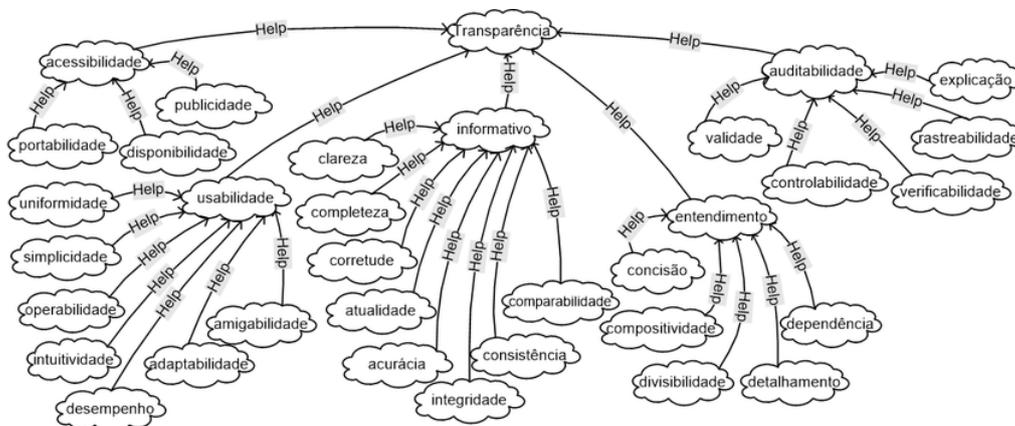


Figura 3.1: SIG de Transparência [Cappelli and Leite 2009]

No SIG são determinados 5 grupos de características que contribuem para alcançar a transparência: acessibilidade, usabilidade, informativo, entendimento e auditabilidade. Em (ENGIEL; ARAÚJO; CAPPELLI, 2012), a autora, ao elaborar o Catálogo de Entendimento de Modelos de Processo, explica que cada característica é definida por meio de um conjunto de outras características que, por sua vez, contribuem para atingir a característica imediatamente anterior. Exemplificando, para alcançar a característica de Entendimento é necessário realizar as atividades descritas na Figura 3.2

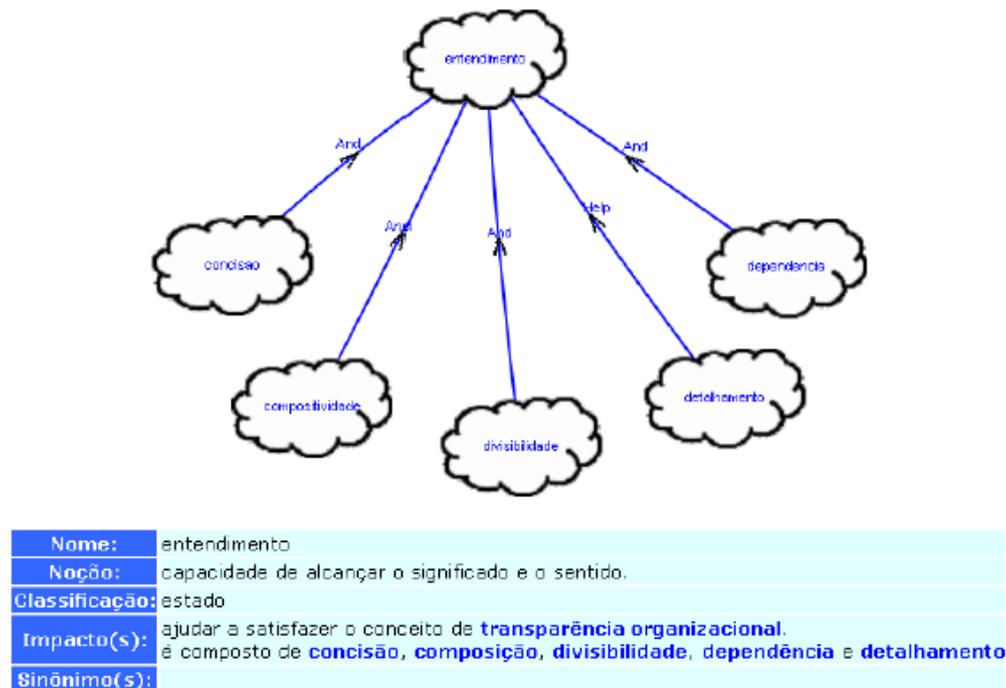


Figura 3.2: Operacionalização da Característica de Entendimento, (CAPPELLI, 2009)

Dentro desse paradigma, é fato que os Sistemas de Informação (SI) têm sido amplamente utilizados para dar suporte à execução e gerenciamento de processos, acesso e uso de informações. O objetivo principal é fornecer melhores maneiras de apoiar, por exemplo, a tomada de decisão, a colaboração, a análise e melhoria de como os processos são executados.

O fluxo de informações abertas é destacado pelos Sistemas de Informação, visando colaborar com um acesso mais democrático às informações (HOLZNER; HOLZNER, 2006). Portanto, a transparência é um requisito não funcional que o SI deve considerar durante sua elicitação, design e implementação. Esse requisito se espalha por diferentes partes funcionais de um SI e é uma preocupação nova e importante ao projetar um artefato de software que processa informações e faz interface com as pessoas. Isso é o que chamamos de “transparência por meio de software”.

Por outro lado, a transparência também se refere ao próprio software e seu processo de desenvolvimento, conforme explica (SERRANO; LEITE, 2011), ou seja, requisitos, projeto de design, codificação, código, teste, etc. Sua aplicação sistemática ainda é uma grande preocupação ao projetar sistemas que operacionalizam serviços e atuam como mecanismos de interface com pessoas que exigem saber como o SI foi desenvolvido e como ele executa suas ações. Isso é o que chamamos de “transparência no software”.

### **3.1.3 Transparência de Informações no Contexto Governamental Brasileiro**

Os governos têm mudado a forma de expor suas informações e seus processos de trabalhos, pois, incentiva e promove adoção de práticas de transparência organizacional, de informações e dados abertos nos órgãos, autarquias e as demais composições nas esferas públicas federais, estaduais e municipais juntamente com as empresas de economia mista.

Para que a abertura de dados de fato acontecesse, o governo brasileiro fez o Decreto nº 8.777 de 11/05/2016 que institui a Política de Dados Abertos. Em 2017, o governo orientou os órgãos da administração pública federal, autárquica e fundacional, através da resolução nº3 de 13/10/2017 que trata do Plano de Dados Abertos (PDA), como instrumento que instruiu a implementação e promoção das atividades que levam a abertura de dados. De acordo com (BRASIL, 2016), o objetivo do plano é organizar e padronizar os processos de publicação de dados abertos do estado, resultando em maior disponibilidade, acesso, qualidade e amplo reuso dos dados abertos às partes interessadas, tanto na sociedade quanto na própria administração pública. Segundo o Open Government Data (GROUP et al., 2007), os dados do governo serão considerados abertos se forem tornados públicos de uma maneira que esteja de acordo com os princípios abaixo:

- **Completos:** todos os dados públicos estão disponíveis. Dado público é o dado que não está sujeito a limitações válidas de privacidade, segurança ou controle de acesso.
- **Primários:** os dados são apresentados tais como coletados na fonte, com maior nível de granularidade e sem agregação ou modificação.
- **Atuais:** os dados são disponibilizados tão rapidamente quanto necessário à preservação do seu valor.
- **Acessíveis:** os dados são disponibilizados para o maior alcance possível de usuários e para o maior conjunto possível de finalidades.

- Compreensíveis por máquina: os dados são razoavelmente estruturados de modo a possibilitar processamento automatizado.
- Não discriminatórios: os dados são disponíveis para todos, sem exigência de requerimento ou cadastro.
- Proprietários: os dados são disponíveis em formato sobre o qual nenhuma entidade detenha controle exclusivo.
- Livres de licenças: os dados não estão sujeitos a nenhuma restrição de direito autoral, patente, propriedade intelectual. Restrições sensatas relacionadas à privacidade, segurança e privilégios de acesso são permitidas.

Desde a criação da Política de Dados Abertos no Brasil (PBDA) (BRASIL, 2017), foram registrados aproximadamente 119 organizações e 5.400 conjuntos de dados. Para executar o plano de Dados abertos, é necessário utilizar as plataformas tecnológicas disponíveis para essa finalidade. Um exemplo é a *Comprehensive Knowledge Archive Network* (CKAN), uma aplicação *web-based* idealizada para publicação e gerenciamento de dados abertos (ASSOCIATION, 2015). Essa plataforma tem sido utilizada por governos de vários países, tais como Alemanha, Argentina, Austrália, Brasil, Canadá, Estados Unidos, México, Nigéria, Suíça e Uganda.

Com o intuito de acelerar as ações de transparência, os órgãos de controle também começaram a incentivar a criação do plano e operacionalização das plataformas de dados abertos. Como é o caso do TCU que publicou uma cartilha que traz os conceitos de dados abertos e informa cinco motivos para a abertura de dados na administração pública (TRIBUNAL, 2015):

1. Transparência na gestão pública;
2. Contribuição da sociedade com serviços inovadores ao cidadão;
3. Aprimoramento na qualidade dos dados governamentais;
4. Viabilização de novos negócios;
5. Obrigatoriedade por lei.

Com isso, a política da abertura de dados impulsiona a divulgação das informações governamentais oriundas das bases de dados dos órgãos públicos e fortalece a cultura da Transparência de Informações.

Segundo (ANGÉLICO, 2011), a transparência é um meio e não um fim em si mesmo, e o grau de utilidade da transparência depende da sua contribuição para um sistema efetivo de prestação de contas, ou seja, em que sejam disponibilizadas informações úteis e que possibilitem inferências.

A transparência e a disponibilização dos dados são fundamentais para a reusabilidade dos dados públicos. Desse modo, a co-produção de serviços tende a aumentar à medida que os cidadãos contribuem, desenvolvendo soluções inovadoras e inteligentes. Isso promove a participação e um novo olhar sobre a democracia.

A abertura de dados foi um importante propulsor da transparência de informações, que por sua vez introduz uma série de mudanças no cenário atual, alterando e incluindo elementos que agregam valor na esfera pública.

Um governo que se preocupa com a transparência de informações, possibilita que os cidadãos ajam com mais liberdade de ações através de políticas transversais, focadas em inclusão social e participação, e que contribua com a inovação dentro dos princípios da cidadania democrática.

O posicionamento do governo através do *Open Government Partnership* (OGP, 2011) foi registrar os compromissos firmados que tratam projetos com práticas de transparências, participação e inovação. O plano reforçou que os principais objetivos estão intimamente relacionados ao aumento da integridade pública e ao Sistema Federal de Acesso à Informação. Percebe-se que, através desse plano, é possível disponibilizar conteúdos de diferentes formatos, impulsionando a abertura de dados e o aumento da integridade de informações por meio de uma plataforma aberta.

O governo propôs a inclusão de ações referentes à participação social, cujo objetivo, no primeiro momento, foi apresentar à sociedade os futuros mecanismos de participação social e incentivar as esferas municipais e estaduais a iniciar seus processos de abertura de dados. Algumas ações foram executadas a fim de disponibilizar ferramentas para consulta da destinação do dinheiro/recursos públicos, além disso foi proposto um modelo de compras centralizado para aquisição de bens escolares e a melhoria no processo de compra e contratação de serviços de tecnologia da informação.

Após a inserção do Brasil no contexto de Governo aberto, algumas mudanças ocorreram na estrutura organizacional do Governo a fim de sustentar uma estrutura compatível com a complexidade e importância das competências, como as leis de Acesso à Informação (LAI, 2011), Conflito de Interesse (BRASIL, 2013b) e Anticorrupção (BRASIL, 2013a).

Em 28 de Maio de 2003, foi criada a Controladoria-Geral da União (CGU) com a finalidade de ser uma agência de anticorrupção nacional, com a função de promover a transparência da gestão e ser o centralizador dos sistemas de Controle Interno do Governo Federal. Em 2006, a CGU cria a Secretaria de Prevenção da Corrupção e Informações Estratégicas (SPCI) com a função de criar mecanismos para o combate à corrupção e, em 2016, a SPCI passa a ser chamada de Secretaria de Transparência e Prevenção da Corrupção (STPC). Em 2017, a CGU altera sua nomenclatura para Ministério da Transparência e Controladoria-Geral da União.

O principal programa instituído pela CGU foi o Brasil Transparente, oficializado pela Portaria nº 277 de 07 de Fevereiro de 2013. Segundo (UNIÃO, 2013), o cumprimento desse programa é obrigatório para todos os entes governamentais. A Lei de Acesso à Informação produz grandes impactos na gestão pública e exige, para sua efetiva implementação, a adoção de uma série de medidas que podem ser auxiliadas pela CGU, por meio do programa Brasil Transparente.

Ainda em (UNIÃO, 2013), os objetivos do Programa são:

1. promover uma administração pública mais transparente e aberta à participação social;
2. apoiar a adoção de medidas para a implementação da Lei de Acesso à Informação e outros diplomas legais sobre transparência;
3. conscientizar e capacitar servidores públicos para que atuem como agentes de mudança na implementação de uma cultura de acesso à informação;
4. contribuir para o aprimoramento da gestão pública por meio da valorização da transparência, acesso à informação e participação cidadã;
5. promover o uso de novas tecnologias e soluções criativas e inovadoras para abertura de governos e o incremento da transparência e da participação social;
6. disseminar a Lei de Acesso à Informação e estimular o seu uso pelos cidadãos;
7. incentivar a publicação de dados em formato aberto na internet;
8. promover o intercâmbio de informações e experiências relevantes ao desenvolvimento e à promoção da transparência pública e acesso à informação.

O Programa Brasil Transparente (UNIÃO, 2017), divulga, no sítio da CGU, um mapa com as informações referentes à adesão dos entes públicos participantes. Segundo dados

publicados no Portal da CGU, até Abril de 2018, foram 1806 adesões, sendo que 33% estão situadas na região Nordeste, 20% no Sul, 24% no Sudeste, 10% no Norte, 12% no Centro-Oeste, além de dez adesões em âmbito nacional. Neste mesmo sítio, é possível acompanhar o ranking de cumprimento da Lei de Acesso a Informação segundo os seguintes critérios de avaliação:

I - Regulamentação da Lei de Acesso:

- Exposição da legislação no site do avaliado;
- Existência da regulamentação;
- Regulamentação do SIC;
- Regulamentação da classificação de sigilo;
- Regulamentação da responsabilização do servidor;
- Regulamentação de instâncias recursais.

II - Transparência passiva:

- Divulgação do SIC físico (atendimento presencial);
- Existência de um e-SIC (atendimento pela internet);
- Possibilidade de acompanhamento do pedido de acesso;
- Inexistência de pontos que dificultem ou inviabilizem o pedido de acesso;
- Respostas aos pedidos no prazo legal;
- Respostas em conformidade com o que foi solicitado.

A Figura 3.3 apresenta o resultado do ranking por estados da federação brasileira e mostra a evolução na sequência de avaliações. A 1ª avaliação foi realizada no período de 31/03/2015 a 23/04/2015, a 2ª avaliação ocorreu entre 17/08/2015 e 03/11/2015 e a 3ª avaliação ocorreu entre 11/08/2016 a 04/01/2017.

Com a promulgação das leis que promovem a abertura e transparência de informações, o cidadão espera de fato entender e utilizar os mecanismos resultantes dos programas instituídos. Dentre as iniciativas criadas, destacam-se o Sistema Eletrônico de Recebimento e Gerenciamento dos Pedidos de Acesso à Informação (e-SIC). O acesso às informações

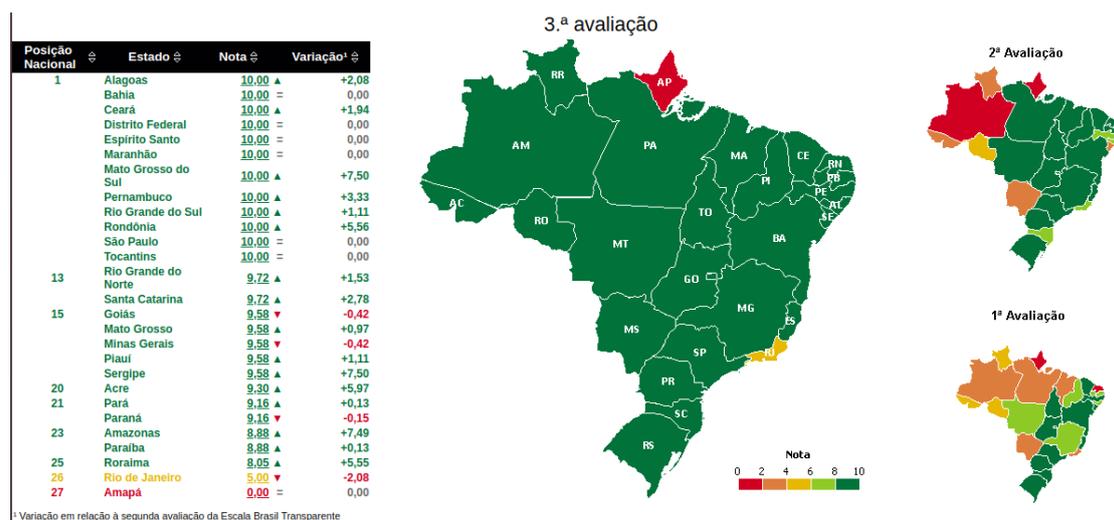


Figura 3.3: Ranking de cumprimento à Lei de Acesso a Informação (UNIÃO, 2017)

públicas é assegurado ao Cidadão por meio desse sistema, nos órgãos e entidade do poder público, assim como informar sobre a tramitação de documentos em suas respectivas unidades.

Dado que a Lei de Acesso à Informação estabelece que o acesso é a regra e que o sigilo é a exceção, através do e-SIC, qualquer cidadão poderá solicitar acesso às informações públicas, aquelas não classificadas como sigilosas, conforme procedimento que observará regras, prazos, instrumentos de controle e recursos previstos. O Cidadão pode fazer um pedido, acompanhar as respostas e prazos, fazer interposição de recursos, fazer sugestões, críticas e reclamações.

Outra ferramenta que respalda a transparência de informações é o Portal da Transparência. Nele é possível ter acesso a informações oriundas de bases de dados governamentais integradas: orçamento anual, receitas e despesas públicas, recursos transferidos, gastos por cartão de pagamento, áreas de atuação do governo, programas de governo, benefícios aos cidadãos, programas e ações orçamentárias, emendas parlamentares, estrutura de órgãos do governo, informações sobre servidores públicos, viagens a serviço, imóveis funcionais, licitações, contratações, convênios e outros acordos e sanções. O Portal de Transparência foi implantado em 2004 e suas informações são oriundas das bases de dados dos sistemas estruturantes do governo, dentre eles: o Sistema Integrado de Administração Financeira do Governo Federal (SIAFI), o Sistema Integrado de Administração de Recursos Humanos (SIAPE), as bases de benefícios sociais, as faturas de Cartão de Pagamentos do Governo Federal, as bases de imóveis funcionais, e diversas outras.

No dia 28 de Junho de 2018, o Ministério da Transparência lançou o Novo Portal da Transparência do Governo Federal. As mudanças focam na integração entre as consultas

e diversas bases de dados; ferramenta de busca; visualização e comparação em gráficos; melhor usabilidade; e linguagem cidadã, e o objetivo é estimular mais brasileiros a exercer o controle social e acompanhar como o recurso público federal está sendo utilizado.

Como consequência do processo de inovação aberta e transparência de informações, os dados abertos exercem influência no crescente movimento de co-produção, onde o cidadão e as organizações criam aplicativos e sistemas que utilizam esses dados, promovendo mais transparência de informações e aproximação do cidadão. Entendemos que essa mudança de comportamento permite mais controle social, por consequência, mais transparência, participação e propostas inovadoras.

### **3.1.4 Teoria do Voto - Modelos Cognitivos que Explicam o Comportamento Eleitoral**

O artefato a ser desenvolvido nesta pesquisa requer a seleção de um conjunto de informações prosopográficas a serem apresentadas. A necessidade de selecionar quais informações um cidadão deseja consultar sobre um político ou candidato, nos levou a investigar quais são as informações que influenciam o voto do cidadão. Identificar o que pode realmente interferir numa decisão de voto é uma tarefa desafiadora, pois os fatores que influenciam a escolha podem ser oriundos do(a) candidato(a), do grupo a que o candidato pertence e de suas parcerias; e, quanto ao eleitor, a escolha pode depender de sua história de vida e do espaço social em que está inserido.

Buscando entender essas influências, através de uma revisão bibliográfica, encontramos algumas explicações no trabalho de (RENNÓ, 2001). O autor relata que há uma vertente que explica os processos cognitivos de tomada de decisão e que está ligada a psicologia do voto. Nessa corrente, os sistemas de crença de massa mostram-se fortemente influentes. Existe também outra abordagem influenciada por conceitos oriundos da economia, que parte do princípio da racionalidade dos indivíduos, onde o objetivo é identificar os fatores que afetam o custo-benefício do voto. Sendo assim, essa última importa-se menos com a identificação dos modelos psicológicos da decisão.

Para compreender a vertente psicológica, recorreremos aos conceitos sobre estrutura das crenças de massas. No trabalho de (CONVERSE, 2006), o autor apresentou um paradigma no qual mostra que indivíduos, independentemente do ambiente social, distribuem-se e se relacionam com o mundo político de acordo com os níveis de conceitualização que são capazes de elaborar.

No artigo *The Nature of Belief Systems in Mass Publics – A Natureza dos Sistemas de Crenças na Massa* –, (CONVERSE, 2006), o autor enfatiza a importância dos aspectos

simbólicos na configuração dos fenômenos políticos. Por sistema de crenças, entende-se a configuração de ideias e atitudes nas quais os elementos estão juntos por alguma forma de coerção ou interdependência funcional. No sentido estático, é o grau de sucesso que teríamos em prever, sabendo de alguma atitude de uma pessoa, se a mesma terá certas ideias e atitudes compatíveis. Do ponto de vista dinâmico, refere-se à probabilidade de que uma mudança no status percebido de uma ideia-elemento leve a mudanças em ideias-elementos iniciais (LUZ, 2008).

Seguindo a perspectiva psicológica, (RENNÓ, 2001) relata que, dentre os intelectuais da Escola de Michigan, o psicólogo social Angus Campbell conjectura sobre o fato de os indivíduos não serem influenciados por aspectos ideológicos mais gerais. De acordo com (CAMPBELL et al., 1980), o que prevalece na memória do indivíduo são assuntos (temas) próximos de sua realidade e são esses assuntos que influenciam o voto.

São três os principais requisitos para que um tema (*issue*) venha a influenciar a decisão de voto. Primeiramente, esse tema deve ser cognizado, ou seja, os indivíduos têm que saber de sua existência. Segundo, o eleitor deve ter uma opinião em relação a ele. Finalmente, deve também poder identificar as posições dos diferentes candidatos sobre o tema (CAMPBELL et al., 1980). Em (RENNÓ, 2001), o autor conclui que, se algum desses requisitos não forem alcançados, então a influência de temas na eleição é muito reduzida, e outros fatores, como personalidade do candidato ou preferência partidária do eleitor, passam a ser preponderantes.

Ao nos aproximarmos das explicações acerca do comportamento eleitoral, estudamos o Modelo Cognitivo de Michigan e de Identificação do Partido, conforme apresentado na Figura 3.4. Este modelo foi proposto por cientistas políticos na década de 1950 no Centro de Pesquisas da Universidade de Michigan e procurava explicar o comportamento eleitoral em termos de apego psicológico de um eleitor a um partido político, que seria construído em um longo período de tempo.

De acordo com (HIMMELWEIT; HUMPHREYS; JAEGER, 1985), no Modelo de Michigan, a seleção eleitoral está diretamente dirigida pelas atitudes do eleitor em relação a três aspectos da política: candidatos, partidos e grupos sociais. Este modelo explica que as pessoas de um mesmo ambiente social, com atitudes e comportamentos parecidos, tendem a votar da mesma maneira.

Segundo (TAVARES, 2013), o Modelo de Michigan prevê que as orientações políticas começam muito antes de se adquirir idade para votar, sendo que a decisão de votar e a direção do voto refletem o ambiente social imediato das pessoas. Já (PERFEITO, 2007),

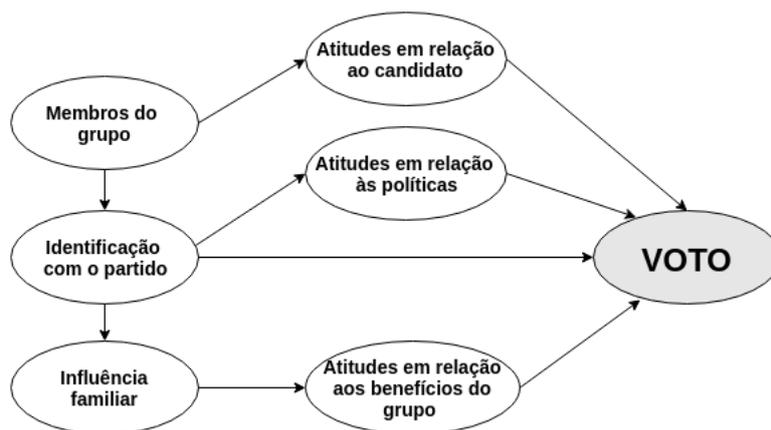


Figura 3.4: Modelo de Michigan (PERFEITO, 2007)

explicando que a proposta do modelo de Michigan tem a seguinte composição:

1. Muitos eleitores sentem fidelidade para com um partido, herdada de sua família;
2. A função da identificação do partido é permitir ao eleitor fazer frente a informação política e conhecer o partido em que vota;
3. A identificação com o partido se fortalece com o passar do tempo;
4. Os eleitores influenciados pelas forças que, em curto prazo, crescem dentro de uma campanha eleitoral, voltam depois ao partido com o qual se identificaram a priori;
5. A distribuição dos partidos dentro do eleitorado proporciona uma base através da qual se pode calcular o voto normal;

Para exemplificar a caracterização da identificação partidária, em (COSTA, 2016), a autora relata que, nos Estados Unidos, a identificação com o partido político pode ser medida com a seguinte escala unidimensional: 1 a 7, onde, (1) fortemente democrata, (2) simplesmente democrata, (3) pobremente democrata, (4) independente, (5) pobremente republicano, (6) simplesmente republicano e (7) fortemente republicano.

Outro modelo cognitivo muito observado pela área das ciências políticas, é o Modelo de Anthony Downs. Para Downs (DOWNS, 1957), os eleitores votam nos partidos que apresentam propostas políticas capazes de maximizar suas funções e encontram nos candidatos a solução para seus interesses individuais no futuro próximo (em termos de subsídios, saúde, educação, etc.) e não por encontrarem afinidades ideológicas. Nesse modelo, os eleitores escolhem os(as) candidatos(as) em função de quem melhor atenda suas preferências. O Modelo de Downs considera que os eleitores avaliam a utilidade que recebem

da política implementada pelo governante atual, bem como a utilidade que receberiam da oposição se esta pudesse estar no poder, e, por fim, calculam o diferencial dos partidos e votam no partido que maior utilidade lhes poderá proporcionar (ENELOW; HINICH, 1984). Tais modelos indicam algumas atitudes, opiniões, eventos e avaliações do governo que subsidiam a escolha do voto. Em (RENNÓ, 2001), o autor afirma que os eleitores avaliam o que seus representantes fizeram durante seus mandatos e o que eles podem vir a fazer se reeleitos. Nesse caso, a alternância de partidos experienciada pelo eleito é o que pode determinar o diferencial partidário. Este autor ainda discute que o nível de incerteza do eleitor, ao estimar o diferencial partidário, é definido pela presença de informações falsas e erros ou ignorância do eleitor. De acordo com (DOWNS, 1957), essa avaliação é condicionada pela informação que o eleitor dispõe sobre as preferências ideológicas dos candidatos e pela compatibilidade entre estas e as preferências do eleitor.

Outro modelo estudado nesta pesquisa foi o modelo de Thomas Holbrook, um modelo “on line” de avaliação dos candidatos que investiga o modo como a informação é disseminada. Em (BORBA, 2012), o autor explica que as campanhas eleitorais influenciam a opinião pública por meio da disseminação de informações que os eleitores consomem e utilizam para decidir em quem votar. O motivo pelo qual o termo “on line” é utilizado, deve-se ao fato do monitoramento das informações ocorrerem em tempo real e o seu processamento ser instantâneo. (LODGE; MCGRAW; STROH, 1989), um dos principais autores que dialogam sobre esse modelo, afirma que os eleitores iniciam o período eleitoral com uma “impressão” superficial dos candidatos. À medida que a campanha avança e novas informações são disseminadas, os eleitores alteram as suas percepções iniciais e formam novas opiniões sobre os candidatos – avaliações positivas no caso de informações positivas, ou avaliações negativas no caso de informações negativas.

Ainda em (BORBA, 2012), o autor destaca 3 variáveis que compõem o Modelo de Holbrook: eventos de campanhas, a cobertura eleitoral da mídia, e a situação do país. Esse modelo interpreta que essas variáveis, quando atuam mutuamente, promovem a propagação das informações e, por consequência, provocam a oscilação da opinião pública afetando o resultado das eleições.

Ao investigar os modelos cognitivos apresentados, surgiram algumas indagações sobre a adequação desses modelos na atualidade e no contexto brasileiro, sobretudo quanto às formas de interações, e as peculiaridades dos ambientes virtuais, alguns mais opacos ou não, outros mais suscetíveis a manipulações por agentes artificiais ou até mesmo impulsionados pelo comportamento humano. Os modelos de Michigan, Downs e Holbrook apresentados, foram criados há muito tempo (respectivamente em 1950, 1957 e 1996), não foram pensados considerando especificamente a cultura brasileira, além de não considerarem,

em tempos de cibercultura, as influências provocadas pelas tecnologias e ecossistemas digitais contemporâneos. Tais modelos foram importantes para o entendimento do comportamento diante de um processo de escolha, mas é preciso compreender as novas formas de interações e novos hábitos do(a) eleitor(a) brasileiro considerando o ciberespaço atual.

## **3.2 Quadro Técnico**

Esta seção apresenta o conjunto de conhecimentos técnicos considerados no desenvolvimento do artefato, e foi organizada em quatro subseções: conceituação sobre modelagem conceitual (subseção 3.2.1), conhecimentos sobre *Data Warehouse* (3.2.2), fundamentos sobre modelagem multidimensional (subseção 3.2.3) e por fim, na seção 3.3 são apresentadas as conjecturas comportamentais da pesquisa.

### **3.2.1 Modelagem Conceitual**

Importante salientar que esta pesquisa faz uso de modelagem multidimensional de dados, sendo importante discutir esse conhecimento técnico importante para a concepção da solução aqui proposta. A modelagem conceitual contribui para manutenção da base de conhecimento e sobretudo para distribuição e organização do conteúdo a fim de facilitar o entendimento do escopo e a gestão dos metadados que apoiam a concepção do artefato.

Para (GUIZZARDI, 2007), um dos principais fatores de sucesso por trás do uso de uma linguagem de modelagem está na capacidade da linguagem de fornecer aos seus usuários-alvo um conjunto de “primitivos de modelagem” que podem expressar diretamente conceitos de domínio relevantes, compreendendo o que nomeamos aqui de conceituação de domínio. Esse mesmo autor diz que os elementos que constituem uma conceituação de um dado domínio são usados para articular abstrações de certos estados de coisas na realidade, denominado abstração de domínio.

Tomemos como exemplo o domínio de operação de uma compra. A conceituação desse domínio pode ser construída elencando conceitos como pessoa, cliente, funcionário, vendedor, caixa, pedido, item de pedido, produto, pagamento, nota fiscal. Utilizando esses conceitos, podemos articular um abstração de domínio de certos fatos da realidade: um cliente chamado José faz um pedido de um produto e é atendido por um funcionário Pedro na função de vendedor e dirige-se ao caixa pra executar o pagamento. Ao final recebe a nota fiscal que contém o item de pedido adquirido.

Este último autor afirma que:

“Conceituações e Abstrações são entidades imateriais que só existem na mente do usuário ou comunidade de usuários de um idioma. Para serem documentados, comunicados e analisados, eles devem ser capturados, isto é, representado em termos de algum artefato concreto. Isso implica que uma linguagem é necessária para representa-los de forma concisa, completa e inequívoca” (GUIZZARDI, 2007, p.3, tradução nossa).

A Figura 3.5 apresenta o Triângulo de Ullmann, que representa a relação entre um objeto e seu conceito na mente do observador e uma possível representação em uma linguagem (ULLMANN, 1979).

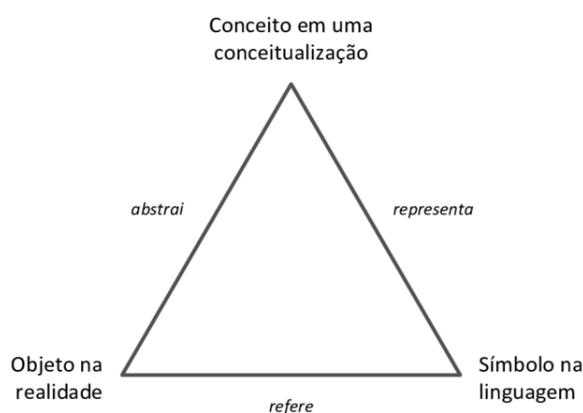


Figura 3.5: Triângulo de Ullmann (ULLMANN, 1979)

Sobre essa relação ilustrada no triângulo, e ao pensar na importância da cognição humana como fator importante na relação representação-realidade, (ABEL; FIORINI, 2013) considera que os estados internos de um sistema de computador, por exemplo, possam também corresponder a uma realidade externa, porém essa será ancorada através da percepção de uma pessoa que restringirá essa realidade e a representará em um sistema de símbolos também definido pela cognição humana.

Ao qualificar um modelo conceitual como semanticamente rico, (CASTRO, 2009) ressalta que a realidade deve ser expressa em uma linguagem que possua os construtos (símbolos) necessários à representação de todos os conceitos integrantes do domínio modelado, visto que esses conceitos abstraem a realidade (coisa) sobre a qual se deseja armazenar informações; em outras palavras, a linguagem que o representa deve ser adequada.

### 3.2.2 Data Warehouse

Por meio desta pesquisa, analisamos as informações oriundas de várias fontes de dados oficiais e, para nos apoiar, buscamos algumas definições sobre tecnologias que servem para armazenamento, consolidação e integração de dados. Antes de discutirmos sobre técnicas e ferramentas de modelagem multidimensional, numa perspectiva maior, incluímos, em nossa base de conhecimentos técnicos, conceitos sobre *Data Warehouse* (DW). Conforme explica (INMON, 2005), o DW é um conjunto de dados baseado em assuntos, ele é integrado, não-volátil e variante em relação ao tempo, utilizado para apoiar as decisões gerenciais. Logo, o DW deve atender aos requisitos do negócio. No princípio, a expressão representava simplesmente um armazém de dados, como é a tradução de DW; porém, ao longo do tempo, vem recebendo diversos incrementos em sua estrutura. Um DW é construído por meio de outros bancos de dados operacionais que podem estar implementados em diferentes plataformas na organização.

De acordo com (MOODY; KORTINK, 2000), o DW é baseado em uma metáfora da cadeia de suprimentos. O “produto” de dados é obtido de “fornecedores” de dados (sistemas operacionais ou fontes externas) e é temporariamente armazenado em um “armazém” central de dados.

Um DW tem por objetivo oferecer organização, gerenciamento e integração de bancos de dados, assim como ferramentas de exploração dos dados, para se obter vantagens competitivas no mercado. Para essa pesquisa, DW consolida o ambiente que abrigará os dados organizados conforme a solução que desenvolvemos nesta pesquisa.

Em seus estudos, (KIMBALL; ROSS, 2011) nos alertam sobre os cuidados que devemos tomar ao construir um DW, pois percebe-se uma certa confusão em relação a utilização, visto que algumas práticas consideram DW como cópias de sistemas de processamento transacional em uma arquitetura de hardware dedicada. Em (KIMBALL; ROSS, 2011), os autores reforçaram que esse fato merece atenção, pois os interesses dos usuários, as estruturas e os ritmos de um DW são muito diferentes, se comparado aos sistemas operacionais de registros. Este mesmo autor elencou alguns requisitos de DW que podem auxiliar os profissionais da área de negócios (o que, na presente pesquisa, respalda a preparação do ambiente que receberá o artefato concebido através do modelo dimensional), são eles:

- O DW deve permitir que as informações de uma empresa possam ser facilmente acessadas. Isso significa que os dados oriundos do DW devem ser compreensíveis. Em nosso contexto, os dados prosopográficos sobre os políticos devem ser intuitivos

para os usuários que buscam a informação. Não somente para os desenvolvedores, mas também para as pessoas que fazem uso de sistemas de informação que, por sua vez, precisam ser fáceis de serem manuseados. Trazendo esse requisito para o contexto de transparência de informações, essa propriedade está alinhada com as características de acessibilidade e entendimento.

- O DW deve apresentar as informações da empresa de modo consistente. Os dados devem oferecer confiabilidade. Deve haver um controle de qualidade, pois os dados podem ser obtidos de modo minucioso, de várias fontes de dados e filtros, e passam por um controle de suas liberações para o devido uso. Esse requisito tem similaridade com as características de transparência relacionadas a consistência e completeza.
- Adaptabilidade também faz parte do SIG de transparência. No contexto do DW, isso implica em dizer que as mudanças num artefato ou ambiente ocorrem eventualmente. No decorrer do tempo, solicitações de mudanças são feitas pelos usuários, os cenários e as tecnologia também são passíveis de alterações. Em (KIMBALL; ROSS, 2011), o autor orienta que os dados e os sistemas de informações não devem ser mudados nem danificados quando a comunidade de negócio fizer novas perguntas ou novos dados forem inseridos no DW.
- O DW deve ser um baluarte seguro que protege as informações. Ainda em (KIMBALL; ROSS, 2011), o autor reforça que o DW deve controlar, de modo eficaz, o acesso às informações confidenciais da empresa, e isso está associado à característica de controlabilidade.
- O DW deve funcionar como a base para uma melhor tomada de decisão. Um dos objetivos dessa pesquisa é a organização da informação para apoiar o cidadão na busca da informação, sobretudo em tarefas de análise e apoio a tomada de decisão, por exemplo, em um processo eleitoral. Por isso as características de transparência relacionadas à informativo e entendimento devem ser preservadas para atender o corrente requisito.
- Para que um DW seja considerado um sucesso, a comunidade de negócio deve aceitá-lo. Após o período de implantação e treinamento, os usuários devem continuar a utilizar o sistema de informação. Um trabalho de conscientização deve ser desenvolvido em prol de um bom resultado no teste de aceitação, para isso o fator simplicidade deve estar presente como característica de usabilidade.

O entendimento dos principais componentes de um DW se faz necessário para construção

das estruturas dimensionais. A Figura 3.6 mostra um ambiente de um *Data warehouse* composto por 4 segmentos.

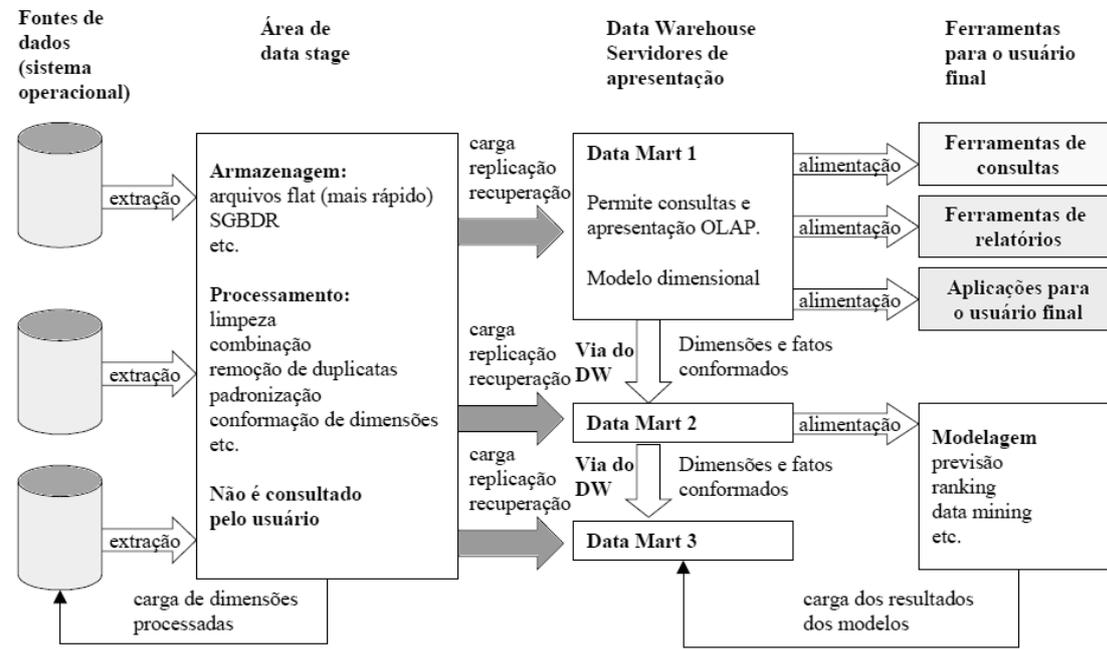


Figura 3.6: Elementos de um *Data warehouse* (KIMBALL; ROSS, 2011)

O primeiro componente de uma Arquitetura de DW é o conjunto de sistemas operacionais de origem que, de acordo com (MOODY; KORTINK, 2000), são sistemas que registram os detalhes das transações da empresa. As fontes externas também fazem parte desse conjunto, pois, eventualmente, dados de fontes externas, como por exemplo, dados do censo e dados econômicos, também constituem esses componentes.

O segundo componente de um DW é a Área de *data stage*, que consiste numa área onde ocorre o processo de ETL (*Extraction, Transformation e Load*). De acordo com (FERREIRA; CALDEIRA; OLIVAL, 2012), o processo de ETL é caracterizado como um subsistema de extração, transformação e limpeza dos dados dos sistemas operacionais no momento da sua disponibilização no DW. Esse autor afirma que o objetivo primordial é a extração dos dados dos sistemas transacionais, transportando-os para a área de estágio, onde ocorrerá a sua transformação e posteriormente o seu carregamento no DW.

Como mostra a Figura 3.6, na primeira etapa do processo ETL, os dados são extraídos de diferentes fontes de dados, consolidados e juntamente reconciliados e armazenados num formato consistente. Neste ambiente ocorre o processo de transformação dos dados, onde são verificados entradas e formatos inválidos, atribuição de chaves de DW.

Em (LOPES, 2017), o autor alerta que os dados extraídos e transformados precisam gerar informações confiáveis, caso contrário, as decisões serão baseadas em informações errô-

neas, podendo assim afetar diretamente a sobrevivência do negócio. Sobre atividades de transformação, por exemplo, é preciso verificar a não ocorrência de “-” ou “(“ em número de telefone, ou a ocorrência de espaços em branco em excesso em campo nome, ou os nomes de cidades escritos de várias formas, como “Belo Horizonte” e “BH”, “New York City” e “New York”. Outro exemplo é a verificação do “@” no endereço de e-mail e a definição de padrões para campos data, “AAAA/MM/DD” ou “DD/MM/AAAA”.

Algumas ferramentas apropriadas, de código aberto ou proprietário que auxiliam o processo de ETL, são recomendadas pelas suas capacidades de atender aos requisitos com eficácia e simplicidade nas atividades de manipulação: PDI (Pentaho BI Kettle), Power BI, IBM, Microsoft, Oracle. Na presente pesquisa, adotamos o Pentaho Data Integrator (PDI) Kettle, por ser uma ferramenta open source que possui os aplicativos de software necessários para desenvolvimento e execução dos processos de ETL.

Por conseguinte, a última etapa da Data Staging Area compreende o processo de load (carregamento dos dados): esse processo assume a forma de tabelas dimensionais (cubos dimensionais), que serão carregados em cada Data Mart (DM). O DM, de acordo com (JARKE et al., 2013), pode ser usado apenas em um departamento específico e contém apenas os dados relevantes para aquele departamento. Quanto à periodicidade do carregamento, em (MOODY; KORTINK, 2000), os autores relatam que as atividades de carregamento podem ocorrer de acordo com a necessidade de cada projeto: diariamente, semanalmente ou mensalmente.

Após a fase de carregamento, (KIMBALL; ROSS, 2011) apontam que os DM precisam indexar os dados para de fato executar as consultas e, após essa fase, a liberação dos novos dados deverá ser comunicada à todos os usuários sobre as mudanças ocorridas nas dimensões.

Em (RAINARDI, 2008), o autor descreve as abordagens para o processo de ETL:

“Poder consultar as fontes de dados de origens regularmente, e efetuar a extração, transformação e carregamento dos dados. A outra abordagem é através dos *triggers* (gatilhos), que vão estar associados a um evento. No momento em que o evento ocorrer, o *trigger* é acionado, sendo responsável pela inserção ou atualização dos dados numa determinada dimensão. E a última, menos utilizável, é através de um leitor de *log*, onde se efetua a leitura de *log* da base de dados para identificar alguma alteração nos dados, e em seguida obtém-se os dados modificados e armazena no DW” (RAINARDI, 2008, p.176, tradução nossa).

O próximo componente é formado pela Área de Apresentação dos Dados que trata da

obtenção dos dados. Para definir essa área, (SALVADOR et al., 2006) afirmam que, cada *Data Mart* é completamente organizado através de apresentação efetiva - modelos dimensionais.

Por fim, o último componente é representado pelas ferramentas de acesso a dados, ou seja, ferramentas utilizadas para a investigação e atividades de análise que incluem consulta ad hoc, geração de relatórios, ferramentas de análise final superior, e *data mining*. De acordo com (KIMBALL; ROSS, 2011), as ferramentas de acesso consultam dados da área de apresentação do DW e a grande maioria dos usuários provavelmente acessará os dados por meio de aplicações analíticas orientadas a parâmetros.

### **3.2.3 Modelagem Multidimensional**

Importante salientar que esta pesquisa, além de fundamentar-se na utilização da modelagem conceitual, tem seu foco na exploração do design de modelos de dados com características de múltiplas dimensões. Para sua construção, utilizamos os fundamentos da modelagem dimensional de dados.

A modelagem de dados multidimensional é uma pré-requisito para criação de um *Data Warehouse* (DW) ou *Data Mart*, nos quais geralmente resulta uma informação melhor, pois os dados normalmente são preparados antes de serem armazenados no DW ou DM. (DIAS et al., 1998).

Em (THOMSEN, 2002), o autor reforça que os bancos de dados multidimensionais são muito mais equipados para acesso a dados, cálculos rápidos e em estilo multidimensional voltado para sistemas de suporte à decisão. Isso está de acordo com a nossa proposta, já que a infraestrutura de tecnologia deve ser compatível com o modelo projetado e, dessa forma, ser capaz de lidar com dados projetados para serem orientados às necessidade analíticas e/ou decisão.

O termo dimensional está associado ao modo como os dados são organizados ou agrupados em um formato de cubo, no qual a ideia é ter várias dimensões de um tópico específico de discussão. Portanto, esse ambiente é profundamente diferente do ambiente transacional, cujas técnicas usadas para projetar bancos de dados não são adequadas para projetar um DW (KIMBALL; ROSS, 2011).

O cubo é uma forma abstrata para visualizar os dados, no entanto ele é limitado a 3 dimensões. (IBM, 2018) explica que um cubo é uma representação multidimensional das informações necessárias para processamento analítico on-line. O modelo dimensional sobrepõe uma estrutura de cubo em uma definição de contexto de monitoramento, que

define como as informações coletadas pelas instâncias de contexto de monitoramento são armazenadas.

Já (ZIULKOSKI, 2003) argumenta que um modelo dimensional não pode ser definido como um cubo e que é necessário estender o número de dimensões, com seus atributos e hierarquias, e em termos de suas medidas ou fatos que estão nos pontos de interseção das dimensões.

A figura 3.7 exemplifica a representação de um cubo que contém dados de vendas para diferentes produtos em diferentes anos e em diferentes cidades. Os números de vendas são os dados nas células do cubo. Vendas é a medida desse cubo. Produto, hora e cidade são dimensões.

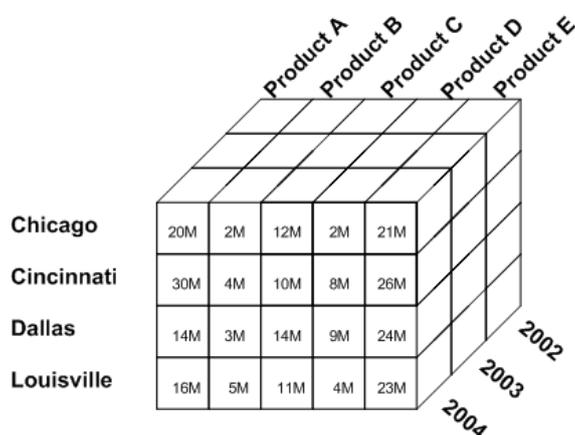


Figura 3.7: Exemplo de um cubo dimensional. (ORACLE, 2018)

Ainda em (ZIULKOSKI, 2003), o autor relata que um cubo possui algumas operações que são definidas da seguinte forma:

“A primeira operação é o *Slice and Dice* ou fatiamento do cubo, é a restrição das coordenadas nas dimensões de acordo com critérios definidos em cima de atributos das dimensões. Por exemplo: visualizar somente os dados dos mercados de pequeno porte. A segunda é a *Aggregation* ou agregação, que permite reduzir a dimensionalidade de um cubo ou de uma fatia de cubo. Por exemplo: ao visualizar a receita total de cada mercado por mês estamos agregando os dados de receita na dimensão produto, eliminando-a do cubo. E a terceira operação é a *Drill-up/Drill-down*, que consiste na navegação entre níveis de agregação, de acordo com hierarquias existentes nas dimensões. Exemplo: o usuário pode começar visualizando dados totais de receita para os mercados e fazer um *drill-up* para visualizar os dados totais por bairro, agregando os mercados próximos” (ZIULKOSKI, 2003, pg.22).

De acordo com (ABELLÓ; SAMOS; SALTOR, 2006), um modelo multidimensional é baseado em fatos (tabela fato), dimensões (tabela dimensões) e medidas. Intuitivamente, um fato representa o foco do objeto de dados da análise.

Em (REIS; TEIXEIRA; ARAÚJO, 2009), os autores explicam que os fatos são representados pela tabela central do modelo e contém os valores (numéricos) que se deseja analisar, geralmente, contendo um grande volume de dados. A tabela central, comumente denominada tabela fato, possui chaves externas que se relacionam com suas tabelas de dimensões, e possui campos numéricos que são os valores (medidas) que serão analisados;

As dimensões possibilitam várias perspectivas nas tarefas de análise e visualização dos dados. As tabelas de dimensões representam um aspecto do negócio que está sendo analisado. Sua chave primária serve para manter a integridade referencial na tabela fato à qual está relacionada (REIS; TEIXEIRA; ARAÚJO, 2009).

As medidas, por sua vez, são cálculos baseados em métricas. Uma medida aponta para uma métrica (como quantidade de pedidos) e destila as informações da instância em algo que seja útil, como a quantidade média de pedidos, a quantidade total de pedidos (IBM, 2018).

Um modelo de representação de dados multidimensional muito popular é o esquema em estrela ou *Star Schema*. Em (INMON, 2005) e (CHAUDHURI; DAYAL, 1997), os autores relatam que este modelo é utilizado sem muitos formalismos. De acordo com (KIMBALL; ROSS, 2011), para a construção de modelo desse tipo, é necessário que os dados para exploração estejam organizados num esquema constituído por uma tabela fato, e muitas tabelas de dimensões não desnormalizadas. (MCCABE et al., 2000), por sua vez, mostram que o esquema em estrela é adequado para aplicações de recuperação de informações. A figura 3.8 exemplifica um modelo *star schema*

No trabalho de (BALLARD et al., 2012), o autor explana sobre o esquema de floco de neve, onde uma dimensão é um floco de neve quando as colunas de baixa cardinalidade na tabela de dimensões foram removidas para separar as tabelas normalizadas e, em seguida, o *link* é criado novamente na tabela de dimensões original. A Figura 3.9 apresenta esse tipo de modelo.

Ao analisar um esquema estrela tradicional, existem casos onde há a necessidade de incluir um ou mais campos chaves, não oriundos das tabelas de dimensões. De acordo com (POE; BROBST; KLAUER, 1997), dessa maneira, consegue-se manter a chave composta da tabela a fim de identificar uma única linha da tabela fato. Esse modelo é denominado multi-estrela e é apresentado como exemplo na Figura 3.10. Observe que os campos data,

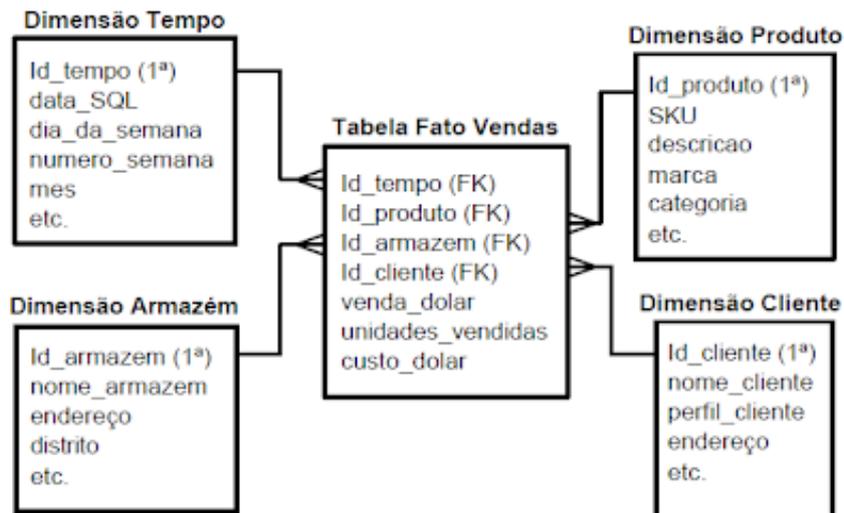


Figura 3.8: Modelo *Star-schema* Adaptado de (POE; BROBST; KLAUER, 1997)

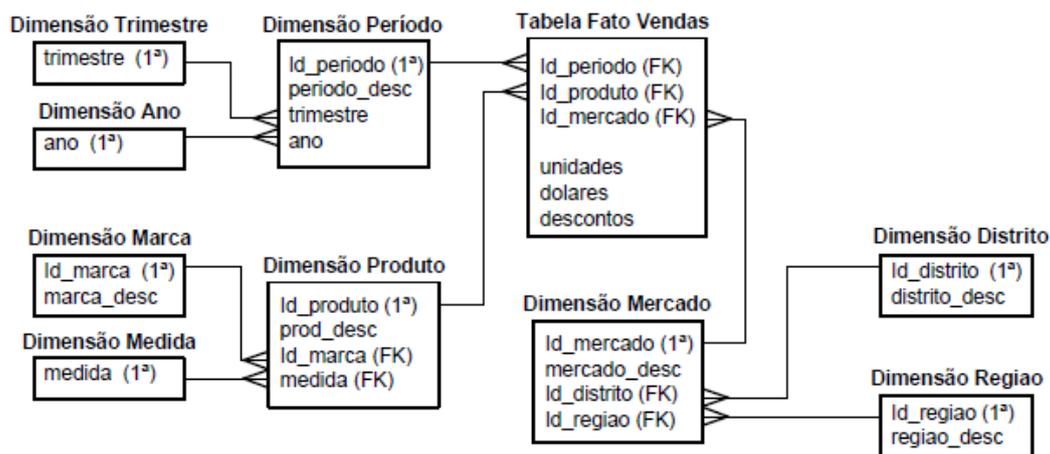


Figura 3.9: Modelo *Snowflake*. Adaptado de (POE; BROBST; KLAUER, 1997)

nr\_recibo e item\_linha\_recibo foram incluídos na chave composta para que o registro possa ser identificado.

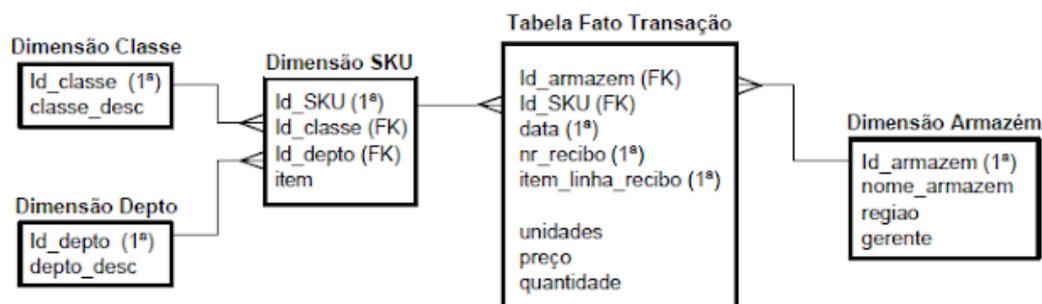


Figura 3.10: Modelo Multi-estrela. Adaptado de (POE; BROBST; KLAUER, 1997)

Nesta pesquisa, optei por usar o modelo de esquema em estrela, pois, de acordo com

(THOMSEN, 2002), apesar da redundância dos dados, o desempenho da consulta de dados para esse tipo de modelo será fornecido mais rapidamente e esse é um dos principais requisitos físicos para suporte a consultas analíticas.

As vantagens de utilizar o modelo *star-schema* para um processamento do tipo analítico, em detrimento ao de estrutura relacional, são destacadas nos trabalhos de (POE; BROBST; KLAUER, 1997) e (KIMBALL; ROSS, 2011) com os seguintes apontamentos:

- O projeto de banco de dados multidimensional provê rápido tempo de resposta;
- Permite otimizar o banco de dados para trabalhar com um projeto mais simples de banco de dados melhorando a execução do planejamento;
- Simplifica o entendimento e a navegação dos metadados por desenvolvedores e usuários finais;
- Permite projetar o banco de dados como o usuário final habitualmente pensa e usa os dados analiticamente; e
- Possibilita um maior número de ferramentas de acesso aos dados, sendo que alguns produtos disponíveis no mercado exigem o projeto de um esquema estrela.

O modelo star-schema é bastante utilizado para os projetos relacionados a banco de dados analíticos. Em (POE; BROBST; KLAUER, 1997), o autor sugere a verificação e/ou uso das variações do esquema em estrela a fim de obter o máximo de vantagens do ambiente de processamento analítico como, por exemplo, a integração e a qualidade dos dados, usabilidade e desempenho, dentre outras.

### **3.3 Conjecturas Teóricas**

Em (PIMENTEL; FILIPPE; SANTORO, 2018), os autores explicam que, em uma abordagem DSR, o artefato é projetado com base nas conjecturas teóricas fundamentadas na Ciência do Comportamento; e a avaliação do artefato fornece novos dados sobre as conjecturas elaboradas visando aumentar o conhecimento teórico.

Nesse sentido, na presente pesquisa foram elaboradas três conjecturas teóricas que nos levaram a projetar o artefato do jeito que o fizemos:

1. Cidadãos consideram que os conteúdos informacionais oriundos das bases de dados oficiais são confiáveis. Conforme será explicado na subseção 6.1.2, o artefato foi concebido com vistas a reconhecer um conjunto de dados oriundos de bases de dados governamentais e isso traz mais confiabilidade a quem consulta, pois estamos lidando com informações oficiais.
2. Cidadãos se interessam pelas informações sobre a vida dos políticos que o governo tem. Portanto, ao projetarmos o artefato, consideramos que seriam de interesse para os cidadãos, os dados pertencentes ao domínio dos sistemas que oficialmente oferecem apoio à organização do governo.
3. Cidadãos irão entender as informações prosopográficas sobre a vida dos políticos.

Foi considerando essas três conjecturas comportamentais, sobre a confiança, interesse e entendimento dos cidadãos, que projetamos a solução proposta nesta pesquisa, apresentada no Capítulo 5.

## 4. Conjecturas comportamentais: Fatores que Influenciam o Voto

Nesta pesquisa, queremos entender os fatores considerados relevantes num processo de busca de informações sobre a vida de uma pessoa que se apresenta como político, e o que é considerado influente para confiar ou decidir o voto. Fizemos um estudo exploratório sobre os fatores que influenciam o voto durante as eleições para presidente em 2018.

Conforme visto na Subseção 2.1.1.3, nossas referências foram oriundas em grande parte das ciências políticas e de conceitos de psicologia cognitiva. Para compreender tais comportamentos, considerando os fenômenos contemporâneos em torno do processo eleitoral, foi preciso repensar os modelos cognitivos referenciados. Através de depoimentos e aplicação de questionários, sob a ótica da *Grounded Theory* ou Teoria Fundamentada em Dados, foi concebido um modelo que contemplou fatores a serem considerados na concepção da solução proposta.

Alguns parâmetros são importantes para analisarmos um perfil social de um determinada pessoa, e no contexto desta pesquisa, uma pessoa política. Na literatura, foram encontrados muitos trabalhos que traçam o perfil social da classe política. (COSTA; CODATO, 2013) informam que, dentre os atributos que compõem um perfil social, temos: origem de classe, tipo e tamanho do patrimônio herdado ou construído, acesso a educação superior e posse de títulos escolares. Classicamente, os estudos sobre a composição social das casas parlamentares analisam o nível de instrução formal dos parlamentares.

Conforme explica (ARAÚJO, 2011), o grau de instrução também é um aspecto de relevância ao pensar no recrutamento parlamentar. Já (POLSBY, 1968) afirma que a profissionalização da política, a competitividade da vida parlamentar e a complexidade das atividades desenvolvidas pelo Legislativo nas democracias contemporâneas aumentam a exigência por níveis de formação elevados, bem como suas habilidades profissionais.

Sobre o gênero, este parâmetro merece atenção quanto ao grau de acessibilidade por gê-

neros no cenários político, pois é bem reduzido o quantitativo de mulheres na carreira política. Conforme (IBGE, 2018), em 2018, o percentual de parlamentares mulheres foi de 16% e de deputadas federais foi de 10,5%, e, em relação às eleições de 2014, o percentual de candidatas foi de 29,3%.

Ainda em (COSTA; CODATO, 2013), também são parâmetros que fazem parte da necessidade analítica da cidadão: a idade de ingresso na política, o número de mandatos antes de chegar a posições superiores na hierarquia política, quantidade de partidos por que passou, cargos estratégicos que dirigiu e etc.

Os modelos cognitivos são importantes para apoiar a identificação de quais são os dados importantes que deveriam constar no modelo a ser elaborado para a presente pesquisa, pois consideramos que os dados mais importantes são os que têm maior influência sobre o comportamento do cidadão diante de um julgamento ou análise de um perfil político. Para complementar esse entendimento, foi realizado um estudo exploratório durante as eleições para presidente em 2018, cujo objetivo foi identificar quais os fatores que influenciam a decisão de voto. Claro que esse estudo seria influenciado pelo contexto daquela eleição em particular.

Conforme apresentado na subseção 3.1.4, os fundamentos teóricos foram oriundos, em grande parte, das ciências políticas e de conceitos de psicologia cognitiva. Para compreender os comportamentos dos eleitores brasileiros, considerando os fenômenos contemporâneos em torno da eleição para presidente de 2018, foi preciso repensar os modelos cognitivos referenciados. Os depoimentos dos eleitores, obtidos a partir da aplicação de questionários, foram interpretados utilizando técnicas da Teoria Fundamentada em Dados, o que resultou em um modelo sobre os fatores a serem considerados na concepção da solução proposta.

O propósito da Teoria Fundamentada em Dados é a possibilidade de imersão mais aprofundada nos dados, trazendo os constructos para a elaboração teórica. Percebemos que a utilização desse método nos permite uma aproximação e análise dos dados de maneira mais antecipada no processo de pesquisa, diferentemente da pesquisa quantitativa onde a aproximação é realizada numa etapa posterior.

O estudo exploratório foi realizada em 3 etapas, cada uma constituindo de um questionário enviado em formato digital para um grupo de participantes. A forma de condução de cada uma das etapas, bem como os resultados encontrados, são descritos a seguir.

#### 4.1 Primeira Etapa: levantamento dos fatores de influência do voto

A primeira etapa deste estudo exploratório ocorreu durante a semana que antecedeu ao primeiro turno das eleições presidenciais de 2018, ou seja, um período bastante oportuno para discutir o tema desta pesquisa. Nessa fase, a pesquisa foi realizada através do envio de uma mensagem por *e-mail* e também pelo aplicativo WhatsApp contendo as seguintes perguntas:

- Quais as razões que o/a levaram a escolher um(a) determinado(a) candidato(a) ?
- O que influenciou/influencia a sua escolha?

A mensagem foi enviada para 91 pessoas que fazem parte do meu círculo social e também do círculo social dos orientadores desse trabalho. Consideramos que seria adequado perguntar para pessoas conhecidas nossas porque esse tema é muito sensível (ao menos na época em que realizamos esta pesquisa) e que estranhos não se sentiriam confortáveis em respondê-las; além disso, sendo pessoas conhecidas, esperávamos ter respostas sinceras e que teríamos a intimidade para poder aprofundar as respostas. Em resposta ao *e-mail*, obtivemos 37 depoimentos. Para organizar os depoimentos, foi criado um relatório único juntando todas as transcrições (cada uma identificada pelo nome da pessoa que deu o depoimento). Em seguida, as respostas foram interpretadas a fim de extrair os fatores de influência.

Nos procedimentos da análise do discurso, procuramos interpretar os sentidos possíveis na fala do entrevistado, não deixando de considerar o sujeito, sua história de vida, o contexto social em que ele pertence, sua ideologia e seus sentimentos. A Figura 4.1 apresenta um trecho do depoimento de um dos participantes. Nos comentários estão os códigos que atribuímos para o que interpretamos ser fatores que o entrevistado aponta ser influente em sua decisão de voto.

Ao final da interpretação do discurso, foram identificados 95 fatores considerados influentes no processo de escolha de um candidato. Após a seleção e classificação de todos os fatores, foi realizado um agrupamento dos mesmos, com base na semelhança dos significados e do contexto no qual cada fator está inserido. Cada agrupamento representa uma macro-categoria teórica, que iremos considerar como sendo fatores de influência. São eles:

- Comportamento do(a) Candidato(a)
- Propostas de Governo

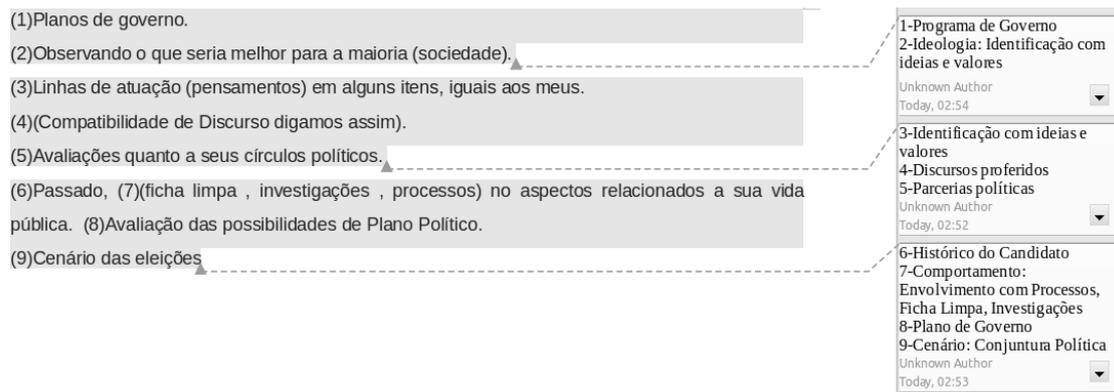


Figura 4.1: Exemplo da interpretação de um depoimento

- Identificação Ideológica
- Experiência do(a) Candidato(a)
- Participação do Eleitor
- Posicionamento Político do Eleitor
- Fontes de Informações relacionadas às Mídias (rádio e TV) e Redes Sociais Digitais
- Fontes de Informações relacionadas às pesquisas eleitorais
- Sentimentos do Eleitor
- Influência Social

Um grupo é representado pelo Comportamento do(a) Candidato(a). Este grupo contém os fatores relacionados à observação da participação do(a) candidato(a) dentro e fora do período eleitoral, a maturidade frente ao movimento político, o estilo de vida, o envolvimento em processos de improbidade administrativa, casos de corrupção ou processos investigativos. Observa-se, também nesse grupo, o discurso proferido do(a) candidato(a), a maturidade intelectual e as relações entre as parcerias formadas.

Outro grupo traz as Propostas de Governo, que concentram os fatores que fazem referência às propostas que tratam: do estatuto de desarmamento, das necessidades da população, de prioridades para a educação, saúde, redução de maioridade penal, inovação na gestão e fortalecimento do estado.

O grupo Identificação Ideológica compõe os interesses nas ideias e valores do eleitor em relação ao candidato ou partido, no que diz respeito aos campos de atuação política. Já o grupo Experiência do Candidato, inclui as informações sobre o histórico da vida profissional e política do(a) candidato(a), o seu grau de instrução, o histórico de atuação

no partido e seu desempenho quanto a elaboração e aprovação de projetos durante seu mandato.

O grupo Participação do(a) Eleitor(a) considera a estratégia de voto, o voto protesto e o voto por eliminação como um fator de influência. Por conseguinte, temos o grupo Posicionamento Político do(a) eleitor(a) que trata da concordância ou não com os candidatos do governo atual, com candidatos novos e observa as articulações políticas para permanência de candidatos no poder. Este grupo também inclui o posicionamento do eleitor no espectro político, econômico, social, bem como sua orientação política (se é de direita ou de esquerda).

Um grupo de fatores está relacionado às Fontes de Informação. Dividimos esse grupo de influência em duas partes; as Fontes de Informações relacionadas às Mídias (rádio e TV) e Redes Sociais Digitais (Facebook, Twitter, Instagram, Whatsapp ); e as Fontes de Informações associadas às Pesquisas Eleitorais. Consideramos que as mídias e redes sociais são englobadas como uma questão de acesso à informação, e o conteúdo das fontes sobre as pesquisas eleitorais pertencem a um espectro diferente da informação. Quando se trata de uma pesquisa de opinião, cabe observar que a metodologia divulgada é bastante sucinta, limitando-se à indicação do instituto que efetuou a pesquisa, o número do registro (obrigatório) do levantamento no Tribunal Eleitoral, a data em que o trabalho de campo foi realizado, a quantidade de entrevistados e em quantos municípios eles se encontravam (MIGUEL; MOTA, 2011).

Outro grupo é representado pelos Sentimentos do Eleitor. Pelo relatos coletados, identificamos os sentimentos de insatisfação, confiança, insegurança, afeto, paixão, repulsa, indignação, desesperança, emoção, etc.

Outro grupo é representado pela Influência Social que inclui a opinião de amigos, parentes, colegas de trabalho, a sociedade patriarcal, a influência regional e demais grupos sociais do indivíduo.

Vale informar que os fatores aqui identificados foram obtidos de minha interpretação sobre o que os voluntários da pesquisa disseram, ou seja, esta é uma pesquisa baseada na subjetividade dos indivíduos. Em (ROLNIK; GUATTARI, 2006), o filósofo Félix Gattari compreende a subjetividade como um processo de produção no qual comparecem e participam múltiplos componentes que resultam da apreensão parcial que o humano realiza, permanentemente, de uma heterogeneidade de elementos presentes no contexto social. Logo, não queremos chegar numa “verdade absoluta” sobre a influência na decisão do eleitor, e sim queremos construir uma narrativa sobre o que as pessoas-respondentes acre-

ditam ser influentes.

Segundo (MANSANO, 2018), a reunião dos componentes que formam a subjetividade possibilita fazer conexões díspares e inesperadas, precipitando movimentos que insistem em suas misturas e desvios. Dessa forma, há uma reinvenção e desuso recorrente dos componentes, cada um com seu valor e duração históricos.

Com o propósito de ilustrar os conceitos e organizar os 95 fatores, através dos agrupamento aqui descritos, foi possível a criação de um modelo conceitual que nominamos de Fatores que Influenciam o Voto no Brasil em 2018 (FIV.Br2018). A versão final do modelo está apresentada na 3ª Etapa (subseção 4.3), pois cada fase do estudo exploratório permitiu que fossem introduzidos elementos a fim de refinar e explicar melhor a composição do modelo FIV.Br2018.

## **4.2 Segunda Etapa: avaliação dos fatores**

Essa etapa ocorreu no período entre os 2 turnos das eleições presidenciais de 2018. O objetivo aqui era identificar quais são os fatores considerados mais influentes, dentre os identificados na Etapa 1. Para avaliar a adequação dos fatores identificados na etapa anterior, elaboramos um questionário, conforme o modelo em Apêndice A, contendo todos os elementos e devolvemos, aos mesmos participantes, com a seguinte pergunta:

- No processo para escolher um(a) candidato(a), dentre os fatores listados abaixo, em que nível ou aproximação, eles influenciam a sua decisão de voto?

Ao responder essa pergunta, os participantes avaliaram cada fator do modelo, pontuando um grau que o participante considera para a influência que cada fator tem sobre a sua decisão de voto. Para isso, eles utilizaram a seguinte escala nominal: “Não influencia”, “Influencia Pouco” e “Influencia Muito”.

Aguardamos uma semana após o envio do questionário, e realizamos a análise dos dados que coletamos como resposta ao questionário obtidos até aquela data. Para a análise dos resultados, convencionamos um mapeamento da escala nominal para uma pontuação definida arbitrariamente como sendo 0 (“não influencia”), 1 (“influencia pouco”) e 2 (“influencia muito”). Para estabelecer uma comparação entre os fatores, utilizamos uma média ponderada com os valores 0, 0.5 e 1.0.

A tabela 4.1 apresenta o ranking dos 10 fatores mais influentes e os 10 fatores menos influentes, segundo a opinião dos participantes.

Tabela 4.1: 10 Fatores mais influentes e menos influentes .

10 Fatores mais influentes	10 Fatores menos influentes
1º Programa de governo proposto	1º Influência de políticos na família
2º Acesso às informações dos candidatos	2º Sociedade patriarcal
3º Comportamento político	3º Influência Regional
4º Conjuntura política	4º Voto protesto
5º Ficha Limpa	5º Preferência de voto ou opinião de amigos
6º Identificação com as ideias e valores	6º Rejeição ao candidato
7º Insatisfação com a realidade da educação, com os desafios de ser professor e suas condições de trabalho	7º Preferência de voto ou opinião de parentes
8º Insatisfação com o a situação atual	8º Voto por eliminação.
9º Entendimento e interesse pelo prestação de serviços públicos	9º Voto parental
10º Envolvimento em processos de corrupção	10º Vota-se por ordem de outrem

Dando continuidade ao trabalho, e por entender que a pesquisa precisava ter maior alcance naquele período, foi incluída a terceira etapa, com algumas modificações no formato e dimensão do questionário. Na próxima seção, detalho como as atividades foram realizadas e como os resultados contribuíram para a evolução e apresentação do modelo FIV.Br2018.

### 4.3 Terceira Etapa: avaliação da influência do agrupamento de fatores

A terceira etapa da pesquisa foi realizada através de um questionário estruturado com perguntas fechadas de identificação, sobre as percepções e atitudes a respeito dos grupos de fatores de influências sobre a decisão de voto, gerados no final da segunda etapa – (APÊNDICE A).

Nesta etapa, as avaliações foram realizadas por meio de 4 níveis de respostas: “Não influencia”, “Influencia Pouco”, “Influencia Muito” e “Influencia Totalmente (Decisivamente)”. Para cada nível, convencionamos os respectivos valores 0, 1, 2 e 3 a fim de sumarizar as avaliações dos respondentes. Diferentemente da fase 2, o questionário foi criado com base no agrupamento dos fatores. A pergunta apresentada na Figura 4.2 exemplifica esta versão do questionário aplicado nessa terceira fase da pesquisa.

Um link para preenchimento do questionário confeccionado, a partir do programa Google Docs, foi enviado para os meus contatos e os contatos dos orientadores, por meio das interfaces de mensagem instantânea dos sistemas Facebook, LinkedIn e WhatsApp.

O objetivo, nessa etapa, foi coletar o maior número de respostas possíveis de maneira democrática e aberta a todos(as). Nesta etapa, além das redes sociais de amplo uso citadas, tivemos a especial contribuição do Espaço UFF-Pesquisa, que é coordenado pela Superintendência de Comunicação Social da Universidade Federal Fluminense (UFF).

No processo para escolher um(a) candidato(a), quais os fatores que mais influenciam a sua decisão de voto? Informe o grau de influência que cada um dos fatores a seguir tem na sua decisão de voto.

12- Os interesses pelas ideias e valores do candidato ou partido, exercem qual nível influência em sua decisão de voto? Estes fatores estão relacionados à Identificação Ideológica

- Não influencia
- Influencia pouco
- Influencia muito
- Influencia totalmente (Decisivo)
- Não quero responder

Figura 4.2: 3ª Etapa: questões sobre o agrupamento de fatores

Esse espaço é destinado à divulgação das pesquisas desenvolvidas na comunidade acadêmica, através de mensagens enviadas ao *e-mail* institucional de todos os alunos, docentes e servidores técnico-administrativos da Universidade.

O questionário foi disparado para aproximadamente 79 mil pessoas, e permaneceu habilitado para respostas durante 15 dias: ao longo da semana que antecedeu o segundo turno das eleições presidenciais de 2018 e ao longo da semana que sucedeu este evento. Após esse período, encerramos o recebimento de respostas, tendo sido coletado um total de 1305 respostas. A tabela 4.2 apresenta o resumo estatístico em função da distribuição sócio-demográfica da pesquisa.

Em relação aos aspectos sócio-demográficos, dos 1305 respondentes, 50,6% eram mulheres, com maior ocorrência na faixa etária entre 18 e 24 anos (39,9%), a maioria delas de pele branca (62%) e a maior parte com renda entre 1 a 3 salário mínimos (27%) e sem religião (36,6%). Se compararmos os indicadores de Gênero e Cor/Grupo Étnico aos valores relativos do Censo demográfico de 2010 (IBGE, 2018), respectivamente (52,3%) e (48,6%), a pesquisa apresentou um cenário aproximado quanto a participação das mulheres, porém distante quanto a faixa etária de 18 a 24 anos (12,24%) e quanto a religião (13%) .

A pesquisa teve uma forte divulgação e, conseqüente, um grande número de respondentes na UFF, que fica localizada em Niterói. Como esse é um município da região metropolitana do Rio de Janeiro, espera-se que algumas distribuições apresentassem resultados

Tabela 4.2: Distribuição Sócio Demográfica: gênero, idade, orientação sexual, concordância entre gênero e sexo, cor/etnia, estado, localização, grau escolaridade, profissão, renda bruta e religião.

Variável	Frequência		Variável	Frequência	
	Absoluta (n)	Relativa (%)		Absoluta (n)	Relativa (%)
<b>Gênero</b>			<b>Localização</b>		
Homem	616	48,1	Zona Rural	61	4,8
Mulher	660	50,6	Centro Urbano	1082	85,3
Não binário	4	0,3	Periferia	125	9,9
Não Informados	25	1,3	Comunidade Indígena	0	0,0
<b>Idade (anos)</b>			Comunidade Quilombola	1	0,1
18 a 24	460	37,0	Não Informados	36	2,8
25 a 29	226	18,2	<b>Grau de Escolaridade</b>		
30 a 34	148	11,9	Sem Educação Formal	0	0
35 a 39	111	8,9	Ensino Fundamental Incompleto	0	0
40 a 44	82	6,6	Ensino Fundamental Completo	4	0,3
45 a 49	65	5,2	Ensino Médio Incompleto	2	0,2
50 em diante	152	12,2	Ensino Médio Completo	62	5,1
<b>Orientação Sexual</b>			Ensino Superior Incompleto	535	44,4
Heterossexual	1035	82,4	Ensino Superior Completo	313	26,0
Homossexual	105	8,4	Mestrado Incompleto	132	11,0
Bissexual	107	8,5	Mestrado Completo	60	5,0
Panssexual	5	0,4	Doutorado Incompleto	48	4,0
Assexual	2	0,2	Doutorado Completo	49	4,1
Demisssexual	1	0,1	Não Informados	100	7,7
Interssexual	1	0,1	<b>Renda Mensal</b>		
Não Informados	49	3,8	Até 1 salário mínimo	126	11,3
<b>Concordância Gênero e Sexo Biológico</b>			De 1 a 3 salários mínimos	297	26,6
Cisgênero	1085	97,0	De 12 a 15 salários mínimos	45	4,0
Transgênero	7	0,6	De 3 a 6 salários mínimos	216	19,4
Não entenderam	26	2,3	De 6 a 9 salários mínimos	112	10,0
Não Informados	187	14,3	De 9 a 12 salários mínimos	91	8,2
<b>Cor/Etnia</b>			Mais de 15 salários mínimos	49	4,4
Branca	787	61,5	Nenhuma renda	179	16,1
Preta	130	10,2	Não Informados	143	11,4
Amarela	14	1,1	<b>Religião</b>		
Parda	345	27,0	Acredito em Deus	3	0,1
Indígena	3	0,2	Agnóstico	9	0,4
Não Informados	26	2,0	Ateu	11	0,4
<b>Estado</b>			Budista	4	0,2
Alagoas (AL)	1	0,1	Candomblé	3	0,1
Amazonas (AM)	5	0,4	Católica	353	14,1
Bahia (BA)	8	0,6	Crença Sincrética	1	0,04
Ceará (CE)	5	0,4	Cristão	9	0,4
Distrito Federal (DF)	16	1,2	Eclética	1	0,04
Espírito Santo (ES)	5	0,4	Esotérica	1	0,04
Goiás (GO)	2	0,2	Espírita	141	5,6
Mato Grosso (MT)	2	0,2	Espiritualista	2	0,1
Mato Grosso do Sul (MS)	3	0,2	Evangélica	207	8,3
Minas Gerais (MG)	65	5,1	Hindu	1	0,04
Pará (PA)	16	1,2	Indefinido	1	0,04
Paraná (PR)	3	0,2	Judaica	4	0,2
Pernambuco (PE)	3	0,2	Messiânica	1	0,04
Rio de Janeiro (RJ)	1099	85,4	Ortodoxo	1	0,04
Rio Grande do Norte (RN)	3	0,2	Pagão	1	0,04
Rio Grande do Sul (RS)	14	1,1	Politeísta/Neopaganismo	1	0,04
Rondônia (RO)	1	0,1	Satanista	1	0,04
Santa Catarina (SC)	5	0,4	Sem religião	458	18,3
São Paulo (SP)	26	2,0	Testemunha de Jeova	2	0,1
Sergipe (SE)	3	0,2	Umbanda	26	1,0
Tocantins (TO)	2	0,2	Wicca	1	0,04
Não Informados	18	1,4	Não Informados	62	2,4

com influência de um grupo com características comuns entre si. Apesar da pesquisa ter alcançado 77% dos estados brasileiros, a maior parte dos respondentes (85,4%), residem no estado do Rio de Janeiro, seguido de Minas Gerais (5,1%) e São Paulo (2%). No que se refere à localização dos participantes, 85,3% residem em centros urbanos, 9,9% em periferia e 4,8% em zona rural, enquanto que no cenário Brasil, conforme (IBGE, 2018), isso representa 84,36% para localização urbana e 15,64% para zona rural.

Quanto ao grau de escolaridade, devido o grande número de respondentes oriundos de um

ambiente universitário, o maior número dos participantes tem nível superior incompleto (44,4%), seguidos por ensino superior completo (26%) e mestrado incompleto (1%).

A Figura 4.3 reúne um conjunto de gráficos que representam as percepções dos participantes acerca de cada fator de influência.

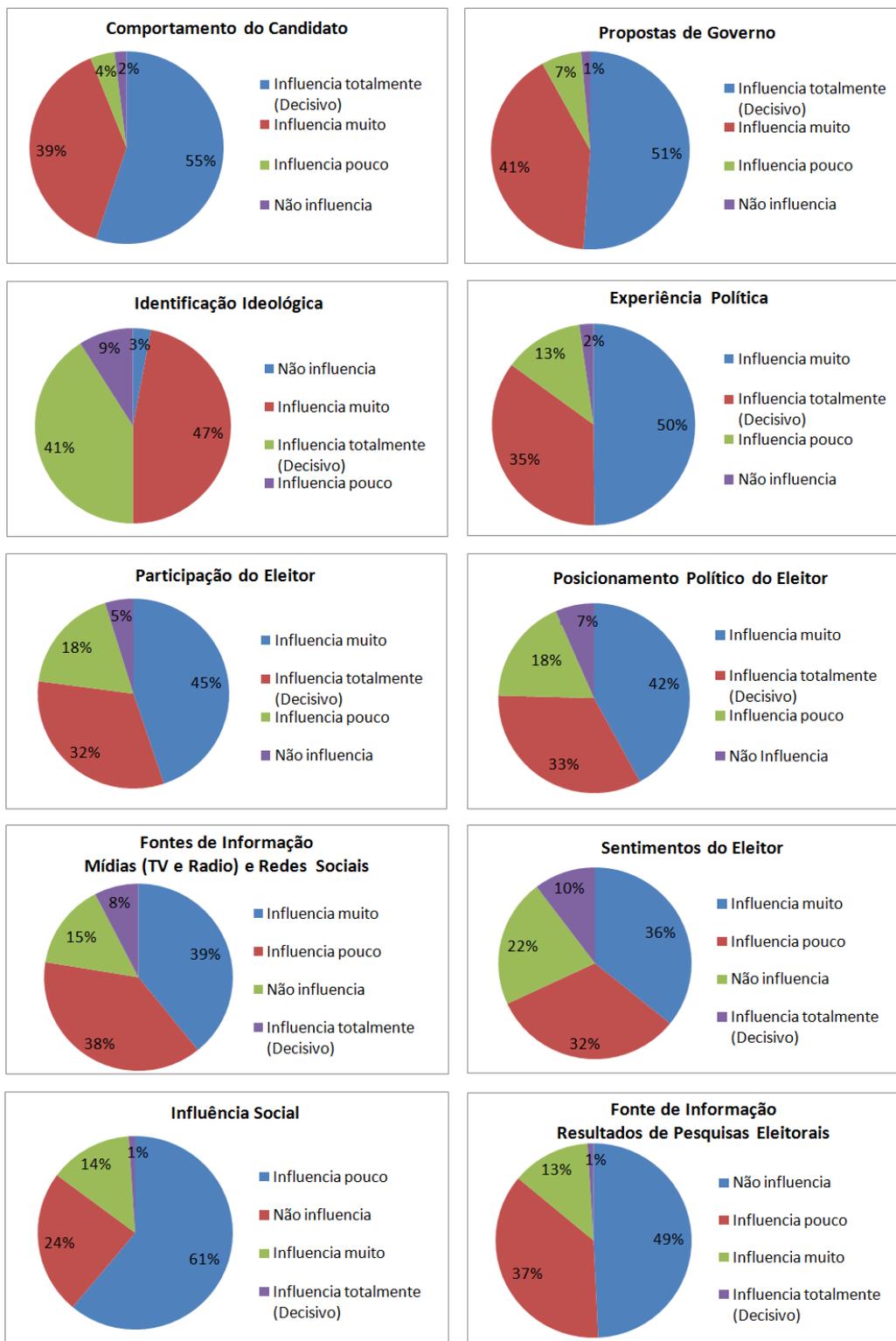


Figura 4.3: Terceira Etapa - Estatísticas dos Fatores de Influência

Como podemos observar, o Comportamento Político foi o fator considerado mais influente, sendo que 55% dos participantes mencionaram esse fator como totalmente decisivo. Em relação às Propostas de Governo, 51% dos respondentes acreditam que este fator influencia muito. Sobre o fator Identificação Ideológica, 47% das pessoas disseram que influencia muito. Em relação à Experiência Política do Candidato, 50% dos respondentes acreditam que influencia muito. Sobre à participação do eleitor, a percepção foi de 45% para influencia muito. Na distribuição da amostra no que diz respeito ao fator Posicionamento Político do eleitor, 42% dos respondentes consideraram que este fator influencia muito. Para o fator Fontes de Informações relacionadas a mídias (rádio e TV) e redes sociais, 39% consideram que o fator influencia muito. A opinião em relação ao grupo Sentimento do Eleitor, indicou que 36% acreditam que influencia muito. Já para o grupo Influência Social, 61% das pessoas afirmaram que se trata de um fator pouco influente. Por fim, sobre o grupo Fontes de Informações sobre Resultados de Pesquisas Eleitorais, 49% dos respondentes acreditam que este fator não influencia.

Ao analisarmos os dados, obtivemos uma métrica, num *score* de 0 a 3, para cada fator, com base nas médias aritméticas das avaliações e obtivemos a classificação em ordem decrescente de grau de influência. Os resultados mostraram que, de acordo com as percepções do grupo, o fator “Comportamento do candidato” foi considerado o mais influente e o fator “Resultados das pesquisas eleitorais” foi o menos influente. A tabela 4.3 apresenta o ranking das avaliações nessa fase.

Tabela 4.3: Ranking das avaliações sobre os grupos de fatores (3ª Etapa)

Grupo de Fatores	Média	Desvio Padrão
1º – Comportamento do Candidato(a)	2,47	0,67
2º – Propostas de Governo	2,42	0,68
3º – Identificação Ideológica	2,26	0,74
4º – Experiência do Candidato(a)	2,18	0,73
5º – Participação do Eleitor	2,05	0,83
6º – Posicionamento Político do Eleitor	2,02	0,88
7º – Mídia (rádio e TV) e Internet (Redes Sociais)	1,40	0,83
8º – Sentimentos do Eleitor	1,35	0,93
9º – Influência Social	0,92	0,64
10º – Resultados de Pesquisas Eleitorais	0,66	0,74

A partir dessa classificação, foi consolidado o Modelo dos Fatores que Influenciam o Voto (FIV.Br2018), apresentado na Figura 4.4. No modelo, os grupos de fatores estão dispostos de baixo para cima de acordo com o grau de influência (os mais influentes aparecem primeiro). Cada grupo de fatores possui um subconjunto de elementos, também

classificados em ordem decrescente por seus pesos<sup>1</sup>, que foram calculados na fase 2 do estudo exploratório.

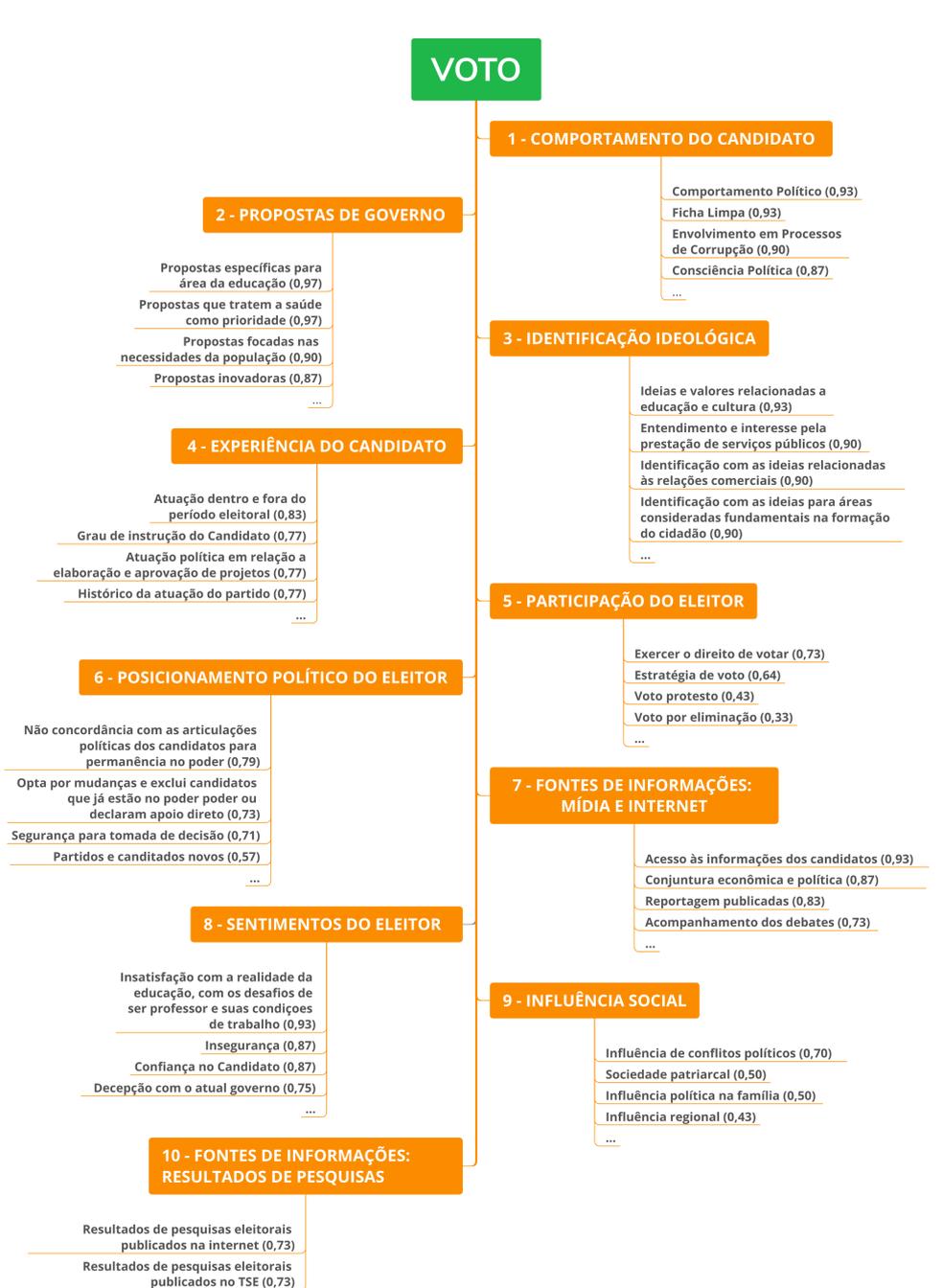


Figura 4.4: Identificação dos fatores que influenciam a decisão de voto

Por questões de espaço, os desdobramentos foram delimitados em 4 elementos. A versão completa do Modelo FIV.Br2018<sup>2</sup> é composta de 95 fatores.

<sup>1</sup>Os números que aparecem ao lado de cada fator dentro dos grupos, referem-se às médias aritméticas dos indicadores de avaliação. Na tentativa de sintetizar a opinião das pessoas sobre cada fator, atribuímos arbitrariamente aos elementos “Não influencia”, “Influencia pouco”, “Influencia muito”, os valores de 0, 0.5 e 1.0 respectivamente. Para uma avaliação qualitativa, sabemos que esses valores são apenas aproximações.

<sup>2</sup>Versão v1.0 do Modelo de Fatores que Influenciam o Voto. Disponível em: <http://bit.ly/2uNeT1R>

#### 4.4 Relação entre o FIV.Br2018 e os modelos cognitivos

Ao comparar o Modelo FIV.Br2018 com os modelos cognitivos referenciados, observamos a ocorrência de algumas características comuns entre eles, apresentadas na tabela 4.4.

Tabela 4.4: Comparativo dos modelos cognitivos com o modelo FIV.Br2018

Modelo FIV.Br2018	Modelo de Michigan	Modelo de Downs	Modelo de Holbrook
1-Comportamento do Candidato	X	X	
2-Propostas de Governo		X	
3-Identificação ideológica	X		
4-Experiência do Candidato	X	X	
5-Participação do Eleitor			
6-Posicionamento Político do Eleitor	X		
7-Mídia (rádio e TV) e internet (Redes Sociais)			X
8-Sentimentos do Eleitor			
9-Influência Social	X		X
10-Resultados de pesquisas eleitorais			X

Com relação ao modelo FIV.Br2018, identificamos o comportamento (Fator 1) e a experiência do candidato (Fator 4) como fatores que também são tratados pelos modelos de Michigan e Downs. As atitudes em relação ao candidato e a observância de suas competências comportamentais antes e durante o período eleitoral, o histórico de atuação (inclusive partidária), são consideradas importantes nos sistemas de preferências dos cidadãos.

Outro fator em comum entre o FIV.Br2018 e o Modelo de Downs é a avaliação das propostas de governos (Fator 2), onde o eleitor maximizará suas vantagens comparativamente aos candidatos da situação e da oposição. A equivalência do fator identificação ideológica (Fator 3) foi percebida entre o FIV.Br2018 e o Modelo de Michigan, oriundo da vertente psicológica, a compatibilidade de ideias é fortemente predominante. Já o posicionamento político do eleitor (Fator 6) é um fator que é fortalecido, segundo a escola de Michigan, ao longo da vida do eleitor e, desse modo, está relacionado ao que foi encontrado no modelo FIV.Br2018. A inclinação política do eleitor e sua maturidade política exercem impacto sobre a escolha do voto.

O fator fontes de informações (Fatores 7 e 10), de acordo com o modelo FIV.Br2018 e que também ocorre no modelo de Holbrook, provoca em nossos sistemas de crenças uma certa oscilação, quando a informação é disseminada, pois a característica on-line influencia os resultados das eleições, à proporção que alimenta os nossos filtros-bolha que, segundo

(PARISER, 2012), são mecanismos de previsão que criam e refinam constantemente uma teoria sobre quem somos e sobre o que vamos fazer ou desejar seguir e assim criam um universo de informações exclusivo para cada um de nós.

De acordo com (SANTAELLA, 2018), é plausível a hipótese de que, mesmo que os algoritmos fossem eliminados (o que é impossível), as pessoas ainda tenderiam a criar suas próprias bolhas de filtro como garantia de aproximação de pessoas que funcionam como espelhos de suas crenças, o que fortalece as crenças na medida e que o espelho cumpre a função de devolver as mesmas crenças de modo redobrado, e assim progressivamente. Isso explica os resultados das eleições, quanto ao uso de informações manipuladas pelos filtros e espalhados pelos motores de busca e mídias sociais digitais, que são utilizados para fins de interesses políticos.

Por fim, percebe-se que a influência Social (Fator 9) do Modelo FIV.Br2018 mostra-se semelhantes às visões dos modelos de Michigan e Holbrook, quando abordam o comportamento do eleitor em relação ao grupo ao qual ele pertence. Conforme aponta (CAMPBELL et al., 1980), os autores analisam a função das classes sociais no comportamento do eleitor.

Em (JUNIOR, 2009), o autor relata que a pesquisa teórica e metodológica da escola de Michigan foi utilizada para a realização de uma série de trabalhos sobre as eleições na Grã-Bretanha, conduzidos por Butler e Stokes, autores da obra *Political Change in Britain*. Conta (BUTLER; STOKES, 1969) que, ao longo do século XX, a influência partidária decaiu, e o crescimento das preferências eleitorais pelos trabalhistas cresceu em relação aos liberais. Nesse sentido, a influência social pode ser percebida em função do movimentos sociais, da situação sócio-econômica e da forma como a política é conduzida em cada país.

O estudo exploratório aqui apresentado foi fundamental para a compreensão do fenômeno do processo das eleições, bem como a observação do comportamento do eleitor ao decidir sobre sua opção de voto. Foi importante considerar os eventos contemporâneos, a influência dos mecanismos artificiais e ecossistemas digitais, sobretudo as deturpações espalhadas pelas redes sociais. Através dos depoimentos dos participantes da pesquisa, conseguimos fazer emergir alguns fatores que até então não foram tratados pelos modelos clássicos e criamos o modelo dos Fatores que Influenciam o Voto no Brasil em 2018 (FIV.Br2018).

Como parte integrante da proposta da pesquisa, o Modelo FIV.Br2018 visa contribuir no contexto da organização informacional e apoiar a construção do modelo dimensional, cujo

propósito é combater a ignorância, fomentar a solidez, a transparência e a confiabilidade das informações que são divulgadas, sobre os políticos, principalmente no decorrer das campanhas eleitorais.

## **5. Proposta da Solução: modelo dimensional e o sistema de informações prosopográficas**

Neste capítulo será apresentada a proposta da solução que é composto pelas seguintes seções: modelo e artefato (Seção 5.1), dicionário de dados (Seção 5.2) onde é apresentada a documentação referente as estruturas das fontes de dados; (Seção 5.3) no qual se discute sobre o comportamento do eleitor; (Seção 5.4) que apresenta o modelo conceitual do domínio estudado e finalmente a (Seção 5.5) que apresenta o modelo dimensional e o sistema de informações de consulta às informações prosopográficas.

### **5.1 Etapas da Solução Proposta**

A Figura 5.1 descreve as etapas consideradas no projeto da solução proposta nesta pesquisa. A primeira etapa do processo compreende a inspeção de todas as bases que alimentarão o DW, bem como a engenharia reversa dessas bases afim de entendermos o significado e qual o propósito de cada entidade no modelo dessas bases. Para essa etapa, selecionamos as bases de dados mantidas pelo grupo Transparência Brasil que são oriundas dos sistemas transacionais do senado, base de dados da Câmara dos Deputados e a base de dados que contém os dados de processos judiciais. Adicionalmente, foi utilizada a *Application Programming Interface* (API) desenvolvida pelo Departamento de Ciências da Computação da Universidade Federal de Sergipe (UFS), em que eles pesquisam sobre registros de processos judiciais e posicionamento político e a plataforma de dados abertos do Tribunal Superior Eleitoral.

As especificações (dicionário de dados) de cada fonte de dados, estão documentadas e disponíveis nos Apêndices C, D e E deste trabalho.

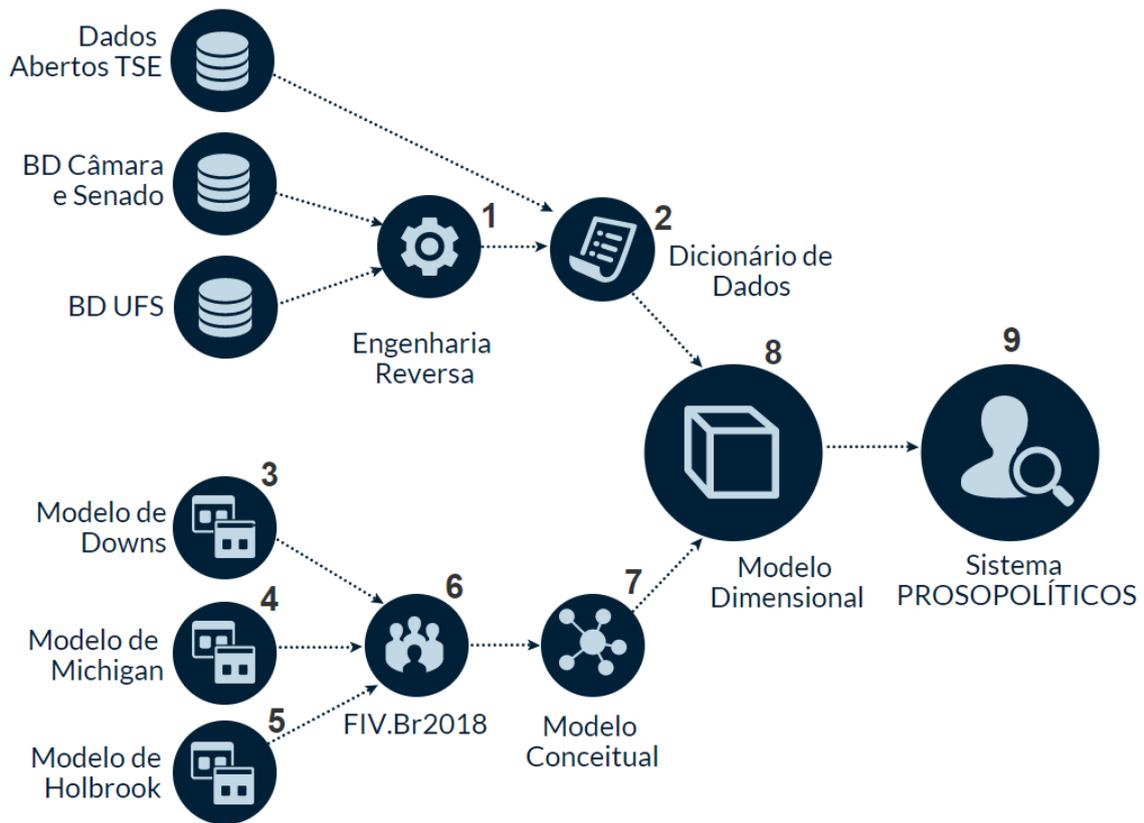


Figura 5.1: Visão das etapas da solução proposta.

## 5.2 Dicionário de Dados — Base de Dados do Transparência Brasil, Tribunal Superior Eleitoral e Universidade Federal de Sergipe

A segunda etapa envolveu a documentação de todas as bases, constituindo um dicionário de dados de todos os esquemas. Foi documentada a estrutura da base de dados da API-UFS, que fornece os dados sobre os candidatos, bem como a estrutura do arquivo CSV referente aos dados de candidatura, coletados do repositório de dados abertos do Tribunal Superior Eleitoral.

## 5.3 Pesquisa sobre os Modelos Cognitivos — Comportamento Eleitoral

Entre a terceira e a quinta etapa, foi importante pesquisar sobre os fundamentos teóricos relacionados aos modelos cognitivos que explicam o comportamento eleitoral à luz da Teoria do Voto. Tais princípios compuseram a base de conhecimento, com o objetivo de fundamentar as conjecturas teóricas da pesquisa.

Os modelos que explicam o comportamento eleitoral investigados neste estudo foram: Modelo de Anthony Downs (DOWNS, 1984), Modelo de Cognitivo da Escola de Michi-

gan (HIMMELWEIT; HUMPHREYS; JAEGER, 1985) e o Modelo de Thomas Holbrook (LODGE; MCGRAW; STROH, 1989) e foram detalhados no Capítulo 4.

A investigação desses modelos é justificada pela necessidade de um olhar socio-técnico sobre a construção do artefato, uma vez que a participação direta do usuário final traz como característica interdisciplinar uma aproximação com o mundo real. Para (BOEHM, 2006), a busca e o entendimento de tais fundamentos para construção de um artefato de *Software* é explicada da seguinte forma:

“Os sistemas serem úteis às pessoas, implica que é forçoso incluir as ciências do comportamento, da gestão e da economia entre as ciências relevantes, assim como a própria ciência da computação [...] o software e sua engenharia serão essenciais para o sucesso, mas também serão críticas as novas competências para integrar a engenharia de software à engenharia de sistemas, ao marketing, às ciências contábeis e às competências próprias ao núcleo central da ciência da computação.”(BOEHM, 2006, pg. 12, tradução nossa)

A etapa 6 consistiu na realização do estudo exploratório detalhado no capítulo 4. O objetivo desse estudo foi buscar entender os fatores considerados relevantes num processo de busca de informações sobre a vida de uma pessoa que se apresenta como político, e o que é considerado influente para confiar ou decidir o voto.

Ao fazer um diagnóstico dos Fatores que Influenciam o Voto no ano de 2018, este estudo contribuiu com um conjunto de conhecimentos úteis para concepção do artefato que foi denominado FIV.Br2018. Ao mesmo tempo, indicou níveis de influência que podem servir de parâmetro para identificar e direcionar ações de aperfeiçoamento em pesquisas relacionadas à política, processos e comportamento eleitorais no ciberespaço.

#### **5.4 O Modelo Conceitual de Informações Prosopográficas**

Para possibilitar a organização dos dados, a etapa 7 foi necessária no desenvolvimento da solução proposta e consistiu na definição do modelo conceitual. Este modelo foi orientado segundo o FIV.Br2018, desenvolvido como resultado do estudo exploratório discutido anteriormente e que contribuiu para determinar os elementos a serem representados pelo modelo dimensional.

Especificamos, em alto nível, entidades orientadas ao usuário e que são usadas para capturar a semântica, a fim de compor o modelo conceitual. Na Figura 5.2, representamos

graficamente as diferentes entidades do modelo proposto. Para a representação dos conceitos, entidades, atributos e seus relacionamentos, foi empregada a notação *Unified Modeling Language* (UML), que é o padrão para visualização, proposto por (RUMBAUGH; JACOBSON; BOOCH, 2004) para especificar, construir e documentar os artefatos de um sistema de software.

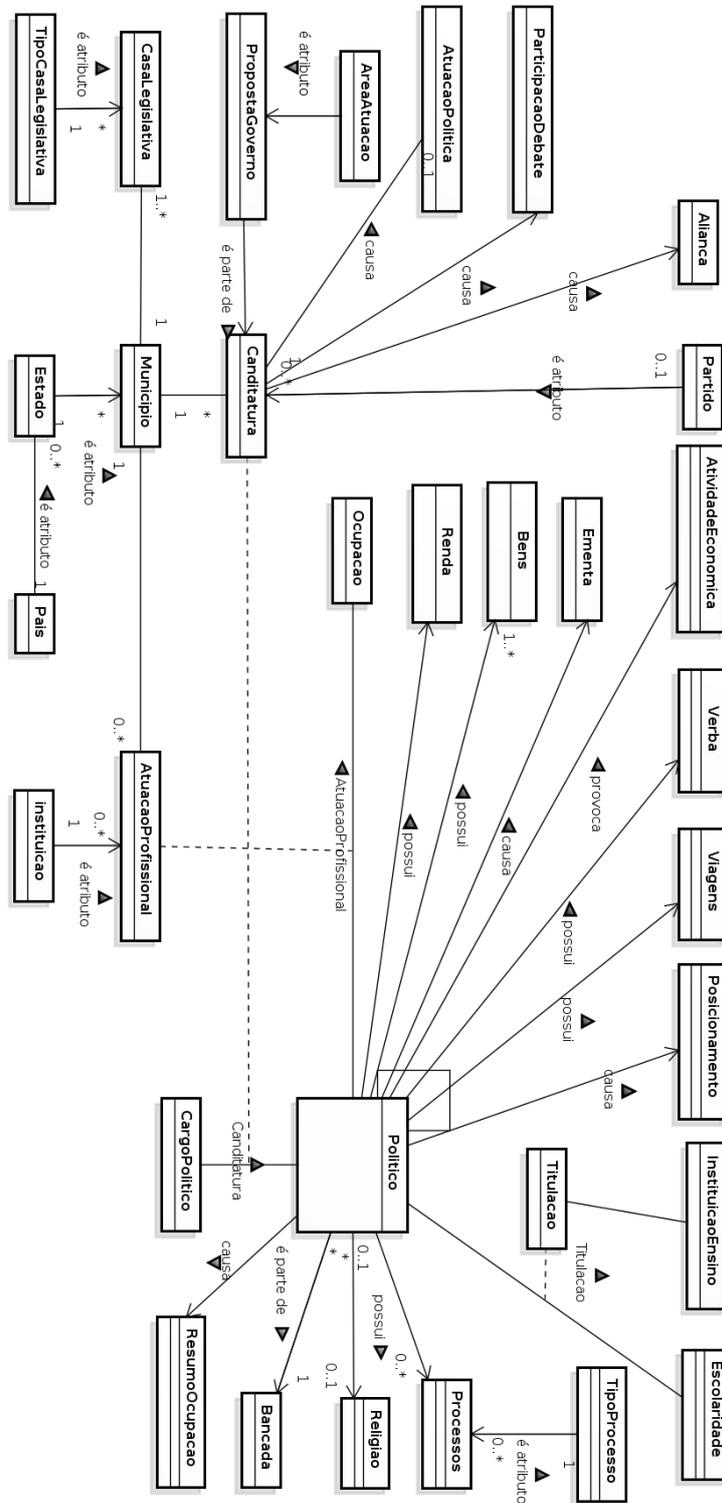


Figura 5.2: Modelo Conceitual (representado como um Diagrama UML)

O modelo conceitual é composto pelas seguintes entidades:

- Político: entidade que representa a pessoa em período de mandato eletivo.
- Titulação: titulação acadêmica do político.
- Escolaridade: A titulação está ligada a um grau de instrução/educação formal.
- InstituiçãoEnsino: a instituição de ensino associada à titulação do político.
- Religião: religião professada pelo político.
- Bens: representação dos bens acumulados pelo político.
- Renda: representação das fontes/renda.
- Candidatura: entidade que reflete a candidatura política.
- Partido: corresponde ao partido político.
- CargoPolitico: cargo no qual o político se candidata ou é eleito.
- PropostaGoverno: corresponde as propostas de governo de cada candidato.
- AreaAtuação: descreve a(s) área(s) em que o político tem como foco de atuação em suas propostas de governo.
- AtuacaoPolitica: representa o histórico de atuação política.
- ResultadoPesquisas: representa os resultados das pesquisas eleitorais.
- AtuacaoProfissional: reflete as experiências profissionais do político.
- Instituição: representa o local vinculado a experiência profissionais.
- ResumoOcupacao: histórico de ocupações do político;
- Ocupação: reflete o cargo ou posição.
- Bancada: representa a bancada política.
- Aliança: representa quais as parcerias formadas ao longo da candidatura.
- CasaLegislativas: representa o local onde o político de determinado cargo desempenha suas funções.
- TipoCasaLegislativa: reflete a classificação das casas.

- **Município:** local onde situa a casa legislativa.
- **ParticipacaoDebate:** entidade para representar uma participação em debate durante uma campanha.
- **DesempenhoMandato:** reflete os indicadores de desempenho de um mandato.
- **Processo:** representa os processos administrativos ou judiciais no qual o político está envolvido ou interessado.
- **Tipo Processo:** reflete a classificação de algum processo.
- **Posicionamento:** configura a opinião do político acerca de algum tema.
- **Viagem:** corresponde a uma viagem realizada pelo político.
- **Emenda:** configura a emenda parlamentar votada pelo político.
- **AtividadeEconomic:** representa a produção de bens e serviços do político nos últimos anos.
- **Estado:** representa o estado no qual ocorre a candidatura.

Vale ressaltar que as entidades desse modelo foram elaboradas em função de algumas fontes de informação. Algumas entidades são decorrentes de conceitos de modelagem de dados prosopográficos que encontramos na literatura, e vimos que as classes, sobretudo as relacionadas aos eventos, vivências e construção de relacionamentos, são pertinentes ao domínio tratado. Outras entidades estão respaldadas na literatura que trata modelos ontológicos e vocabulários peculiares às biografias, conforme explana (SHAW; TRONCY; HARDMAN, 2009), (LAGOZE; HUNTER, 2001), (BRICKLEY; MILLER, 2007). As entidades desse diagrama também são resultantes de uma fusão de dados providos e transformados de uma estrutura governamental oficial. Além disso, recebeu a contribuição do modelo FIV.Br2018 apresentado anteriormente.

## **5.5 O Artefato: o modelo multidimensional e o sistema de informação de apoio ao cidadão**

Nesta seção, destacamos as principais atividades referentes as etapas 8 e 9, ou seja o que foi feito para concepção do artefato. Tais atividades consistiram em, definição dos requisitos de negócio, a prototipação do modelo multidimensional, as atividades de extração e

tratamento dos dados, e o sistema gerenciador de banco de dados selecionado. O desenvolvimento do modelo multidimensional e do sistema de informação serão detalhados no capítulo 6.

### 5.5.1 Definição do Escopo

Com o principal objetivo do projeto em mente, levar informação transparente para os cidadãos brasileiros, foi preciso definir quais seriam as informações mais relevantes para o cidadão (o que decidimos com base no modelo FIV.Br2018) e quais delas seriam selecionadas como regra de negócio para construção do modelo multidimensional.

Com base nos conjuntos de dados existentes, os dicionários de dados foram analisados, e a partir desse conhecimento, foram elencadas as regras de negócio apresentadas na Figura 5.1, afim de construir cada modelo de dimensões.

<b>Regra de Negócio</b>	<b>Modelo FIV.Br2018</b>
Mandato Político	Experiência do Candidato
Candidatura	Experiência do Candidato
Atividade Econômica	Comportamento do Candidato
Gastos com Verbas	Comportamento do Candidato
Gastos com Viagens	Comportamento do Candidato
Emendas Aprovadas	Propostas de Governo
Posicionamento Político	Identificação ideológica
Ocorrências em Processos	Comportamento do Candidato
Matérias Publicadas	Identificação ideológica
Projetos Idealizados	Experiência do Candidato

Tabela 5.1: Seleção das regras para compor o modelo dimensional

### 5.5.2 Prototipação

Protótipo é a tangibilização de uma ideia, a passagem do abstrato para o físico de forma a representar a realidade e propiciar validações (VIANNA, 2012). Para prototipagem dos modelos multidimensionais, foi utilizada a funcionalidade de EER Diagram do *Software Mysql Workbench*.

Para o desenvolvimento dos protótipos não-funcionais da interface de consulta à base prosopográfica, o *software* escolhido foi o *canva.com*, que é gratuito e *on-line*.

### 5.5.3 Extração e tratamento dos Dados

Conforme visto na seção 3.2.2, um dos processos para construção de Data Warehouse consiste nas atividades de extração, transformação e carregamento dos dados. Para essa etapa, utilizamos a ferramenta de pré-processamento de dados, o Pentaho Data Integration (PDI) que é uma ferramenta Open Source.

Python versão 3 (PYTHON, 2001) foi a linguagem de programação escolhida para o desenvolvimento da interface de consulta neste trabalho. Ela foi escolhida por se tratar da linguagem com maior adoção no contexto da ciência de dados (de acordo com (MATOS, 2017)), ser gratuita, simples de ser utilizada e possuir diversas bibliotecas especializadas em trabalhar com grande volume de dados como o Pandas e o Numpy. Como *framework* para aplicações web, foi escolhido o Django, que é gratuito e de código aberto, e escrito em Python (DJANG, 2005). Além desses pontos de destaque, a estação de trabalho tem: a ferramenta Pycharm Professional Edition, que dá suporte sintático e semântico para o Django; a ferramenta Postman, para testar as chamadas; e o navegador Firefox, para visualizar e fazer o *debug* da interface. As bibliotecas requeridas de terceiros, e suas respectivas versões, são:

1. django (2.2.5)
2. djangorestframework (3.10.3)
3. mysqlclient (1.4.4)
4. psycopg2-binary (2.8.3)
5. pytz (2019.2)
6. sqlparse (0.3.0)

As 3 últimas bibliotecas são internas, ou seja, são dependências do Framework Django.

### 5.5.4 Sistema Gerenciador de BD

Por se tratar de um sistema com poucas inserções e inúmeras consultas, o modelo, *star-schema* foi escolhido para a estrutura de DW. Neste sentido, a escolha pelo MySQL foi uma escolha natural, dada sua natureza livre, sua larga adoção e uma comunidade de desenvolvedores grande e engajada (STACKOVERFLOW, 2018). Para a modelagem, diagramação e estruturação do banco de dados de modo visual, a ferramenta escolhida foi o MySQL Workbench.

## 6. O Artefato: PROSOPOLÍTICOS

Após análise de viabilidade e esforço de aquisição dos dados referentes aos tópicos definidos nas etapas 8 e 9 da proposta de solução (Seção 5.5), iniciou-se o desenvolvimento do modelo dimensional composto por suas dimensões e fatos.

### 6.1 O Modelo Multidimensional: organização dos dados

O desenvolvimento do artefato iniciou-se com a exploração da base de dados do Transparência Brasil. Foi realizado o procedimento de engenharia reversa da base resultando no modelo de entidade e relacionamento apresentado na Figura 6.1. Todos as imagens referente aos modelos apresentados nessa seção estão disponíveis em formato \*.pdf em um diretório compartilhado no GoogleDocs <sup>1</sup>

O acesso a base transacional foi efetuado através do protocolo *Secure Shell* (SSH), cujo objetivo é estabelecer um canal de comunicação cifra entre uma origem e um destino. Este serviço é muito utilizado com o fim de realizar a administração geral de um sistema informático remoto, conforme aponta (LAUDON; LAUDON, 2016). O mesmo procedimento foi realizado para a base de dados do Grupo de Pesquisa da Universidade Federal de Sergipe, conforme é apresentado na a Figura 6.2

Por conseguinte, foram examinados os dados do repositório de dados abertos do Tribunal Superior Eleitoral (TSE)<sup>2</sup> e lá foram selecionados os arquivos referentes as candidaturas de 2004 a 2018. O dicionario de dados está disponível no Apêndice C

Com base nas 3 fontes de dados, os requisitos de negócio foram então selecionados. Conforme a Tabela 5.1, cada linha da primeira coluna representa um fato do modelo dimen-

<sup>1</sup>Diretório de imagens: modelos de dados. Disponível em <http://bit.ly/388Tui4>

<sup>2</sup>Arquivos com informações sobre o perfil dos candidatos nas eleições disponível em <http://www.tse.jus.br/hotsites/pesquisas-eleitorais/candidatos.html>

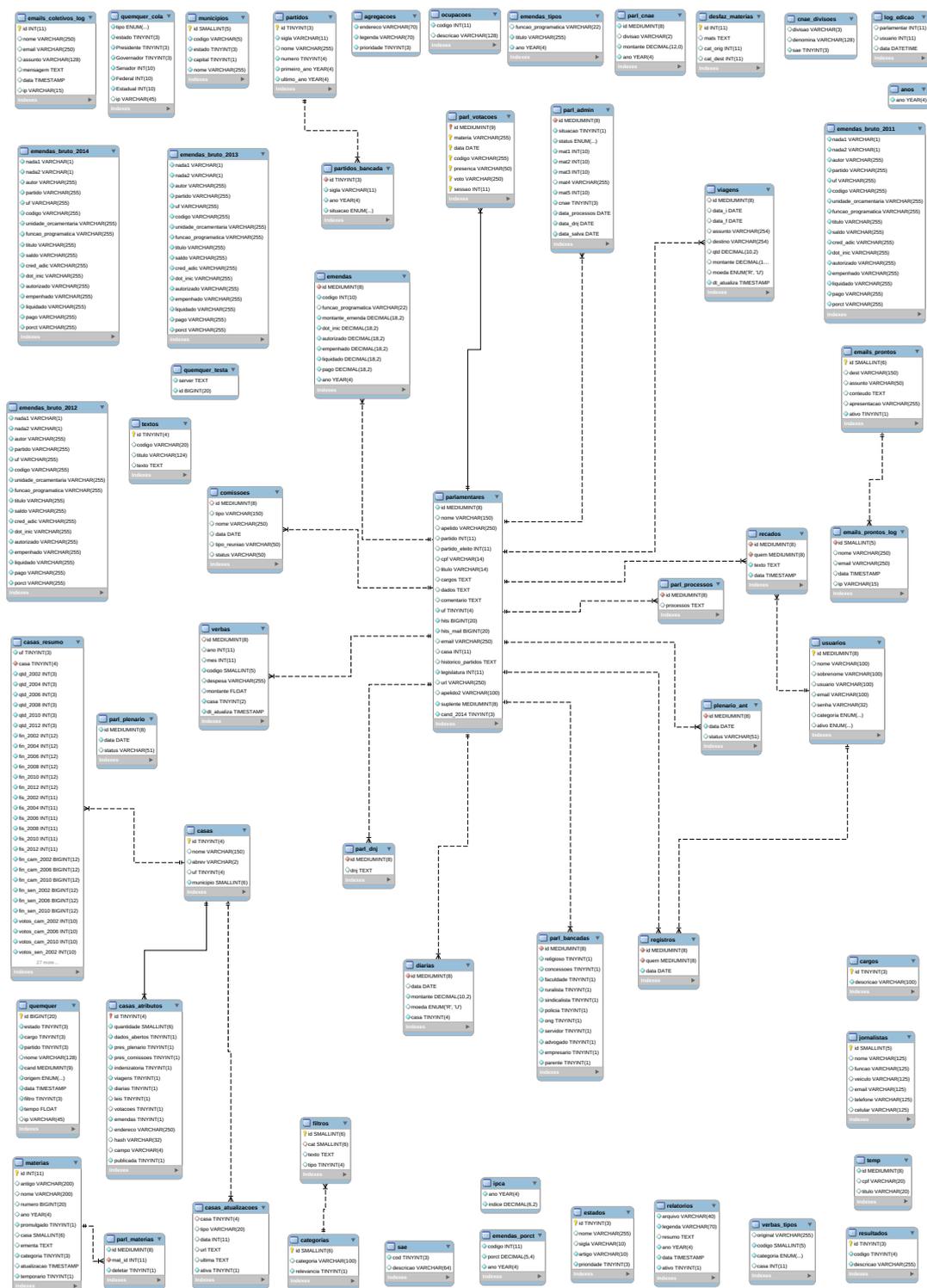


Figura 6.1: Diagrama de Entidade e Relacionamento da Base Transparência Brasil

sional. Esses requisitos, por sua vez são aderentes ao modelo FIV.Br2018 resultante do estudo exploratório.

A partir dos requisitos definidos e os dados necessários já analisados, foi iniciado o trabalho de projeção dos modelos dimensionais, que são pré-requisitos para os processos de

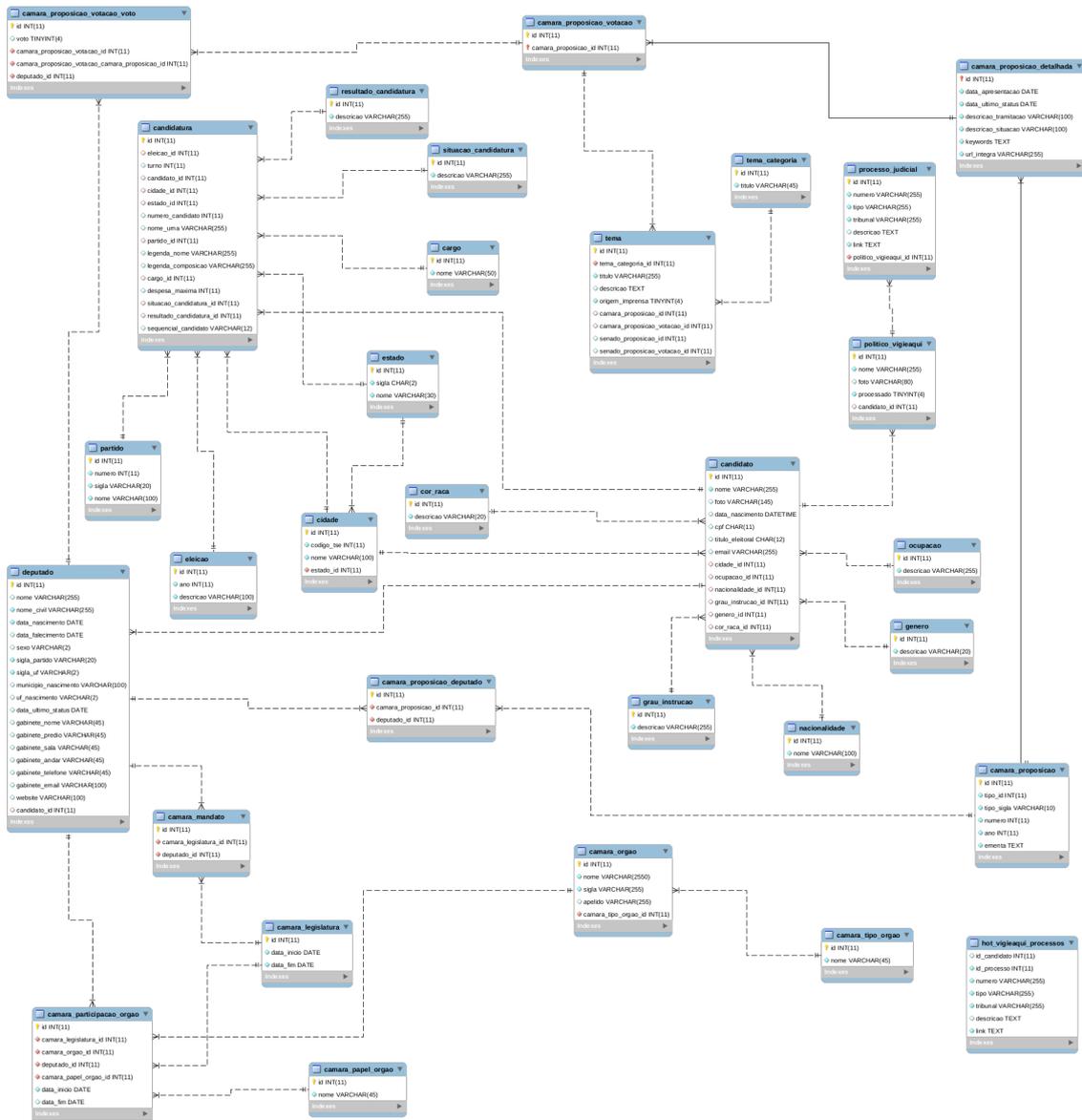


Figura 6.2: Diagrama de Entidade e Relacionamento da Base DCC-UFS

*Extract, Transform, Load - (ETL).*

### 6.1.1 Projeção dos Modelos Dimensionais

Nesta seção, será apresentada a projeção dos modelos dimensionais. O primeiro modelo é o Mandato Político, representado pela tabela FATO\_MANDATO. Na Figura 6.3, a entidade central é o fato e as extremidades são as dimensões.

Por conseguinte o modelo de Atividade Econômica, originou-se de uma entidade denominada Classificação Nacional de Atividades Econômica (CNAE). De acordo com (FEDERAL, 2019), essa classificação é aplicada a todos os agentes econômicos que estão engajados na produção de bens e serviços, podendo compreender estabelecimentos de

empresas privadas ou públicas, estabelecimentos agrícolas, organismos públicos e privados, instituições sem fins lucrativos e agentes autônomos (pessoa física). Nesse sentido, o modelo Atividade Econômica na Figura 6.4 representa as atividades realizadas por um Político.

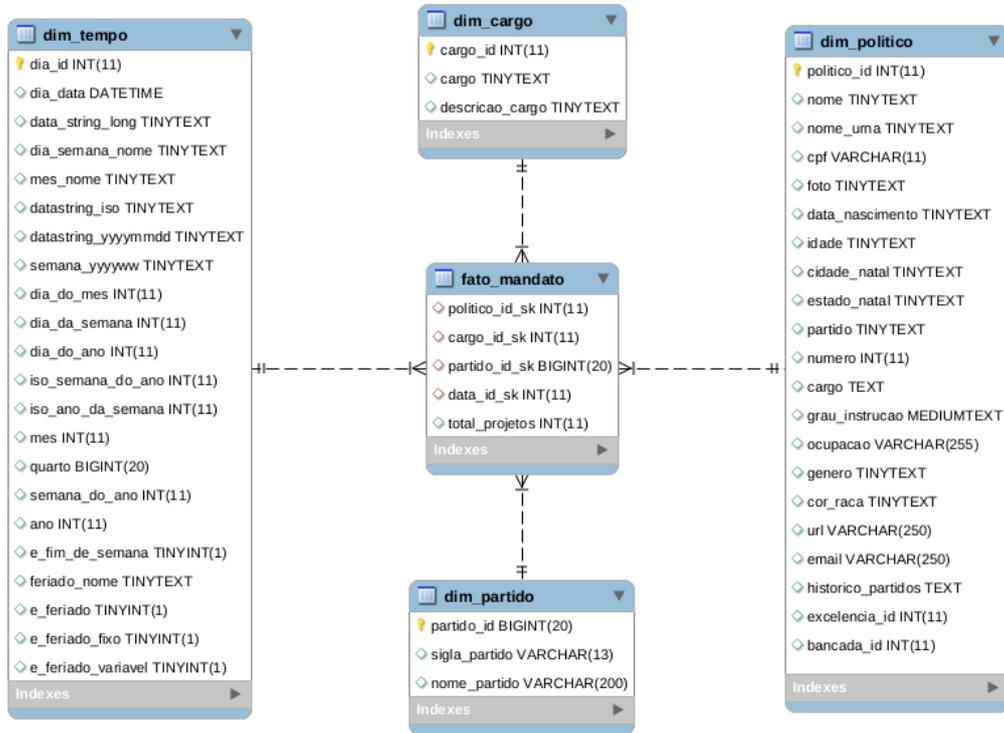


Figura 6.3: Diagrama do Modelo Dimensional Mandato Político

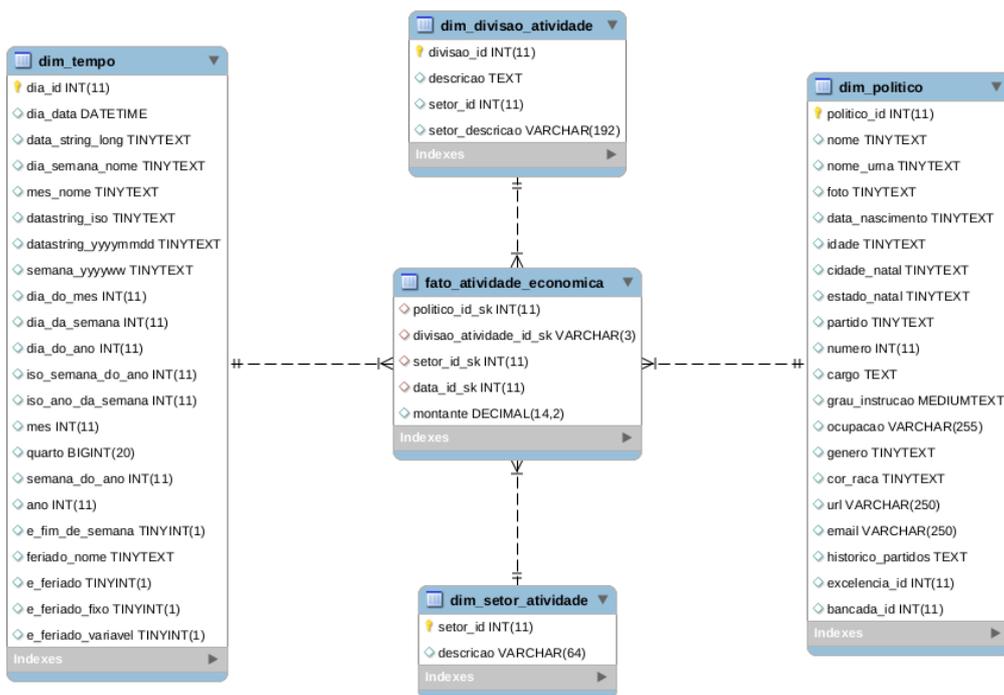


Figura 6.4: Diagrama do Modelo Dimensional Atividade Econômica

O próximo modelo é o Candidatura, conforme mostra a Figura 6.5, ele é composto de 17 dimensões e traz informações sobre a eleições, turnos, coligações, situação da candidatura e conta com a dimensão dim\_político bem detalhada.

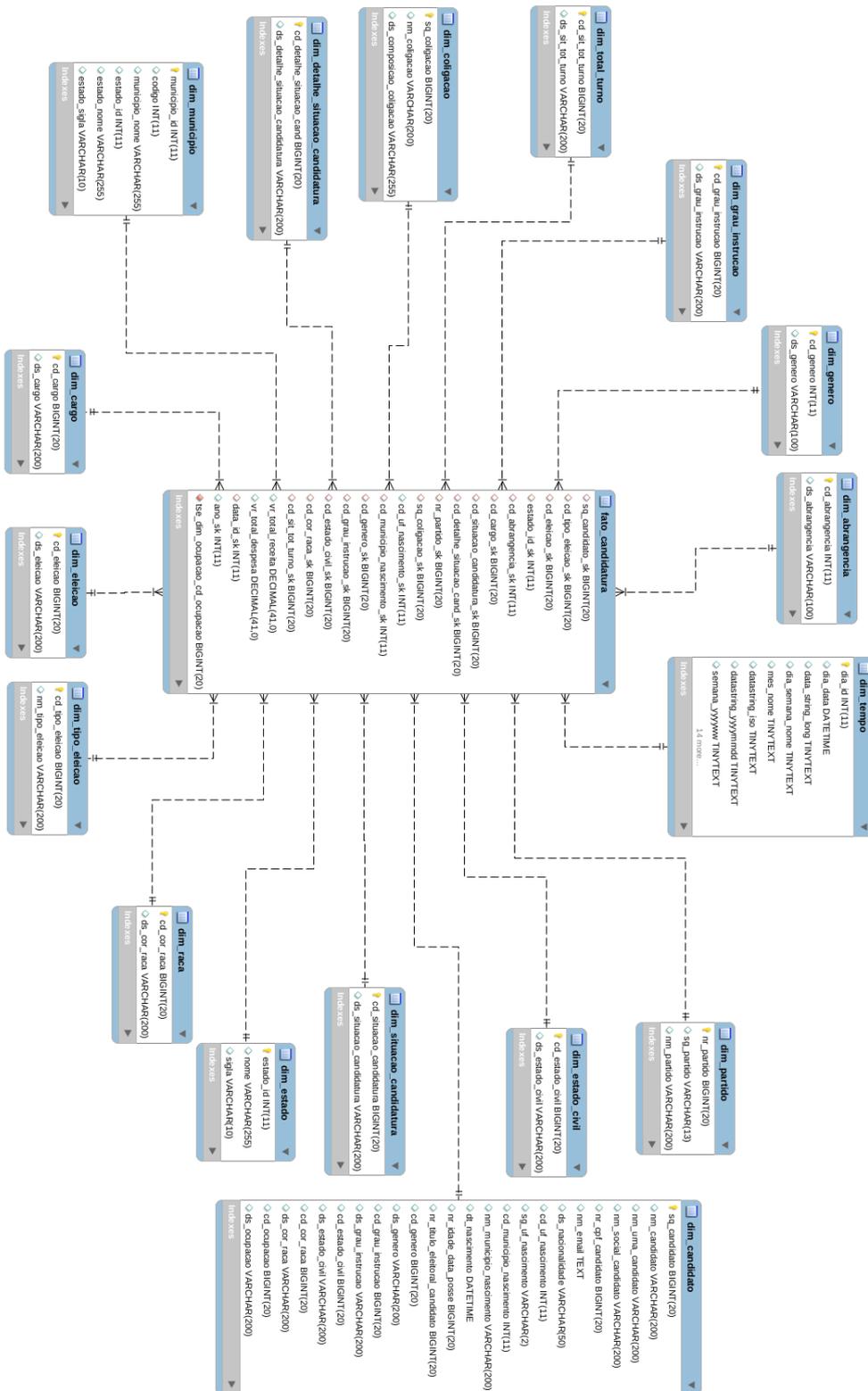


Figura 6.5: Diagrama do Modelo Dimensional Candidatura

Os gastos com verbas também foram considerados fatores críticos na concepção do modelo FIV.Br2018. Por isso foi criado o modelo para representar as verbas executadas pelos políticos bem como suas respectivas subcategorias. A Figura 6.6 representa o modelo dimensional criado para esse fim.

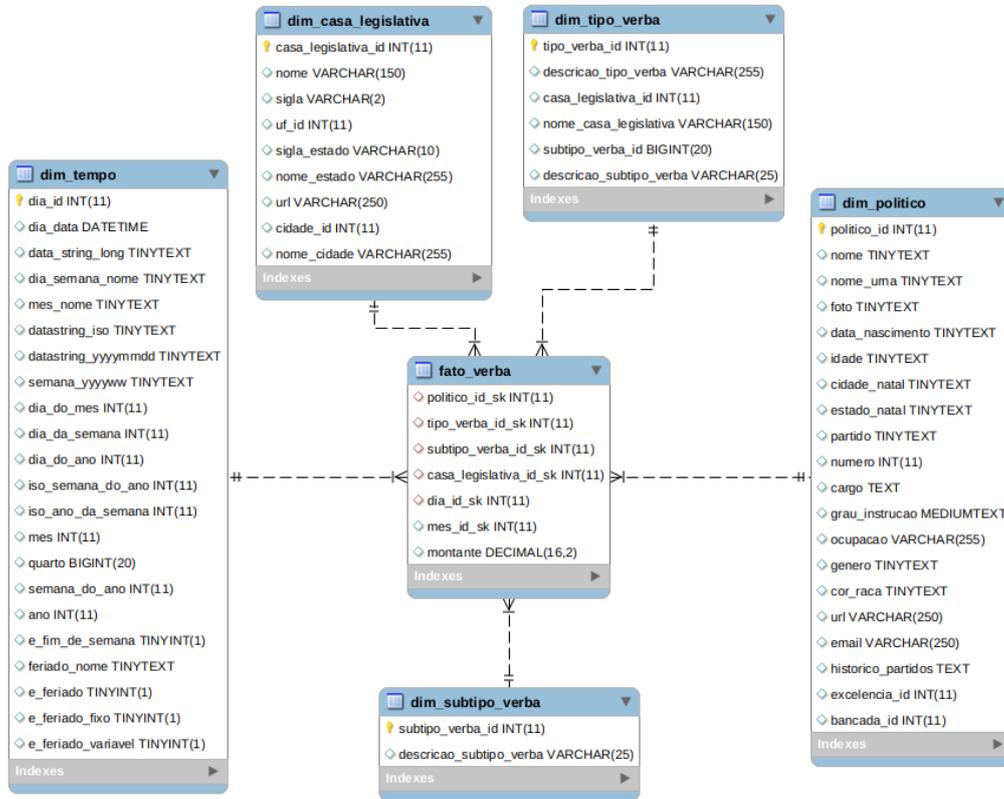


Figura 6.6: Diagrama do Modelo Dimensional Verbas

Outro dado identificado nas bases é o registro de viagens realizadas pelos políticos. Foi criado o modelo Viagens, composto pelas dimensões dim\_tempo, dim\_justificativa\_viagem, dim\_político e para representação do destino, foram incluídas as dimensões, dim\_pais, dim\_estado e dim\_município conforme mostra a Figura 6.7. Em seguida, a Figura 6.8 apresenta o modelo dimensional das emendas parlamentares.

E por fim, conforme apresentado na Figura 6.9 foi criando um modelo dimensional Modelo Candidatura Doações e Voto. O objetivo desse modelo é reunir e apresentar as informações financeiras relativas às candidaturas de cada político e o quantitativo de votos.

Esse modelo é composto pela tabela fato\_candidatura\_doacao\_voto e as dimensões dim\_político, dim\_partido, dim\_estado, dim\_funcao\_programática, dim\_tempo e dim\_unidade\_orcamentaria. Acompanha também as principais métricas que são: recursos recebidos de candidatos, recursos recebidos de pessoas jurídicas, recursos de auto doações, recursos recebidos de empresas e a totalização de votos recebidos.

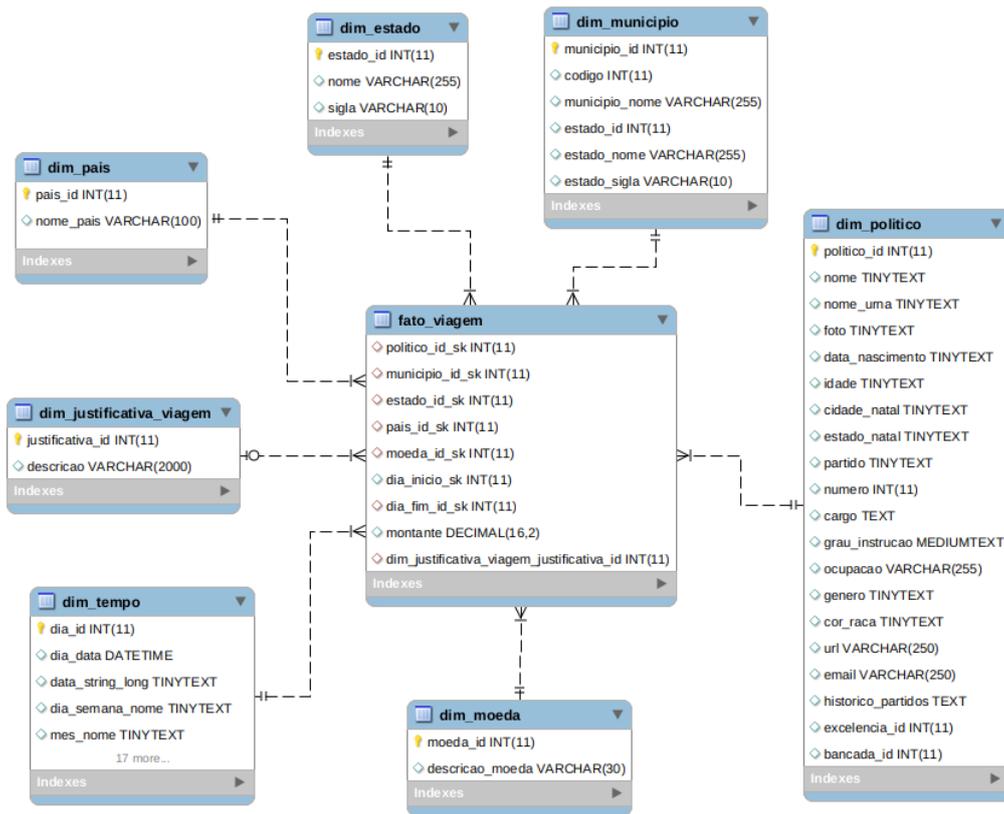


Figura 6.7: Diagrama do Modelo Dimensional Viagem

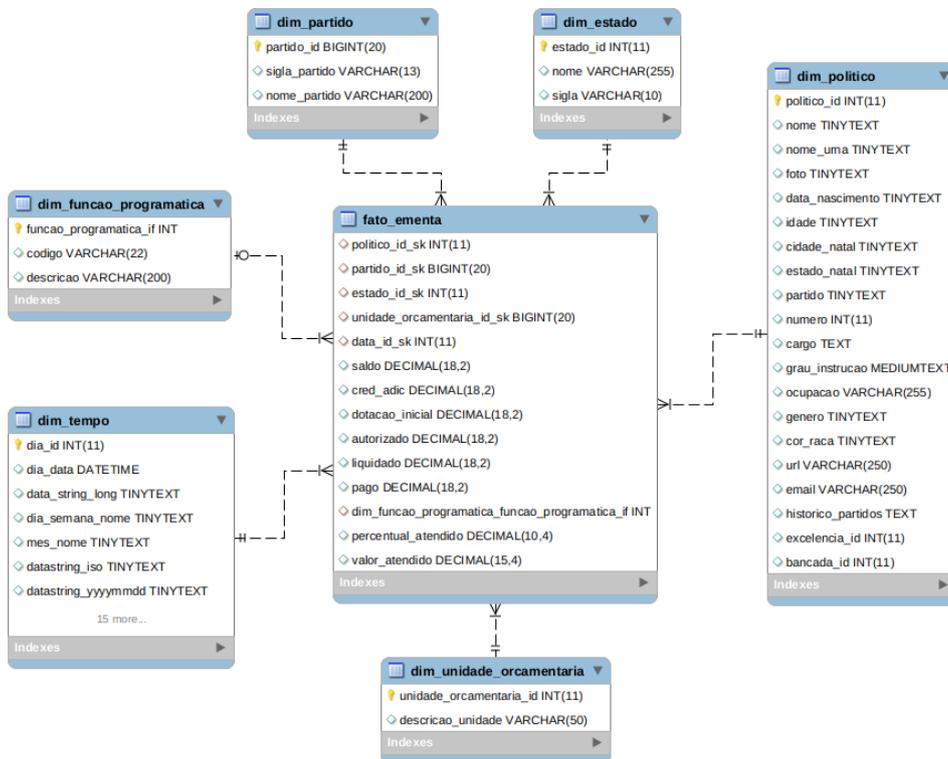


Figura 6.8: Diagrama do Modelo Dimensional Emenda

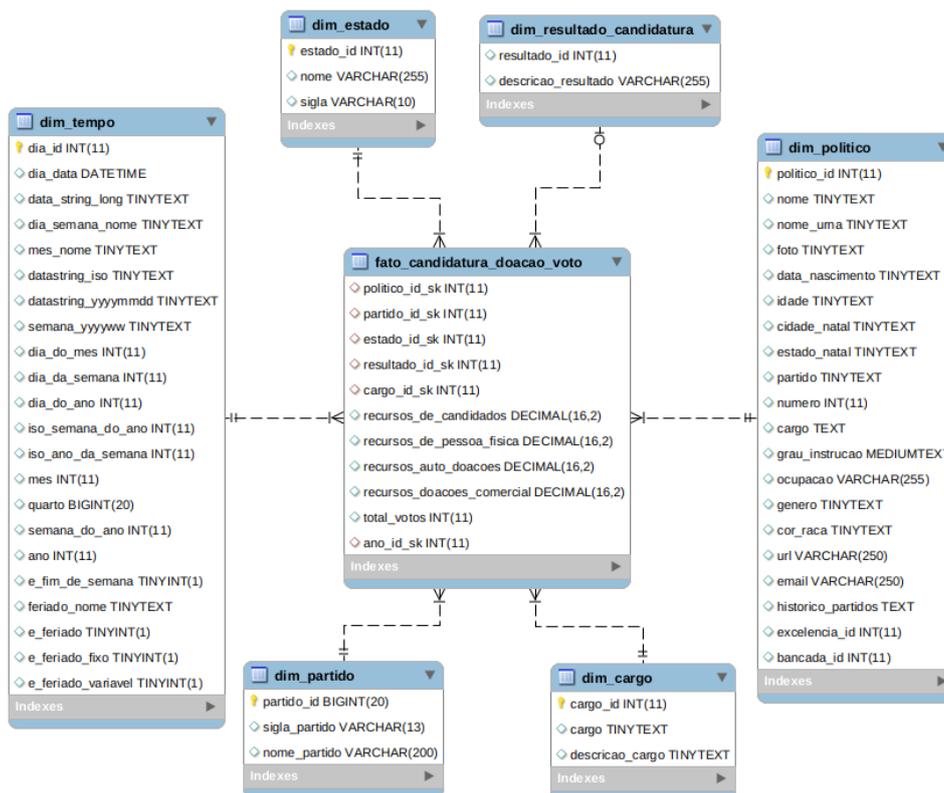


Figura 6.9: Diagrama do Modelo Dimensional Candidatura Doações e Voto

## 6.1.2 Origem dos Dados e Limitações

De posse dos modelos dimensionais projetados, a fase de pré-processamento foi realizada com foco nas 3 fontes de dados, que serão detalhados na próxima seção.

### 6.1.2.1 Base Transparência Brasil: dados da câmara dos deputados e do senado brasileiro

A base de dados da Câmara e do Senado podem ser utilizadas tanto para acompanhar a tramitação e as votações de projetos e pareceres como para verificar a atuação e os gastos dos deputados e senadores eleitos (ASSOCIATION, 2018).

No momento em que a análise das bases de dados foi realizada (entre Janeiro e Março de 2019), a conexão para execução das consultas estava bem estável, contudo foram encontradas algumas inconsistências e estruturas desordenadas. Assim, as transformações e cargas foram realizadas através da Ferramenta Pentaho Data Integrator (PDI), com o uso da Linguagem Javascript e PL/SQL. Também foi necessária a implementação de expressões regulares para recuperação das informações relacionadas aos mandatos eletivos de cada político, conforme apresentado na Lista de Código ??

As Figuras 6.10 e 6.11 mostram como os dados referentes aos mandatos dos políticos estavam estruturados em suas bases originais e como ficaram organizados após a aplicação dos algoritmos de expressões regulares, conforme consta no Apêndice G, a fim de carregar os dados para a tabela FATO\_MANDATO.

Form Editor Navigate: [Navigation icons] 31 / 501 Edit: [Edit icons]

Id: 78

Apelido: Alice Portugal

Cargos: Foi deputada federal outras três vezes pelo PC do B (2003-2007/ 2007-2011/ 2011-2015). Ainda exerceu dois mandatos de deputada estadual pelo mesmo partido (1995-1999/ 1999-2003).

Figura 6.10: Mandato eletivo antes: forma original como campo texto

#	politico_id	nome_urna	descricao_cargo	sigla_partido	ano
13	1274062	Alice Portugal	Deputado Federal	PC do B	2003
14	1274062	Alice Portugal	Deputado Federal	PC do B	2006
15	1274062	Alice Portugal	Deputado Federal	PC do B	2007
16	1274062	Alice Portugal	Deputado Federal	PC do B	2008

Figura 6.11: Mandatos eletivos depois: visualização dos dados transformados

Outra dificuldade encontrada na extração dos dados, foi a disposição dos dados nas tabelas. Os dados que geralmente ficam em formato horizontal (linhas) foram colocados em formato vertical (Colunas). A Figura 6.12 apresenta um exemplo da disposição dos dados.

#	id	religioso	concessoes	faculdade	ruralista	sindicalista	policia	ong	servidor	advogado	empresario	parente
1	95482	0	0	0	0	0	0	0	0	0	0	0
2	95481	0	0	0	0	0	0	0	0	0	0	0
3	95480	1	0	0	0	0	0	0	0	0	0	0
4	95479	0	0	0	0	0	0	0	0	0	1	0
5	95478	0	0	0	1	0	0	0	0	0	0	1
6	95477	0	0	0	0	0	0	0	0	0	0	1
7	95476	0	0	0	0	0	0	0	0	0	0	0
8	95475	0	0	0	0	0	0	0	0	0	0	0
9	95474	0	0	0	1	0	0	0	0	0	0	1
10	95473	0	0	0	0	0	0	0	0	0	1	0

Figura 6.12: Bancadas política antes: tabela original

Na sequência, a figura 6.13 ilustra os dados tratados e atualizados para carregamento e construção da dimensão DIM\_POLITICO.

#	politico_id	politico_nome	bancada_id	bancada_descricao
1	1274197	Abelardo Luiz Lupion Mello	4	Ruralista
2	1274589	Abmael Sousa Lacerda	4	Ruralista
3	1274591	Ada Lili Faraco de Luca	7	Servidores públicos
4	1274592	Adahil Barreto Cavalcante Sobrinho	8	Advogados
5	1274593	Adalberta da Rocha dos Santos Pe...	7	Servidores públicos
6	1274595	Adalto de Freitas Filho	4	Ruralista

Figura 6.13: Bancada política depois: visualização dos dados transformados

Outro exemplo de problema encontrado que mereceu atenção, foi a estrutura da tabela que contem o conjunto de dados relativos às emendas parlamentares. A Figura 6.14 apresenta um recorte da tabela de emendas em sua estrutura original.

Como a disposição dos dados na tabela dificultava a manipulação dos dados, foi necessária a implementação de uma *procedure* em PL/SQL e a partir daí, o carregamento para tabela FATO\_EMENDAS foi possível, conforme documentado na Lista de Código 6.1. Os dados passam e ser organizados conforme mostra a Figura 6.15

```

1 //Procedure para manutencao dos dados das emendas
2 CREATE DEFINER='root'@'localhost' PROCEDURE `AtualizaAutor`()
3 BEGIN
4     DECLARE v_finished INT DEFAULT 0;
5     DECLARE pautor, ppartido, puf, a, b, c VARCHAR(100);
6     DECLARE pcodigo, d INT;
7     DECLARE curl CURSOR FOR SELECT autor, partido, uf, codigo
8         FROM dw.emenda;
9     DECLARE CONTINUE HANDLER FOR NOT FOUND SET v_finished = 1;
10
11 OPEN curl;
12 read_loop: LOOP
13     FETCH curl INTO a, b, c, d;
14     IF v_finished = 1 THEN
15         LEAVE read_loop;
16     END IF;
17

```

```

18     IF d is null THEN
19         SET pautor = a;
20         SET ppartido = b;
21         SET puf = c;
22     ELSE
23         UPDATE dw.emenda set autor = pautor, partido = ppartido,
                uf = puf where codigo = d;
24     END IF;
25
26 END LOOP read_loop;
27 CLOSE curl;
28 DELETE FROM dw.emenda WHERE codigo is null;
29 END

```

Listing 6.1: Tratamento dos dados sobre Emendas Parlamentares

autor	partido	uf	codigo	unidade_orcamentaria	funcao_programatica	titulo	saldo	cred_adic	dot_inic
Abelardo Camarinha	PSB.	SP							
			21640001	56101 M. das Cidades	15.451.2054.1D73.0035	Apoio à Política Nacional de Desenvolvimento...	7243000.00	192947000.00	200190000.00
			21640002	36901 Fundo Nacional de Saúde	10.301.2015.8581.0035	Estruturação da Rede de Serviços de Atençã...	2543000.00	208504000.00	211047000.00
			21640003	55901 FNAS	08.244.2037.2B31.3926	Estruturação da Rede de Serviços de Proteçã...	100000.00	0.00	100000.00
			21640004	36901 Fundo Nacional de Saúde	10.302.2015.8535.0035	Estruturação de Unidades de Atenção Especi...	4550000.00	151530000.00	156080000.00
			21640005	36901 Fundo Nacional de Saúde	10.302.2015.8535.3692	Estruturação de Unidades de Atenção Especi...	250000.00	200000.00	450000.00

Figura 6.14: Emendas parlamentares antes: forma original

#	autor	partido	uf	titulo	saldo	percentual_atendido
1	Abelardo Camarinha	PSB	SP	Apoio à Política Nacional de Desenvolvimento Urbano - No Estado de São Paulo	7243000.00	83.5000
2	Abelardo Camarinha	PSB	SP	Estruturação da Rede de Serviços de Atenção Básica de Saúde - No Estado de São ...	2543000.00	68.1200
3	Abelardo Camarinha	PSB	SP	Estruturação da Rede de Serviços de Proteção Social Especial - No Município de Sã...	100000.00	0.0000
4	Abelardo Camarinha	PSB	SP	Estruturação de Unidades de Atenção Especializada em Saúde - No Estado de São ...	4550000.00	61.2500
5	Abelardo Camarinha	PSB	SP	Estruturação de Unidades de Atenção Especializada em Saúde - No Município de Ma...	250000.00	100.0000

Figura 6.15: Emendas parlamentares depois: dados transformados

Nestas bases, os principais pontos de acesso para execução dos Processos de ETL, foram:

- **Políticos:** ponto principal do *dataset*, são os agentes da câmara e atuam como autores de proposições, membros de órgãos etc. Dependendo da quantidade de votos recebidos, um deputado pode ser titular ou suplente. No Dados Abertos, cada deputado que exerceu mandato, mesmo que por apenas um dia, recebe um número identificador exclusivo;
- **Assembléia Legislativa:** os trabalhos da Câmara são exercidos pelos deputados nas assembleias;

- Bancada: Grupo de parlamentares que pertencem a uma mesma organização partidária;
- Cargo: Cargo no qual o político concorre ao mandato eletivo;
- Divisão de Atividade: classificação das atividades econômicas;
- Mandatos: períodos de trabalhos legislativos, iniciados no dia da posse dos parlamentares após uma eleição e encerrados na véspera da posse dos deputados seguintes;
- Partidos: organizações políticas às quais os deputados são filiados e estão oficialmente registradas no Tribunal Superior Eleitoral brasileiro;
- Projetos: projetos elaborados pelo político;
- Ocorrência de Processos: conjunto de dados que relata o envolvimento do político em processos administrativos e de improbidade administrativa.
- Emendas Parlamentares: As emendas apresentadas por parlamentares ao orçamento são de dois tipos: aquelas que se incorporam a emendas de outros parlamentares num mesmo conjunto, de modo que o montante total resulta da agregação das diversas emendas; e as emendas individuais, que aparecem sozinhas no orçamento. No cálculo das parcelas de cada parlamentar que são atendidas não se faz distinção entre emendas individuais e coletivas, considerando-se sempre o percentual do agregado;
- Matérias: alguns assuntos de diversos tipos são submetidos a votações e algumas dessas votações são nominais, com o registro dos votos de cada deputado. Existe um sumário das matérias propostas desde 2003. Elas são classificadas em categorias temáticas (Educação, Saúde, Mulher etc.);
- Viagem: informações sobre viagens e os destinos, bem como valores e justificativa de viagem; e
- Informações sobre cidades, estados e países.

#### **6.1.2.2 Base UFS: dados sobre os processos judiciais e posicionamento político**

A estrutura organizacional do poder judiciário brasileiro é bem complexa. São diversos tribunais estaduais e federais, quatro instâncias e três ramos especializados da justiça federal da união (STF, 2018), como mostra a Figura 6.16. Apesar de estar passando por um

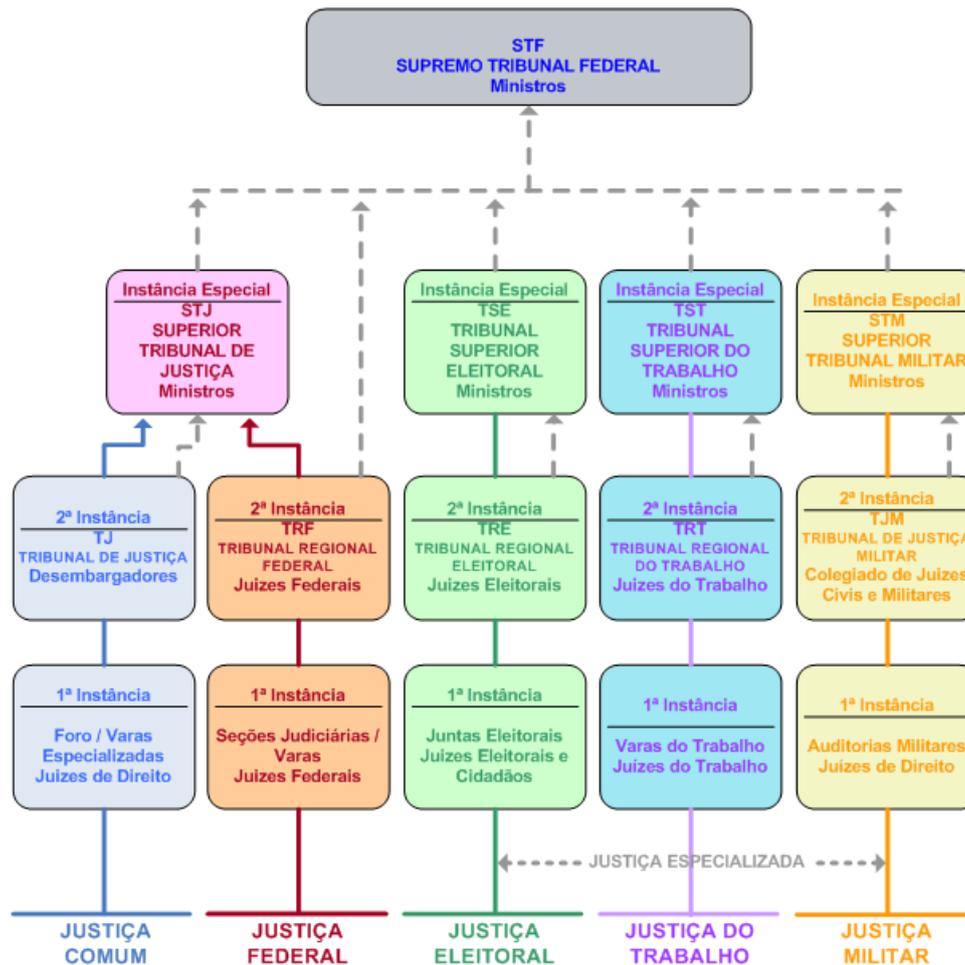


Figura 6.16: Organograma do Poder Judiciário. (MATOS, jfal)

processo de modernização e de abertura de dados nos últimos anos, ainda não é possível obter informações públicas de maneira centralizada e automatizada dessas instituições.

A observância do envolvimento de um político em casos de corrupção consiste num dos fatores que influencia a decisão de voto e chama a atenção do cidadão num processo investigativo. Na última década, o uso da Internet para divulgação de informações de campanhas e para a promoção da transparência na política aumentou substancialmente. A quantidade de informações disponíveis sobre as atividades parlamentares de cada deputado, sobretudo as criminais, nunca foram tão grandes e, dada a rápida velocidade de difusão da informações através de vários meios, tornam-se facilmente acessíveis e atrativas ao cidadão comum.

Algumas pesquisas foram realizadas com o intuito de observar e questionar o comportamento político e eleitoral em virtude do uso da internet como exposição de casos de corrupção. No trabalho de (GONCALVES, 2010), a autora apresentou uma análise comparativa dos escândalos de corrupção e consequentes ações ajuizadas contra os membros

da Câmara dos Deputados do Brasil na 52<sup>a</sup> legislatura (2003-2007) e na 53<sup>a</sup> legislatura (2007-2010). O objetivo foi entender se o projeto de disseminação de informações através do Transparência Brasil teve algum impacto visível quanto a criminalidade dos parlamentares.

Ainda em (GONCALVES, 2010), a autora apresentou que o número de deputados processados por crimes cometidos durante o seu mandato diminuiu mais de 70% entre os dois períodos legislativos analisados. Tal redução do número de deputados julgados ocorreu não somente dentro do grupo de deputados reeleitos, mas também, entre os desafiantes eleitos em 2006, em comparação com os ex-deputados, cujos cargos assumiram em 2007.

Essas evidências apresentadas sugerem fortemente que um maior acesso à informação sobre as ações dos políticos e o seu perfil pode ter um forte impacto não apenas na seleção política, mas também no comportamento dos políticos no cargo.

Dado que registros criminais dos candidatos nos fazem interpretar que a corrupção política não se restringe apenas ao governo federal, mas também aos governos estaduais e municipais, esta pesquisa julga importante alargar a política de transparência e divulgação de informações inserindo mais uma dimensão no modelo dimensional: o Envolvimento em Processos de Corrupção. Com base em fatos objetivos, sejam eles registros criminais nos tribunais, reivindicações de despesas ou detalhes sobre contratos, essa dimensão tem um peso considerável no comportamento eleitoral.

Para conseguir incorporar as informações sobre processos judiciais dos políticos foi preciso buscar fontes secundárias e confiáveis, estabelecer uma parceria e construir uma base mais completa. Durante a apresentação de trabalhos em eventos científicos, tive a oportunidade de conhecer o grupo de pesquisa da Universidade Federal de Sergipe (UFS). Este mesmo grupo estuda e explora informações sobre posicionamento político e envolvimento em processos por parte dos políticos. O contato com o grupo nos possibilitou ter acesso a base de dados sobre políticos, mantida pelo Grupo de Pesquisa do Departamento de Ciências da Computação da UFS.

Sobre a transparência de informações relacionadas ao envolvimento dos políticos em processos, um dos trabalhos deste grupo relata que:

"Uma das iniciativas da sociedade civil que tem demonstrado destaque na união de dados dos diversos tribunais espalhados pelo país é o Vigie Aqui, “o jeito mais fácil de descobrir quem são os políticos enrolados com a justiça” (RECLAMEAQUI, 2018). Idealizado pelo instituto Reclame Aqui, um dos maiores sites do Brasil - com mais de 600.000 acessos diários (RECLAME, 2018),

o projeto conseguiu levantar dados judiciais de presidentes, ex-presidentes, senadores, deputados federais e governadores, formando um *dataset* bem completo da esfera político-judicial brasileira. Após alguns contatos, o LADDRES e o Vigie Aqui firmaram uma parceria e o *dataset* de processos judiciais foi disponibilizado via arquivo .json."(PASSOS, 2018, pg.40).

Estabelecido o acesso aos dados de processos, foi incluída uma dimensão de processos no modelo proposto com os seguintes atributos: identificador do político, número do processo, tipo do processo, tribunal que cuida do processo, descrição do processo e o link para acesso direto às informações do trâmite do processo no sistema de consulta do tribunal responsável por ele.

Para concluir os requisitos de negócio do modelo dimensional, foi incluído o posicionamento político. O objetivo é conhecer o que o candidato pensa/manifesta sobre algum tema. Ainda em (PASSOS et al., 2019), os autores relatam que foi realizado um levantamento acerca de assuntos que se destacaram nos últimos anos e que possivelmente provocam questionamentos e posicionamentos dos políticos. São eles:

“Congelamento dos Gastos em Serviços Públicos; reforma trabalhista; terceirização de todas as atividades; reforma da previdência; taxaço de grandes fortunas e heranças; reforma do ensino médio; cobrança por cursos de especialização em universidades públicas; cotas em universidades públicas; isenção fiscal a petroleiras estrangeiras; escola sem partido; fim do regime de partilha; privatização da Petrobrás; porte de armas; redução da maioria penal; intervenção federal (militar) no Rio de Janeiro; descriminalização do aborto; descriminalização das drogas” (PASSOS et al.,2019, pg. 6).

O conjunto de dados sobre posicionamento do Político foi delimitado aos candidatos que tiveram mandatos na última legislatura e/ou candidatos à presidência do país, pois de acordo com (PASSOS et al., 2019), o motivo foi por não terem recursos e tempo suficiente para cobrir a demanda de catalogar a opinião de todos os candidatos da base. Desse modo, as informações foram coletadas e as transformações necessárias foram aplicadas a fim de criar uma dimensão que refletisse o posicionamento político.

### **6.1.2.3 Base TSE: repositório de dados abertos**

O repositório de dados eleitorais do TSE é uma iniciativa de abertura de dados e reúne informações brutas das eleições brasileiras a partir de 1945. Ele foi desenvolvido como

parte do movimento de abertura e transparência da justiça eleitoral e é voltado para pesquisadores, imprensa e pessoas interessadas em analisar os dados de eleitorado, candidaturas, resultados e prestação de contas (ELEITORAL, 2018).

O formato dos arquivos disponibilizados no portal é o CSV. Ao observar o conteúdo histórico, os arquivos estão disponíveis desde 1945. Porém, foram identificadas estruturas incompletas nos arquivos anteriores a 2003. Conforme explica (ELEITORAL, 2018), “está sendo realizada uma revisão nas fontes de dados e, conforme os trabalhos forem concluídos, os arquivos serão substituídos”, então para este trabalho delimitamos a extração das candidaturas entre 2003 até 2018.

A interação com o repositório se dá através do site institucional de dados abertos do TSE e o usuário pode acessar o ano da candidatura e obter um arquivo compactado no formato ZIP. Este arquivo contém o conjunto de arquivos no formato TXT referente a cada estado brasileiro e acompanha também um arquivo LEIAME com a documentação sobre o layout e as observações sobre possíveis mudanças na estrutura do documento naquela versão.

### 6.1.3 Processos de Extração e Transformação e Carregamentos dos Dados

Através da ferramenta Pentaho Data Integration (PDI) em sua versão Kettle, foi possível realizar as transformações, que são fluxos de trabalho de dados que representam suas atividades ETL. Os processos de ETL são implementados a partir das funcionalidades *transformation* e *job*, sendo que cada *job* pode conter um conjunto de *transformation*. As transformações são armazenadas nos arquivos KTR. Todos os arquivos KTR estão disponíveis ordenadamente a sua sequência de execução no repositório Git<sup>3</sup>.

O processo de ETL para dimensão DIM\_POLITICO foi realizado de forma que as tabelas parlamentares, oriundas do grupo transparência, e a tabela de candidatos, cuja base é mantida pela UFS, fossem mescladas. A justificativa foi que alguns candidatos, embora não tivessem mandatos eletivos, precisariam constar na base de políticos. A Figura 6.17 mostra o processo formalizado. Observe que neste *job* temos 3 transformações: a transformação que busca as chaves/identificadores dos políticos que já possuem um mandato eletivo; seus dados complementares; e a transformação dos dados dos candidatos. Os fluxos que são identificados com um sinal de “=”, significam que podem ser executados em paralelo. Para este *job*, foram inseridos os componentes para escrita de *log*, antes, durante e após cada *transformation*. Dessa forma foi possível observar algum erro ou anomalia no processo de carga e na ocorrência de um evento deste tipo, onde o fluxo é direcionado para o componente “Cancela *job*”.

<sup>3</sup>Repositório Github com a sequência de execução, disponível em: <http://bit.ly/2PGfJ8P>

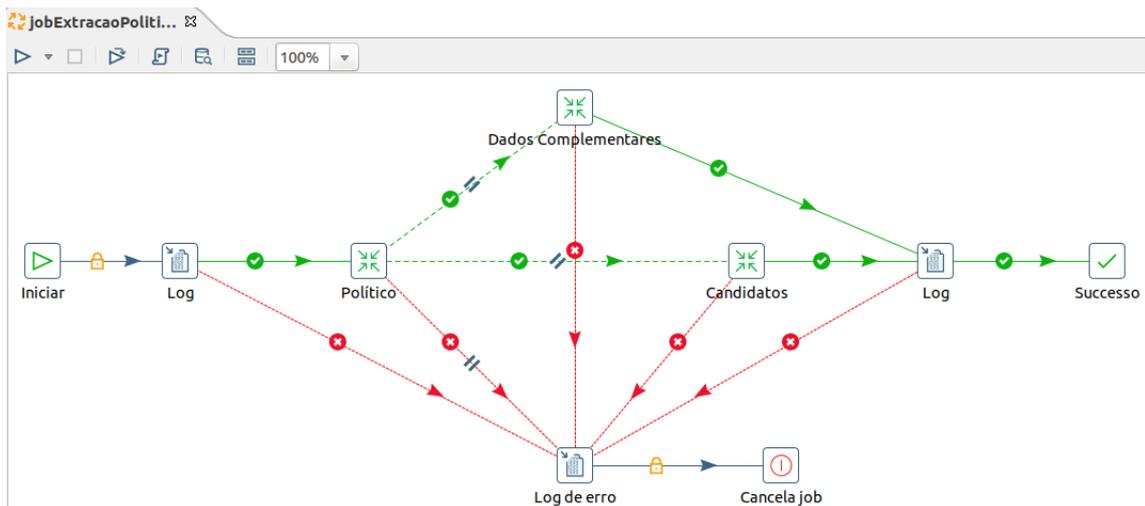


Figura 6.17: Processo ETL para a dimensão DIM\_POLITICO

Em seguida, foi criada a transformação para a extração dos processos jurídicos, cuja fonte de dados é a base da UFS. Foram coletados 1157 processos que apresentam o envolvimento dos políticos. A dimensão DIM\_PROCESSO contém o número do processo, a descrição do assunto, o tipo do processo, o número, o identificador do político, o tribunal de origem e o *link* que direciona para uma consulta pública ao sistema transacional que controla os trâmites de tais processos. A Figura 6.18 apresenta o processo ETL.



Figura 6.18: Processo ETL para a dimensão DIM\_PROCESSO

Conforme explicado na subseção 6.1.2.1, os mandatos dos políticos não estavam claramente organizados na base. Foi necessário extrair as informações de mandato contidas na propriedade cargo de cada político. A Figura 6.19 apresenta a transformação implementada para recuperar os dados de mandato político. A primeira parte do processo consiste na seleção dos cargos. Em seguida é realizada a inclusão de delimitadores “;” para extração das sentenças, conforme apresentado na figura 6.10. Os componentes Mandato1, Mandato2, Mandato3 e Mandato4 contém a implementação das expressões regulares em *javascript*. O resultado da codificação apresentada na Figura 6.11 foi a seleção dos períodos de mandatos eletivos de cada político. A partir de então foi possível fazer a extração de cada ano de mandato, desconsiderar os registros nulos e incluir os registros do mandato na base.

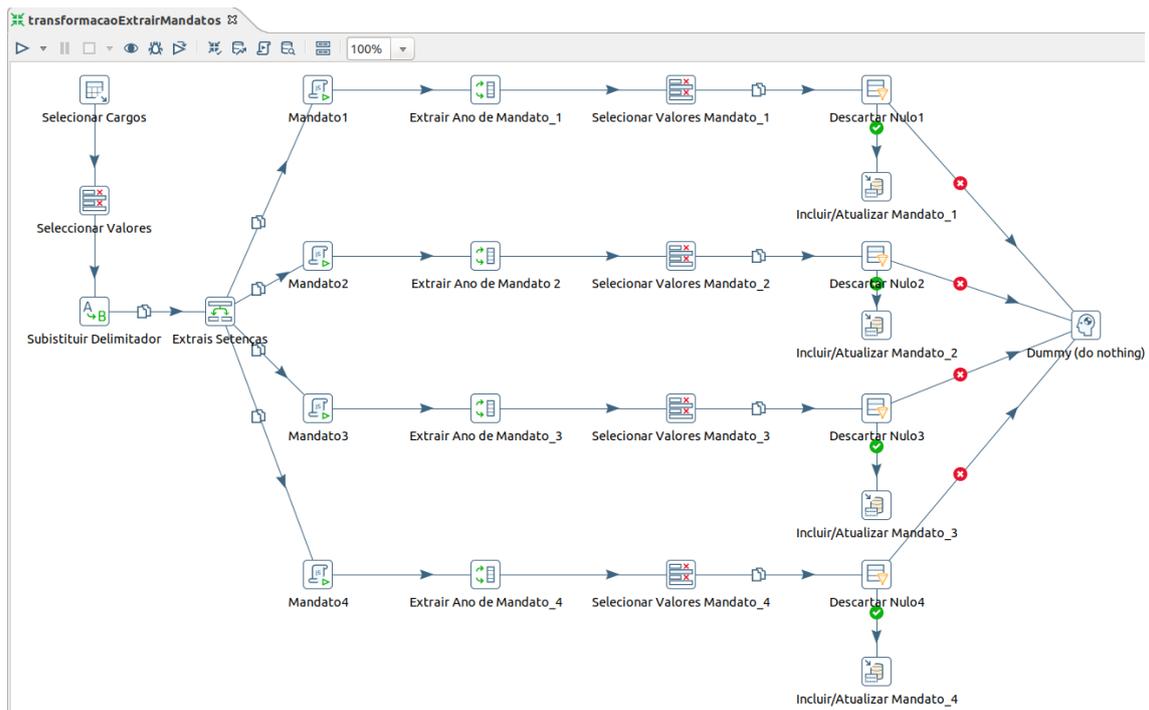


Figura 6.19: Recuperação dos mandatos a partir das descrição do cargo

Para a transformação do FATO\_MANDATO, foi necessária a criação de outro processo conforme mostra a Figura 6.20.

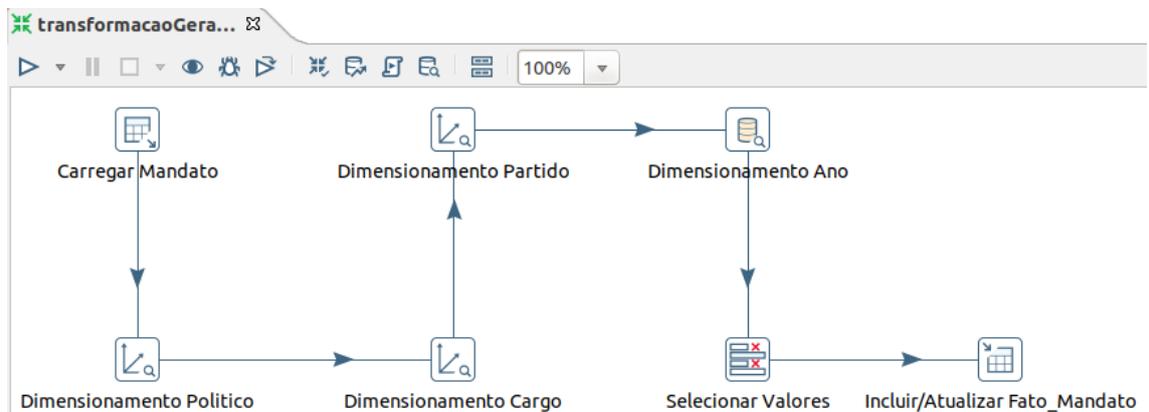


Figura 6.20: Processo ETL para a FATO\_MANDATO

Ao extrair dos dados de candidaturas e suas respectivas receitas/doações e despesas, foi necessário criar um arquivo KTR para os anos de 2014, 2016 e 2018. A interface do Portal de Dados Abertos do TSE, contém menus organizados por ano e cada menu direciona a uma interface que contém os arquivos compactados para cada categoria de informação. A Figura 6.21 ilustra as transformações implementadas com o link apontando para cada arquivo CSV que contém as informações das candidaturas acerca de cada estado brasileiro.

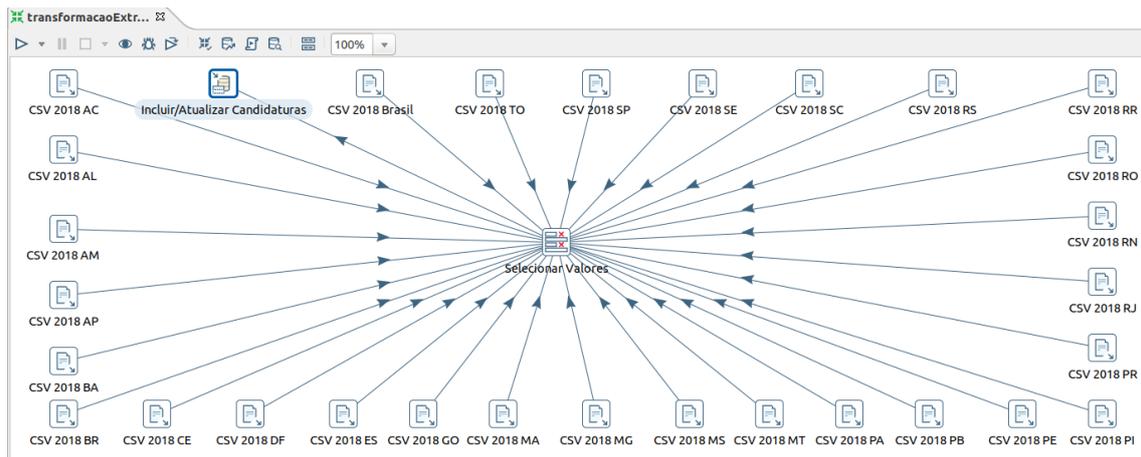


Figura 6.21: Transformação dos dados de Candidaturas

A partir dos dados extraídos do portal foi possível a construção do processo ETL para o carregamento da FATO\_CANDIDATURA conforme apresenta a Figura 6.22.

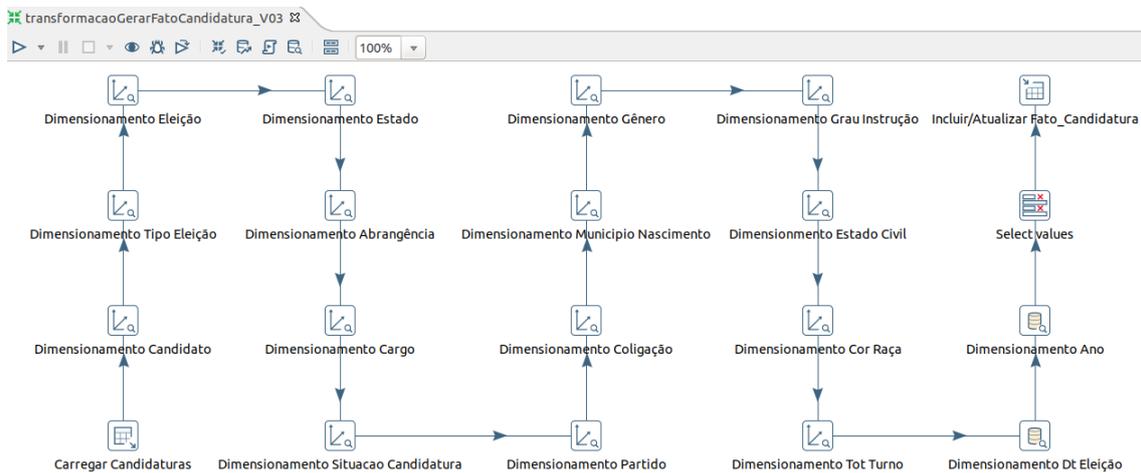


Figura 6.22: Processo ETL para a FATO\_CANDIDATURA

Para o carregamento da FATO\_EMENDAS, foi criada uma transformação para emendas apresentadas pelos parlamentares dentre os anos de 2011 e 2014. A Figura 6.23 mostra o processo ETL.

A Figura 6.24 apresenta a transformação dos dados que representam as atividades econômicas executadas por cada político. Para este modelo multidimensional, foram selecionados apenas os registros que possuíam uma divisão/categoria associada à atividade econômica.

Para extrair os dados referente às verbas utilizadas por cada parlamentar, também foram coletados os tipos de verba e materializados em uma dimensão denominada DIM\_TIPO\_VERBA. O objetivo é identificar qual a finalidade da verba gasta. O processo ETL é ilustrado na Figura 6.25.

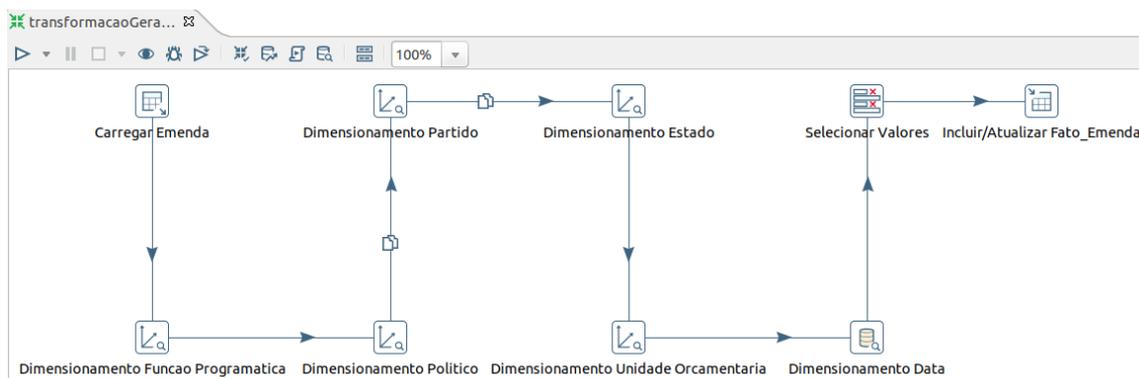


Figura 6.23: Processo ETL para a FATO\_EMENDA

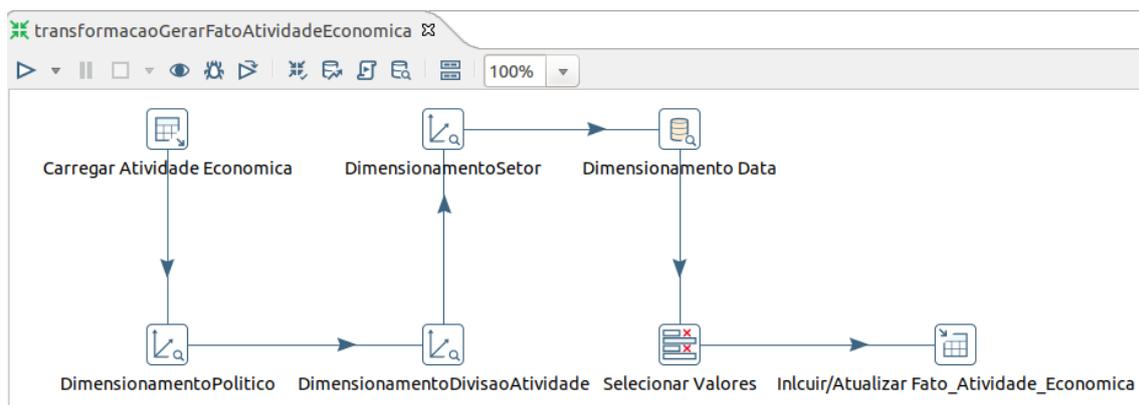


Figura 6.24: Processo ETL para a FATO\_ATIVIDADE\_ECONOMICA

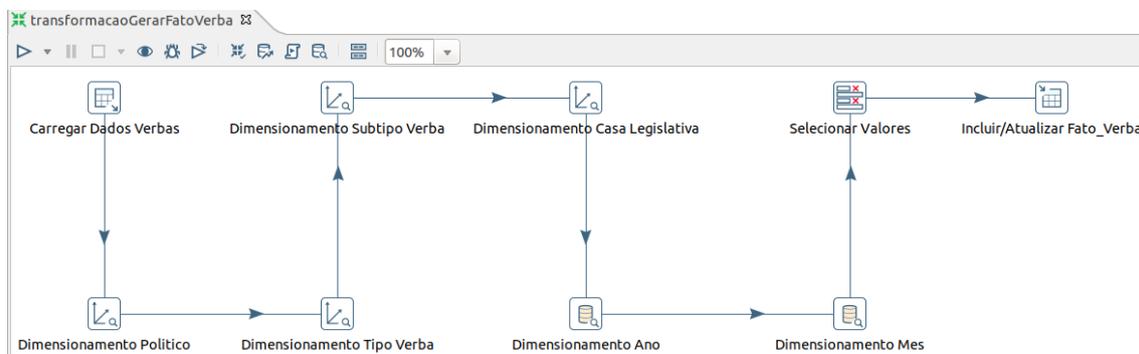


Figura 6.25: Processo ETL para a FATO\_VERBA

No modelo FATO\_VIAGEM, conforme mostra a Figura 6.26, concentramos todos os dados referentes aos gastos com as viagens realizadas pelos políticos. Assim como no modelo FATO\_VERBAS, aqui foi incluída uma dimensão DIM\_JUSTIFICATIVA para a identificação do propósito da viagem.

Outro dado que faz parte da produtividade parlamentar, são as proposições relativas às matérias que são levadas ao plenário. Os dados foram transformados e foi criada a dimensão DIM\_MATERIAS\_PUBLICADAS e a tabela FATO\_MATERIA. O conjunto de dados inclui o sumário das matérias propostas entre 2003 e 2016. Elas são classificadas

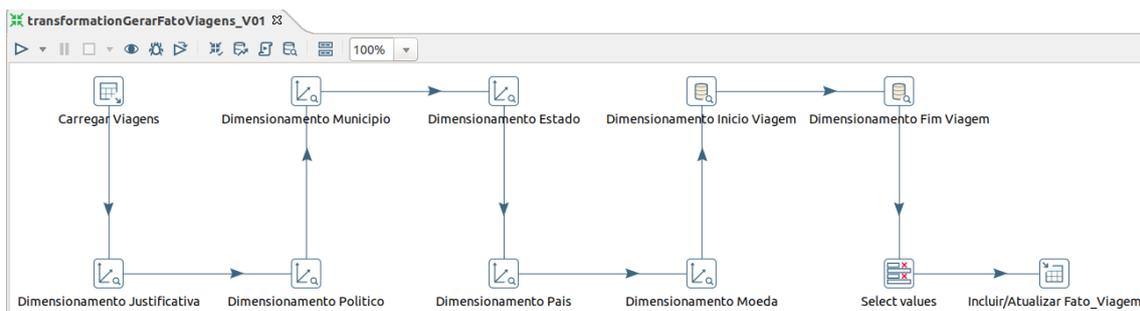


Figura 6.26: Processo ETL para a FATO\_VIAGEM

em categorias temáticas (Educação, Saúde, Agropecuária, Licitações/Contratos, Regulação Política, Regulamentação de Profissões, etc). A Figura 6.27 apresenta o *job* utilizado para carregamento da dimensão DIM\_MATERIAS\_PUBLICADAS e inclui o autor da matéria.

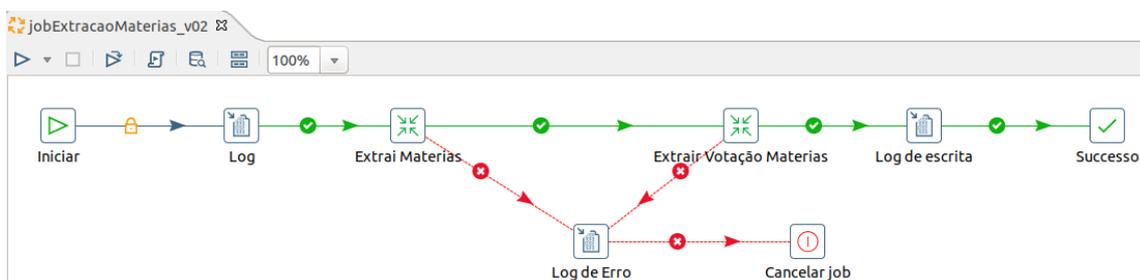


Figura 6.27: Transformação dos dados de matérias apresentadas ao plenário

Já a Figura 6.28 ilustra o processo de transformação das matérias em voto, ou seja, identifica quem votou e o status da participação e o voto do político.

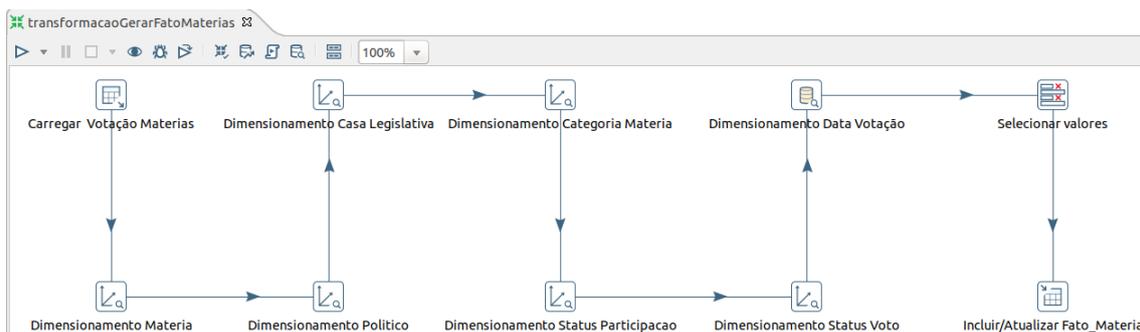


Figura 6.28: Processo ETL para a FATO\_MATERIA

Até o momento, foram apresentados os modelos dimensionais, as fontes de dados, os participantes que permitiram o acesso a essas informações, principais processos de ETL e algumas limitações que identificamos ao longo dessas atividades.

## 6.2 O Sistema de Informações Prosopolíticos

Após a criação do modelo FIV.Br2018 (Estudo Exploratório), a definição do Modelo Conceitual, a projeção dos modelos Multidimensionais e a Execução dos Processos de ETL para o ambiente de Data Warehouse, determinou-se a disponibilização dos dados em um sistema de informações para a consulta do candidato e para a avaliação pelo cidadão. A esse sistema nominamos Prosopolíticos.

Para o desenvolvimento do sistema de informações, foram seguidas as diretrizes conforme a (LAI, 2011), que determina que o acesso ao dado deve ser único, de fácil utilização e estruturado. Outro ponto importante no que tange a acessibilidade foi a utilização do software VLibras. Segundo (ECONOMIA, 2018), o software consiste em um conjunto de ferramentas computacionais de código aberto, responsável por traduzir conteúdos digitais (texto, áudio e vídeo) em Língua Brasileira de Sinais - LIBRAS.

Com o propósito de rastrear os dados resultantes da trajetória de construção do Sistema de Busca, os dados contidos em cada modelo multidimensional foram dispostos dentro da própria estrutura da interface. Dessa forma, a implementação dos requisitos e a disposição das informações na tela foram orientadas com base nos fatores elencados no modelo FIV.Br2018.

Além dos requisitos relacionados aos modelos multidimensionais, outros requisitos foram incluídos, tais como o tratamento de *Trending Topics* para verificar quais eram as tendências das consultas realizadas e as cinco últimas consultas efetuadas no sistema. Além disso, a observância do uso do software na internet foi possível com a inclusão do código de acompanhamento do Google Analytics. A Figura 6.29 apresenta a tela inicial do software.

O Sistema Prosopolíticos foi estruturado na seguinte forma:

- Página de entrada: sinalização de que o sistema está configurado para utilização do aplicativo VLibras disposto do lado direito da tela.
- Campos de busca: busca com quatro (4) parâmetros: nome, cargo eletivo, partido e estado de nascimento.
- Sobre o sistema: uma descrição sobre o sistema de consulta e seus objetivos, bem como da natureza dos dados prosopográficos.

A tela de resultados da busca apresenta um conjunto de registros com referência aos no-



Figura 6.29: Interface de Consulta do Sistema [www.prosopoliticos.uff.br](http://www.prosopoliticos.uff.br)

mes completos do Candidato(s) e seu(s) respectivo(s) nome(s) de urna. O link, quando acionado, conduz o usuário a tela com a informações sobre o político conforme mostra a Figura 6.30

Seguidamente as informações pessoais do candidato, estão os fatos selecionados e apresentados em tópicos. Vale ressaltar que, cada tópico somente é exibido, se houver o registro do candidato na tabela fato.

Quanto aos testes de verificação do sistema, o *framework* Django realiza algumas análises antes de executar o sistema. De acordo com (DOCUMENTATION, 2019), o Django possui um conjunto de verificações estáticas para validar o projeto. Ele detecta problemas comuns e fornece dicas de como corrigi-los. Todas as vezes em que ocorrem alterações em campos, modelos, gerenciadores de modelos e nas estruturas do banco de dados, o processo de *check* é acionado em busca de erros sintáticos ou semânticos. Para os testes lógicos, o sistema foi testado em diferentes cenários

Prosopolíticos - Mozilla Firefox

Prosopolíticos x +

www.prosopoliticos.uff.br/ 80% Search

PROSOPOLÍTICOS Nome Procurar candidato



**Jair Bolsonaro**  
PRESIDENTE PELO PSL

DADOS PESSOAIS:  
Jair Messias Bolsonaro

CARGOS RELEVANTES:  
Presidente

BANCADA POLÍTICA:  
Parentes

POSICIONAMENTO POLÍTICO:  
Terceirização de Todas as Atividades (1998): Fica permitido, a partir de agora, que empresas terceirizem qualquer tipo de mão de obra, inclusive as chamadas atividades fim. [1][2][3] Sua posição foi indefinido ao assunto.

Nascido em 21/03/1955 (64) na cidade de Campinas - São Paulo

Histórico de partidos: PDC, PP, PPR, PPB, PTB, PFL, PP

Membro das Forças Armadas

jairbolsonaro.jb.2018@gmail.com  
<https://www.camara.leg.br/deputados/74847>

Superior Completo

Processos

Matérias

Projetos

Atividade Econômica

Verba

Candidaturas

Emendas

Mandatos

Figura 6.30: Resultado da busca: registro ampliado

## 7. Avaliação

Este capítulo apresenta a avaliação da solução proposta descrevendo o processo de coleta de dados, a seleção e interação com os entrevistados, o contexto e a codificação realizada.

### 7.1 Os Sujeitos Entrevistados

A avaliação do artefato foi realizada nas cidades de Niterói e Rio de Janeiro e por meio da amostragem por conveniência, selecionamos 5 participantes. Este grupo foi composto por perfis diferentes quanto a idade, inclinação/pensamento político, contato com a internet, área de formação e local de residência. A tabela 7.1 apresenta o perfil dos participantes.

Tabela 7.1: Perfil dos Entrevistados

Idade	Inclinação Política	Experiência com Internet e Tecnologias	Área de Formação	Local de Residência
74	Direita	Iniciante	Linguística, Letras e Artes	Petrópolis
33	Centro	Avançada	Ciências Humanas	Volta Redonda
42	Centro	Avançada	Engenharia e Tecnologia	Niterói
46	Esquerda	Intermediário	Ciências da Saúde	Rio de Janeiro
45	Esquerda	Avançada	Engenharia e Tecnologia	Rio de Janeiro

As entrevistas foram feitas individualmente e cada participante recebeu um *e-mail* contendo as instruções de participação e o cenário da pesquisa. Após a confirmação da participação, as datas das entrevistas foram agendadas.

#### 7.1.1 Seleção do Contexto

Foi apresentado, a cada participante, uma lista com 18 políticos. Essa lista foi organizada em 3 grupos de pessoas relacionados ao posicionamento político. Um grupo de candidatos da esquerda, um outro grupo somente de candidatos da direita, e um terceiro grupo de

candidatos do centro.

Foi solicitado que cada participante verificasse na lista se havia algum candidato que já conhecia. Caso houvesse, esses candidatos seriam excluídos do grupo. O objetivo foi selecionar uma tarefa para o participante em que ele tivesse que realizar a busca por informações sobre um candidato desconhecido. Definida a lista de políticos desconhecidos dos participantes, foram selecionados 4 candidatos. A seleção foi aleatória dentro de cada grupo de posicionamento político.

O estudo iniciava com a seguinte mensagem enviada por *e-mail*:

“Estamos nos aproximando das eleições de 2020 para o cargo de presidente da república e, para esse ano, você terá 4 (quatro) candidatos disputando o mandato eletivo.

Com o objetivo de apoiar a sua busca pelas informações, disponibilizamos um sistema de consulta aos candidatos. Você já ouviu falar do site [www.prosopoliticos.uff.br](http://www.prosopoliticos.uff.br), que possibilita realizar pesquisas sobre políticos? Use esse sistema e saiba mais sobre a vida dos candidatos que disputam à presidência do Brasil.”

### **7.1.2 Observação do uso do sistema e realização das entrevistas**

A interação do usuário com o sistema desenvolvido nesta pesquisa foi gravada com o uso do software Camtasia. Com este recurso foi possível verificar todos os passos de interação com o sistema pela gravação da tela do computador durante o uso pelo participante da pesquisa, bem como a gravação do rosto do participante capturado pela *webcam*. O objetivo desse recurso foi capturar as informações em tempo real e observar as feições dos usuários e os caminhos percorridos durante a pesquisa.

Com a autorização de cada participante, as entrevistas foram gravadas via aparelho de celular. Paralelamente, foi utilizada a funcionalidade de digitação por voz do Google Docs. Dessa forma ao final da entrevista já coletamos a primeira versão da transcrição.

Após a realização das entrevistas, foi iniciado o processo de análise/interpretação por meio da técnica de codificação. De acordo com (CORBIN; STRAUSS, 2014) a microanálise, ou análise linha a linha, é um pré-requisito para a codificação. Revisitando (CRESWELL; POTH, 2016), o autor explica que a codificação é o processo pelo qual os dados são reduzidos a elementos significativos de acordo com os objetivos da análise empreendida e são apreendidos nomes ou rótulos para sua identificação.

A *Grounded Theory* (GT) Construtivista, (CHARMAZ, 2009) especifica 3 tipos de codi-

ficações: a inicial, a focalizada e a teórica. De acordo com a autora, a codificação inicial, se fixa com rigor aos dados, dessa forma, esse ciclo é orientado a imersão e análise do texto.

O próximo ciclo compreende, a codificação focalizada (axial), descrita por (LEITE, 2016) como:

Os códigos produzidos nesse ciclo são mais direcionados e seletivos que os da etapa inicial. Para o seu realizar, são utilizados os códigos iniciais mais significativos e/ou frequentes para analisar minuciosamente grandes quantidades de dados. Essa codificação exige tomada de decisão sobre quais dados permitem uma compreensão analítica melhor para categorizar os outros dados de modo pleno. De outra forma, nessa etapa se definem quais dados têm a potencialidade de se coadunar com outros formando, assim, uma categoria.

O processo de codificação foi realizado com o apoio do *software* ATLAS.ti<sup>1</sup>. Amplamente recomendado por (CORBIN; STRAUSS, 2014), o ATLAS.ti é um ambiente de trabalho para a análise qualitativa de grandes corpos de dados textuais, gráficos, de áudio e vídeo. A Figura 7.1 apresenta uma demonstração da codificação realizada por meio deste software. Nela, o trecho realçado em azul é uma citação. Ao criar uma citação, o trecho é identificado como uma barra na vertical, do lado direito da transcrição. Ao criar uma categoria, ela fica disponível na lista de códigos (lista marcada no retângulo em vermelho a esquerda). Basta arrasta-la para frente do texto transcrito, que a codificação axial é identificada com o nome da categoria na horizontal, conforme aponta a seta da “Codificação Focalizada (Axial)”.

Durante a codificação axial, foram criados alguns grupos principais de códigos: educação\_política, facilidade\_uso, fator\_influencia, fonte\_informacao, intencao\_uso, redes\_sociais\_digitais, utilidade\_percebida. À medida que as transcrições e as codificações ocorriam, ao comparar as citações<sup>2</sup>, novos códigos foram sendo construídos. A relação de todos os códigos e suas descrições está disponível no Apêndice F.

Paralelamente à codificação focalizada e aberta, (CORBIN; STRAUSS, 2014) recomendam a prática da criação de memorandos, um “armazém de ideias”, cujo objetivo é registrar as associações, as percepções e os questionamentos que surgem ao interpretar os

<sup>1</sup>ATLAS.ti é um ambiente de trabalho para a análise qualitativa de grandes corpos de dados textuais, gráficos, de áudio e vídeo. <https://atlasti.com/product/what-is-atlas-ti/>

<sup>2</sup>No ATLAS.ti, cada trecho selecionado é denominado citação, e cada citação pode ou não estar relacionada a no mínimo um código.

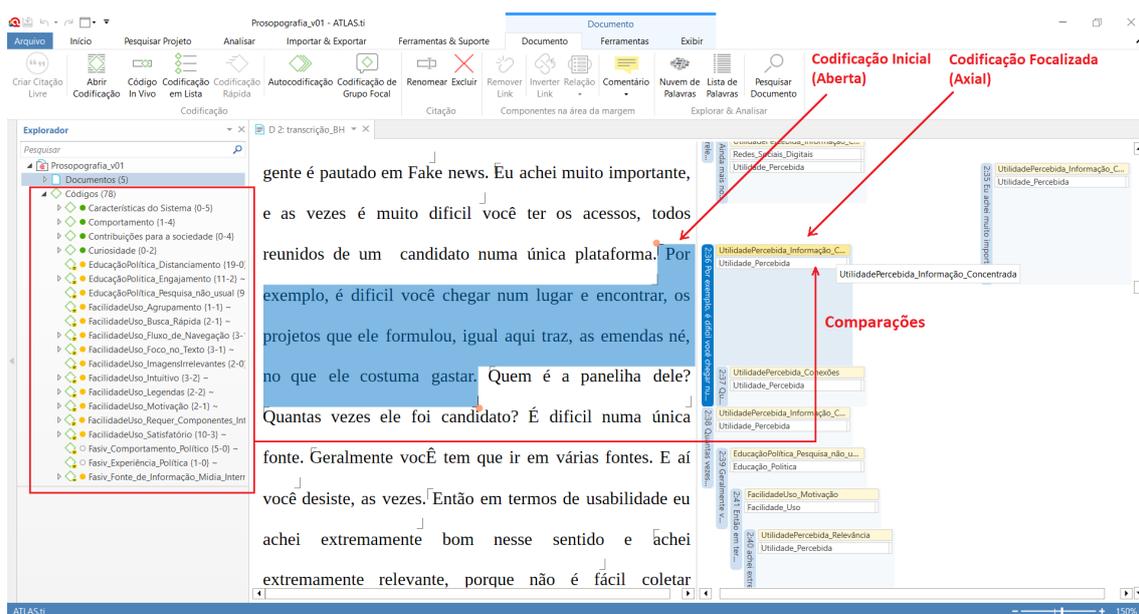


Figura 7.1: Interface do *software* ATLAS.Ti apresentando as etapas de codificação inicial e axial.

dados. Então, para cada transcrição e codificação, foi criado um memorando, bem como os comentários para dar significado aos códigos criados.

A terceira etapa compreende a codificação teórica. De acordo com (TAROZZI, 2011), é nessa fase que a construção das categorias alcança plenitude e a teorização procede para a identificação das categorias centrais, os conceitos-chave em torno dos quais se organizará a teoria. Em (CHARMAZ, 2006), a autora reforça que a formalização da categorização central pode ser representada graficamente, através de diagramas/mapas que revelam situações e processos.

Em busca da categoria teórica/central, além dos registros dos memorandos e todos os comentários descritos tanto nos códigos quanto nos documentos das transcrições, foram utilizados os relatórios gerados pelo ATLAS.ti agrupados pelos códigos e suas citações. Ademais, a construção das redes de códigos contribuiu para a construção da categoria central emergente da pesquisa. A figura 7.2, apresenta um trecho do relatório gerado e de uma rede criada a partir do relato de um dos entrevistados.

Para melhor entendimento e orientação da análise dos dados, as principais questões foram sistematizadas na figura 7.3.

Nessa figura, o grupo marcado em azul representa as características da TAM como tópicos que originaram as questões. Em verde, destacam-se as questões que foram utilizadas como condutores, durante as entrevistas. Em laranja, as principais proposições utilizadas na análise, isto é, as questões que devem ser investigadas para possibilitar uma avali-

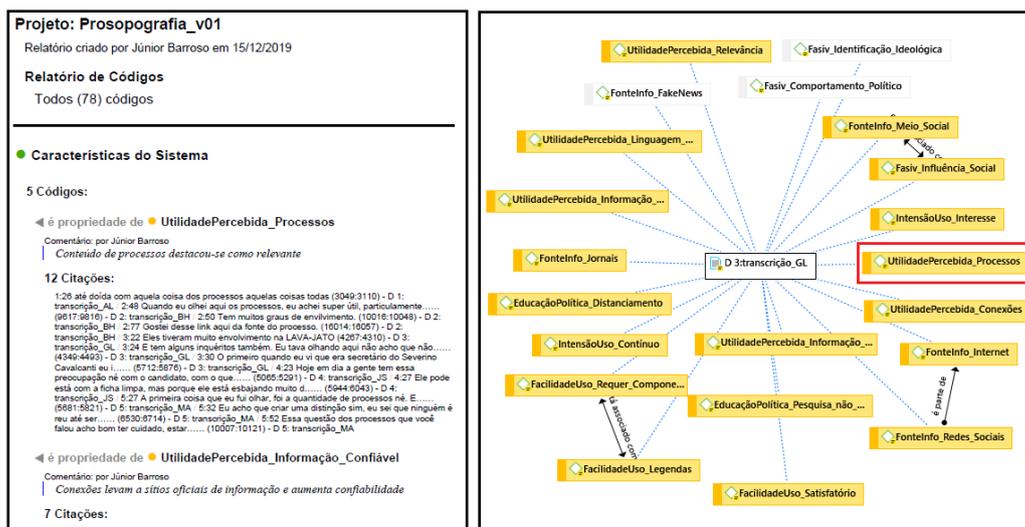


Figura 7.2: Relatório de códigos e criação de uma rede de categorias a partir de um documento

ação do artefato e das conjecturas comportamentais conforme apresentado na instância do Modelo-DSR desta pesquisa. Em preto, as categorias construídas durante a etapa de codificação.

## 7.2 Interpretação dos Resultados

Diante da finalidade do artefato desenvolvido, de promover a transparência e suporte às necessidades informacionais dos cidadãos, a interpretação dos dados iniciou em torno das fontes de informações acerca dos políticos.

A Figura 7.4 mostra as principais fontes utilizadas para busca de informações sobre o cenário político e os candidatos.

Esta rede expõe as conexões e as características dos algoritmos (como sistemas de recomendações) que levam informações ao cidadão. O modelo FIV.Br2018 também compõe a rede juntamente com o fator Influência no meio social a que pertence o cidadão. A intenção é identificar, através dos códigos, as alternativas utilizadas como fonte de informação em tempos de cibercultura, quais as percepções dos entrevistados sobre tais fontes, bem como identificar os hábitos de acesso às fontes governamentais.

As principais fontes de informação política elencadas foram: “meio social”, “internet”, “jornais”, “redes sociais” e as “propostas” de governo. Todas as fontes se relacionam, e o “meio social” foi a fonte de informação mais mencionada entre os participantes (13 no total), seguida de “Internet” (*sites* de política), “Jornais” (TV e *On-line*), “Redes Sociais”

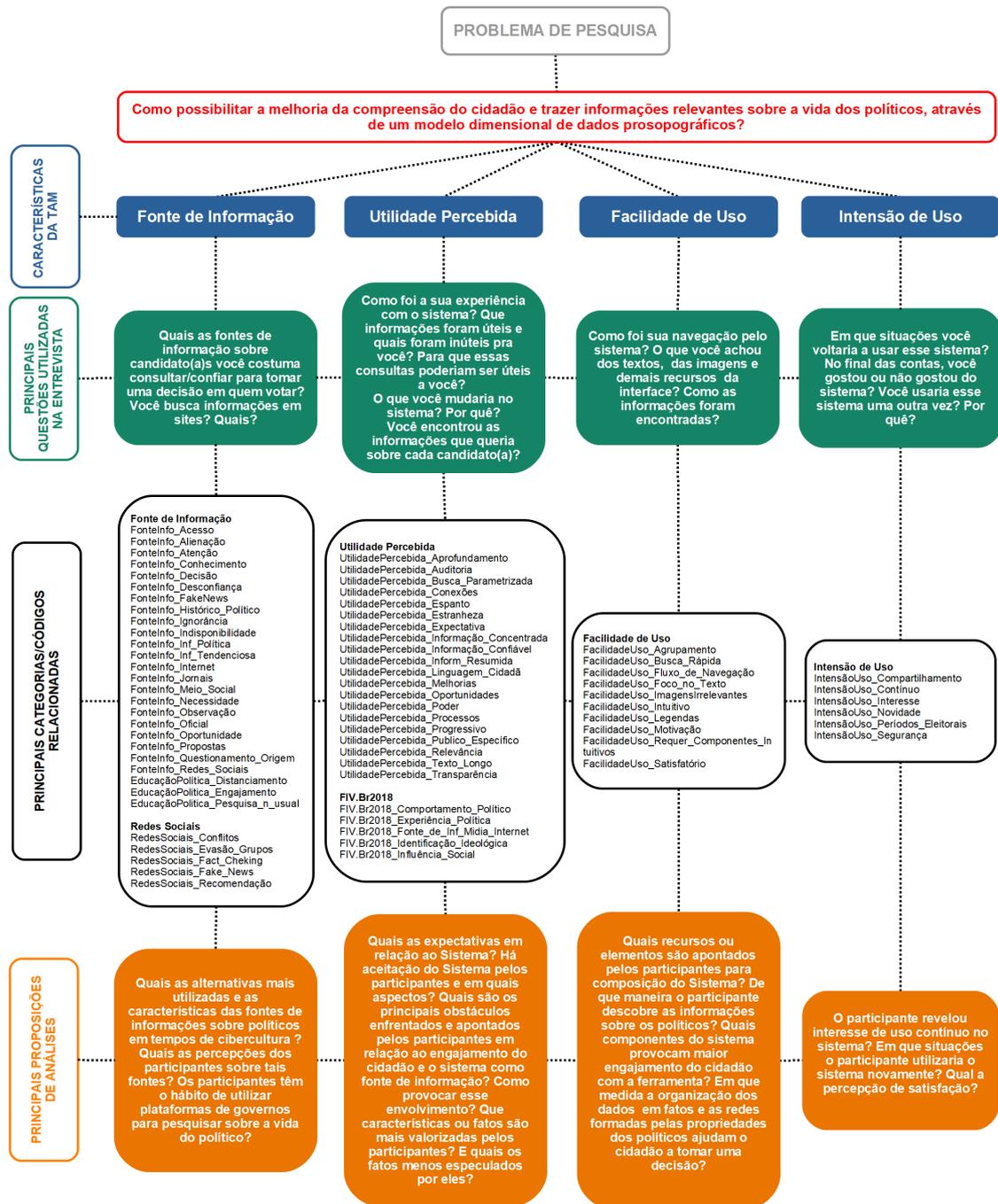


Figura 7.3: Planejamento da análise dos dados

e as “Propostas” de governos. Sobre o meio social, ocorre a observância tanto dentro quanto fora das redes sociais. Conforme o que foi relatado, as pessoas buscam informações entre si. Geralmente através de uma relação de confiança e pela percepção que a outra pessoa tem sobre o entendimento político ou por ser melhor informado.

O comportamento do usuário nas redes sociais, os rastros, as curtidas e as movimentações nas redes, provocam nas pessoas o interesse em interagir nos mesmos canais. O modelo FIV.Br2018 está presente na rede como forma de justificar alguns fatores que foram men-

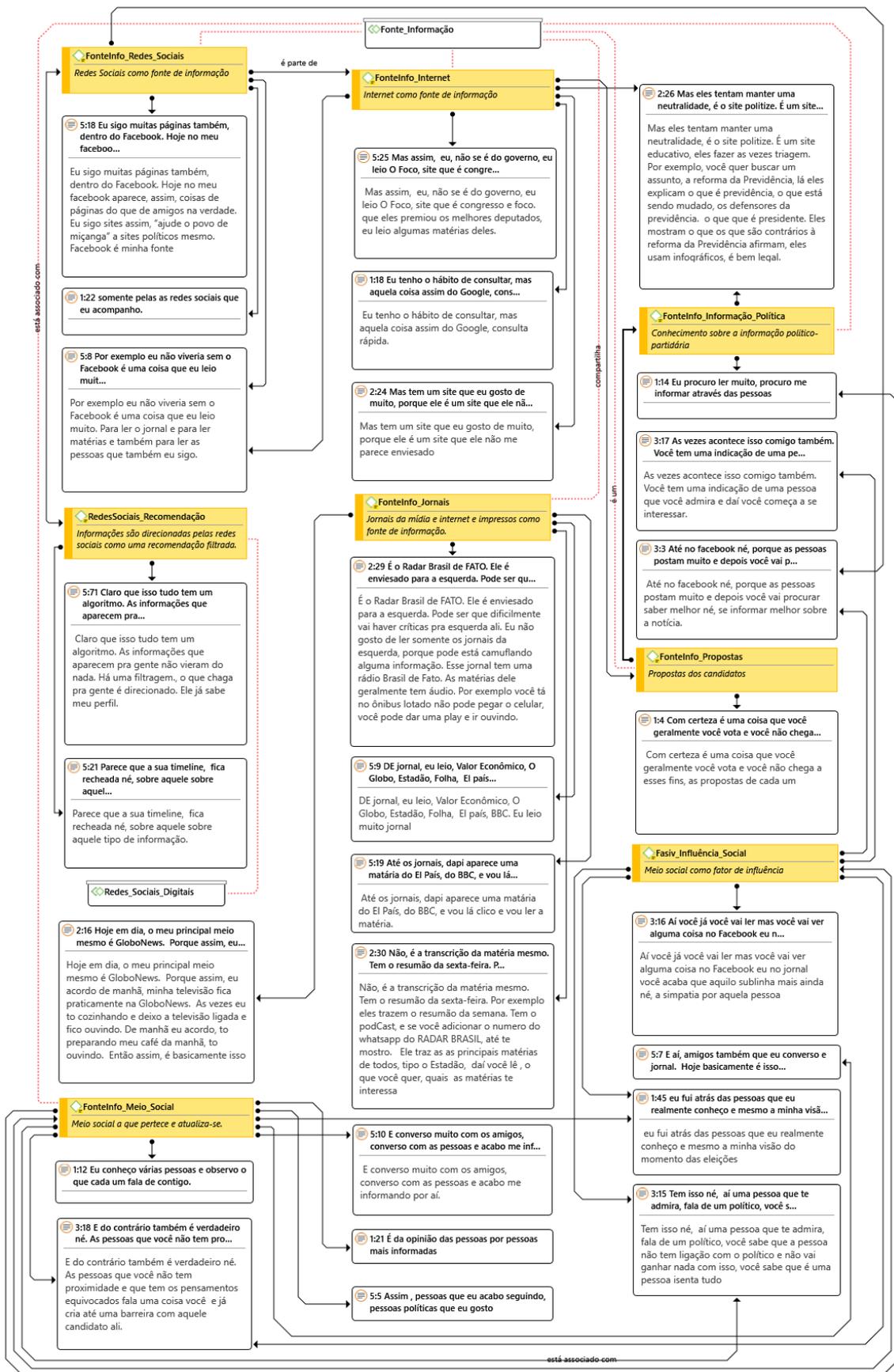


Figura 7.4: Rede sobre as fontes de informações sobre políticos. Fonte: do próprio autor,

cionados, como é o caso do meio social como fonte de informação e como influência para decisão do voto.

Sobre a categoria “Recomendação”, há o reconhecimento dos filtros executados pelos sistemas de recomendação que pulverizam a *timeline* das plataformas digitais, conforme apontado: *claro que isso tudo tem um algoritmo. As informações que aparecem pra gente não vieram do nada. Há uma filtragem, o que chega pra gente é direcionado. Ele já sabe meu perfil* (MA) . Há uma preocupação em selecionar o veículo de informação que apresenta uma neutralidade e não é um hábito a consulta sobre a vida dos políticos em plataformas de governo.

A Figura 7.5 mostra que a internet é o meio em que as informações sobre a política e os políticos chegam aos cidadãos através de suas redes e suas bolhas.

Foi mencionado que há um desconforto em relação a confiabilidade das informações, e isso afeta as construções ou desconstruções das relações sociais.

No primeiro nó da rede, temos a categoria “Internet” como fonte de informação e as “Redes Sociais” como parte desse meio, onde ocorre a prática da disseminação das “*Fake News*”. Essas, por sua vez, movimentam as redes sociais com conteúdos que provocam uma série de questionamentos construtivos ou destrutivos e criam um cenário de dúvidas e desinformação. Por isso elaboramos a categoria “Desconfiança”. Relacionado a essa categoria, foi mencionada a falta de credibilidade nas informações dos sites dos próprios candidatos: *vai estar ali na página dele e tudo serve de algum conteúdo pra saber algo de bom que ele fez* (JS). As informações contidas nessas plataformas pessoais foram consideradas tendenciosas e expressam apenas o lado positivo do perfil do político.

Os componentes dessa rede reportam o desconforto em permanecer em grupos virtuais, pois as manifestações giram em torno de “Conflitos”: *então eu praticamente não utilizo a rede social ou mídia para política, porque eu não tenho paciência para os debates que as vezes são gerados lá* (BH). A categoria “Evasão” revela a saída dos membros dos grupos, provocada pelo desconforto mental do participante.

Também foi mencionado sobre o cuidado em checar um conteúdo suspeito antes de compartilhar ou se posicionar diante de uma informação aparentemente tendenciosa ou falsa. Embora o número de citações relacionadas seja muito baixo, a categoria “Propostas” também emergiu como fonte de informação, o que vai ao encontro do modelo FIV.Br2018, apresentado na subseção 4.3 do capítulo 4, que contém as propostas de governo como fator de influência.

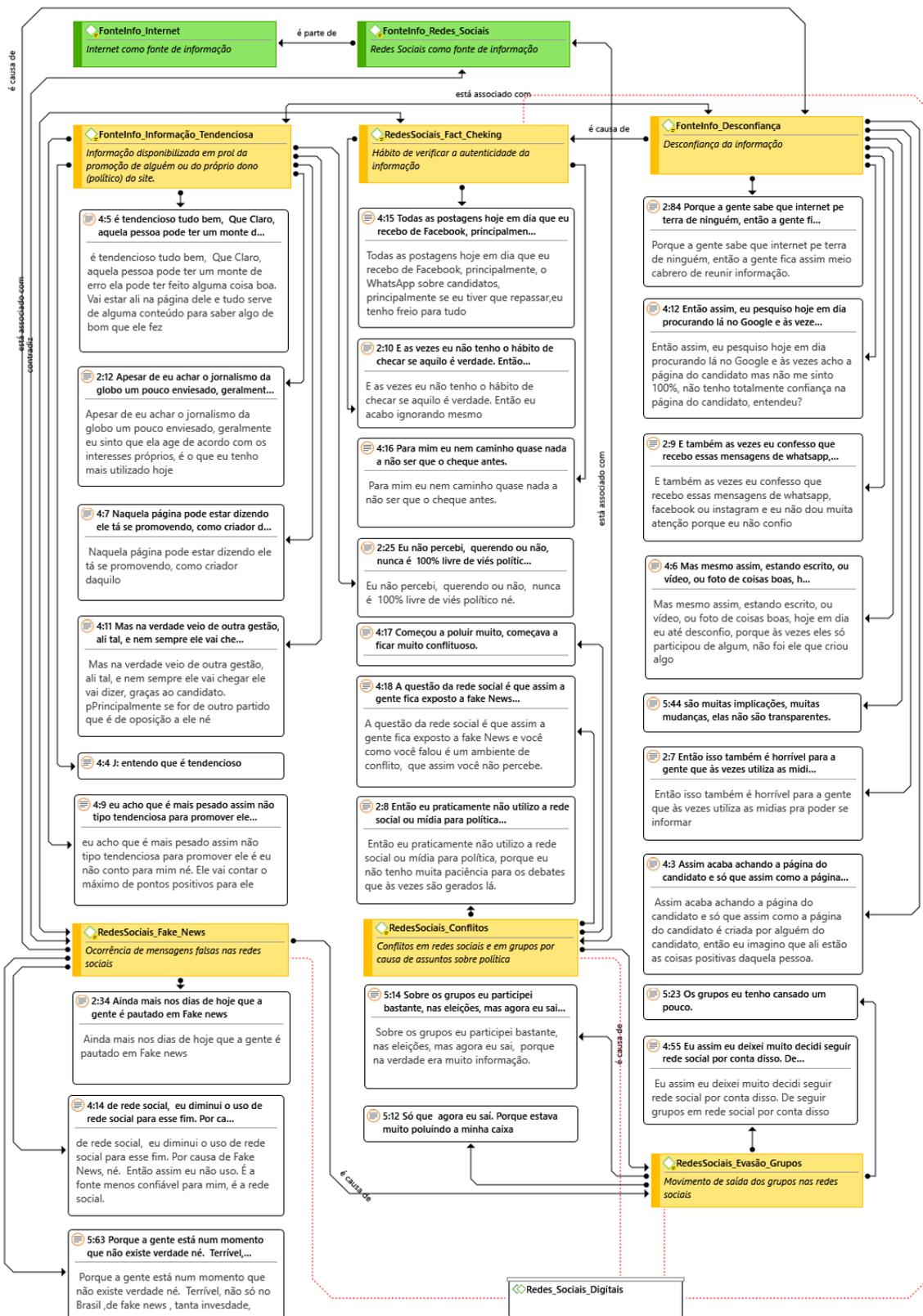


Figura 7.5: Rede sobre as redes sociais e fake news. Fonte: do próprio autor,

A Figura 7.6 mostra as principais categorias que emergiram sobre o comportamento dos entrevistados ao lidar com a pesquisa acerca das informações sobre políticos.

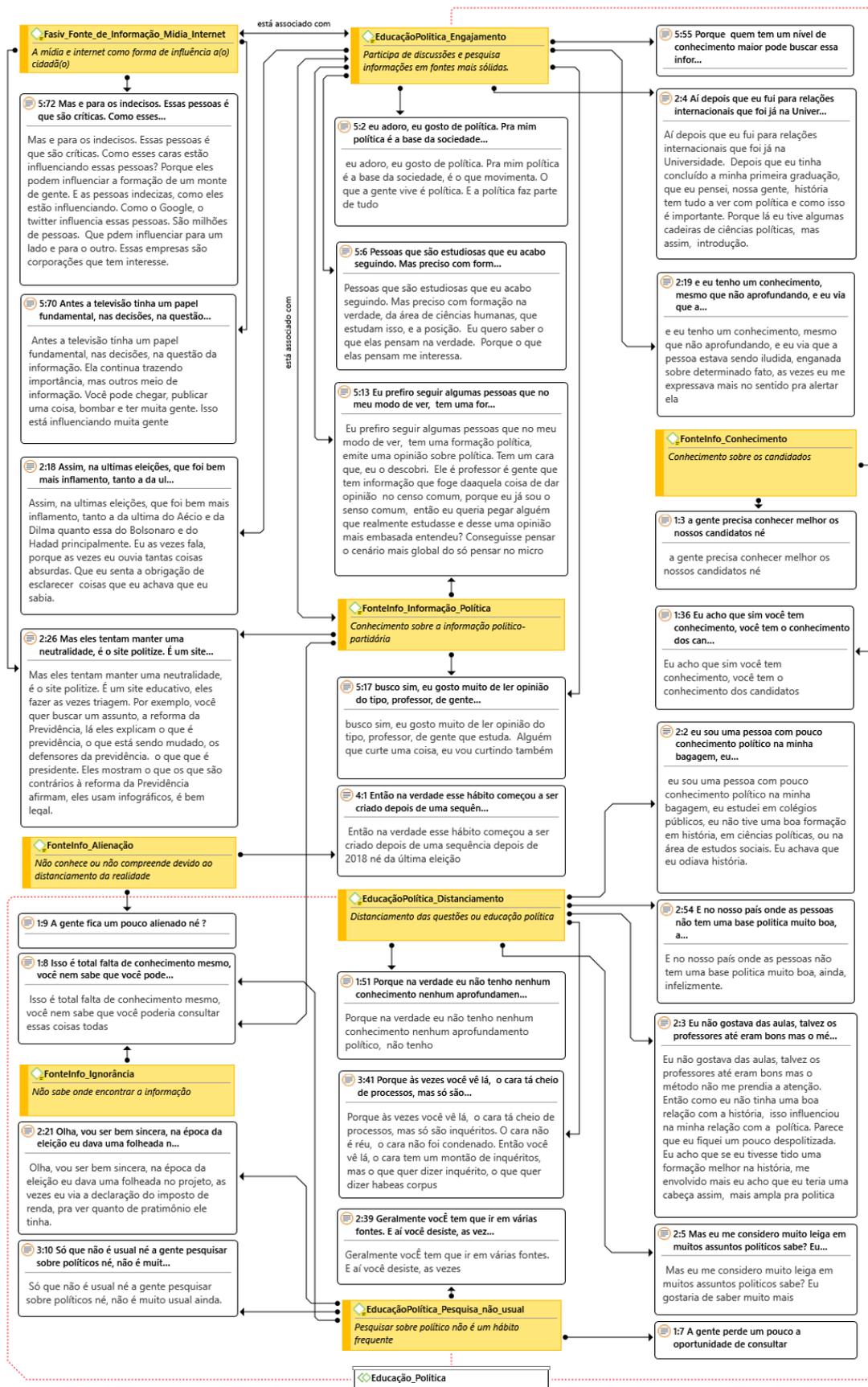


Figura 7.6: Rede sobre educação e engajamento político. Fonte: do próprio autor,

Os dados fizeram emergir a categoria “Distanciamento” como característica do perfil do cidadão em relação a educação política e estudos sobre a democracia e gestão do estado. Muitos se consideraram alienados e mencionam a falta de conhecimento, a falta de entendimento, e a dificuldade em lidar com a “Informação Política”. Os conhecimentos sobre política foram adquiridos em período posteriores à formação básica, conforme o relato a seguir:

*Eu não gostava das aulas, talvez os professores até eram bons mas o método não me prendia a atenção. Então como eu não tinha uma boa relação com a história, isso influenciou na minha relação com a política. Parece que eu fiquei um pouco despolitizada. Eu acho que se eu tivesse tido uma formação melhor na história, me envolvido mais, eu acho que eu teria uma cabeça assim, mais ampla pra politica (BH).*

Interpretamos que a falta de uma formação que inclua esses tópicos em seus currículos cria limitações em relação a pesquisa por informação sobre políticos e no entendimento dos resultados obtidos. Isso possivelmente configura um baixo engajamento com as ferramentas governamentais oficiais. Os dados revelam a categoria “Pesquisa não usual”, pois foi declarado que não é muito comum a utilização de ferramentas do governo que possibilitem investigar/pesquisar sobre a vida dos políticos.

Por outro lado é observada a crescente movimentação das redes sociais digitais, mídias e internet sobre a temática. Foi relatado que, na internet, as pessoas mais engajadas e que detém conhecimento sobre a temática política têm maior chance de serem seguidas.

*Eu prefiro seguir algumas pessoas que, no meu modo de ver, têm uma formação política, emite uma opinião sobre política. Tem um cara que, eu o descobri. Ele é professor, é gente que tem informação que foge daquela coisa de dar opinião no senso comum, porque eu já sou o senso comum, então eu queria pegar alguém que realmente estudasse e desse uma opinião mais embasada, entendeu? Conseguisse pensar o cenário mais global do que só pensar no micro (MA).*

Já a Figura 7.7 aponta três categorias que emergiram relacionadas ao comportamento dos participantes após o uso do sistema de informações.

A primeira delas é o “Espanto”, mencionado em algumas citações, que se refere à situação do participante ao descobrir a quantidade de informação registrada em base oficial e os fatos divulgados em relação ao uso das verbas. A outra categoria construída é a “Estranheza” ao visualizar as atividades e os processos registrados sobre os candidatos nos tribunais.

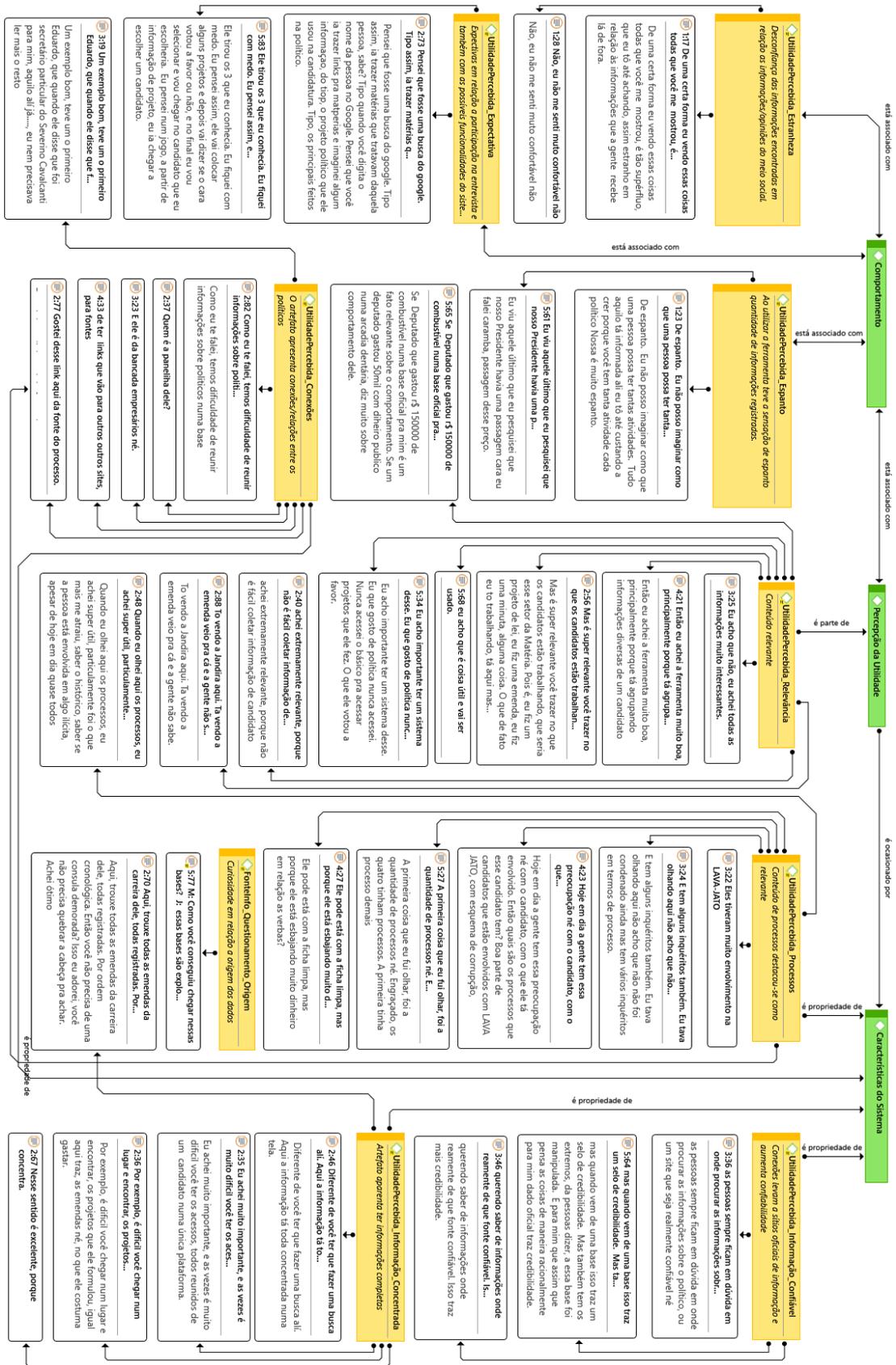


Figura 7.7: Rede sobre as características do sistema e percepção de utilidade. Fonte: do próprio autor,

*De uma certa forma eu vendo essas coisas todas, que você me mostrou, é tão supérfluo que eu tô até achando, assim, estranho em relação às informações que a gente recebe lá de fora (AL).*

A categoria “Expectativa” se refere ao fato de que os participantes se manifestaram ansiosos em 3 aspectos. Um em relação a participação na entrevista, pois eles imaginaram que as perguntas fossem em torno do(a)s candidato(a)s pesquisado(a)s. Além disso, era esperado que a ferramenta funcionasse como um jogo, onde os candidatos veriam resultados de seleção de projetos, de acordo com o seguinte relato:

*Eu fiquei com medo. Eu pensei assim, ele vai colocar alguns projetos e depois vai dizer se o cara votou a favor ou não, e no final eu vou selecionar e vou chegar no candidato que eu escolheria. Eu pensei num jogo, a partir de informação de projeto, eu ia chegar a escolher um candidato. (MA)*

E, por fim, o contato com o sistema gerou expectativa quanto a liberação, pois alguns participantes perguntaram quando a versão do sistema em produção seria liberada para uso e compartilhamento, o que indica uma certa expectativa de uso da tecnologia desenvolvida nesta pesquisa.

A rede também mostra a relevância do sistema a partir da emergência de categorias que refletem as funcionalidades de “Processos”. Para os participantes, essa foi considerada uma das principais características atrativas. Além disso, os links que direcionam o registro para as bases dos tribunais permitem observar o trâmite dos processos em tempo real. O acompanhamento de processo aumenta o grau de transparência sobre o desempenho do candidato. Por outro lado algumas questões sobre a necessidade de uma linguagem cidadã foi um tema que gerou muitas inquietações.

Em seguida, a categoria “Informação confiável” emergiu em virtude dos dados serem oriundos de bases oficiais, o que chamou a atenção dos participantes, que reagiram com confiabilidade após confirmada a origem dos dados. *As pessoas sempre ficam em dúvida em onde procurar as informações sobre o político, ou um site que seja realmente confiável (JS).* De acordo com os participantes, integrar a plataforma com dados oficiais traz um selo de confiança para o cidadão.

Uma característica que também emergiu é a “Informação Concentrada”. Foi relatado, pelos participantes desta pesquisa, que o sistema tem a capacidade de ser completo e apresenta informações de diferentes tipo. Mencionaram a dificuldade existente em localizar os tipos de dados em uma mesma plataforma. Geralmente as buscas são feitas isoladamente em vários sites de diferentes padrões e órgãos do governo, como ilustra o

relato de um dos entrevistados:

*Eu achei muito importante, e às vezes é muito difícil você ter os acessos, todos reunidos, de um candidato numa única plataforma. Por exemplo, é difícil você chegar num lugar e encontrar os projetos que ele formulou, igual aqui traz, as emendas, no que ele costuma gastar (BH).*

A Figura 7.8 mostra que o grau de satisfação em relação a facilidade de uso é elevado.

Os componentes de busca e os filtros utilizados foram considerados intuitivos pelos participantes, como ilustra esse depoimento: *“Então você não precisa ter um conhecimento prévio pra buscar a informação”*. A velocidade de retorno também foi considerada satisfatória bem como a performance dos links para outras páginas. *Acho que a pesquisa foi fácil, a pesquisa foi rápida (JS).*

O agrupamento dos fatos foi considerado uma forma de organização da informação. Porém alguns componentes web precisam ser mais intuitivos e é preciso acrescentar legendas a fim de tornar a difusão da informação mais aberta e democrática. *Não sei, talvez no início, pra ser uma ferramenta mais intuitiva, além dos ícones, deveria haver uma breve explicação pra que serve aquele ícone (GL).*

De acordo com os participantes, o mais importante na etapa de busca foi o conteúdo textual, as imagens dos políticos, os ícones presentes interpolados no texto, as cores de fundo de página e o formato dos componentes de agrupamento, não foram tão criticados. Talvez alguns recursos para tratamento de palavras possam ajudar na experiência do usuário, mas não é o foco deste trabalho.

Ao observar as gravações das interações, percebeu-se que os participantes não mostraram dificuldade em interagir, contudo o uso de legendas são necessárias para melhorar o entendimento dos *links* pra navegação. A expansão das conexões entre fatos ou desdobramentos dos fatos foram alguns requisitos que emergiram do uso do Sistema de Informação.

Na Figura 7.9, a rede mostra que, na sociedade contemporânea onde nós temos o conhecimento de muitos casos de corrupção divulgados na mídia e internet, seja no período eleitoral ou fora dele, a curiosidade é um fator que aumenta a busca por informações.

Se tais buscas forem incentivadas a ocorrerem em bases oficiais, essa prática pode reduzir a disseminação de conteúdos falsos e despertar uma curiosidade nas pessoas. Por consequência, o compartilhamento de informações oficiais poderia se tornar uma prática muito comum. Entendemos que, à medida que uma ferramenta de busca for divulgada de

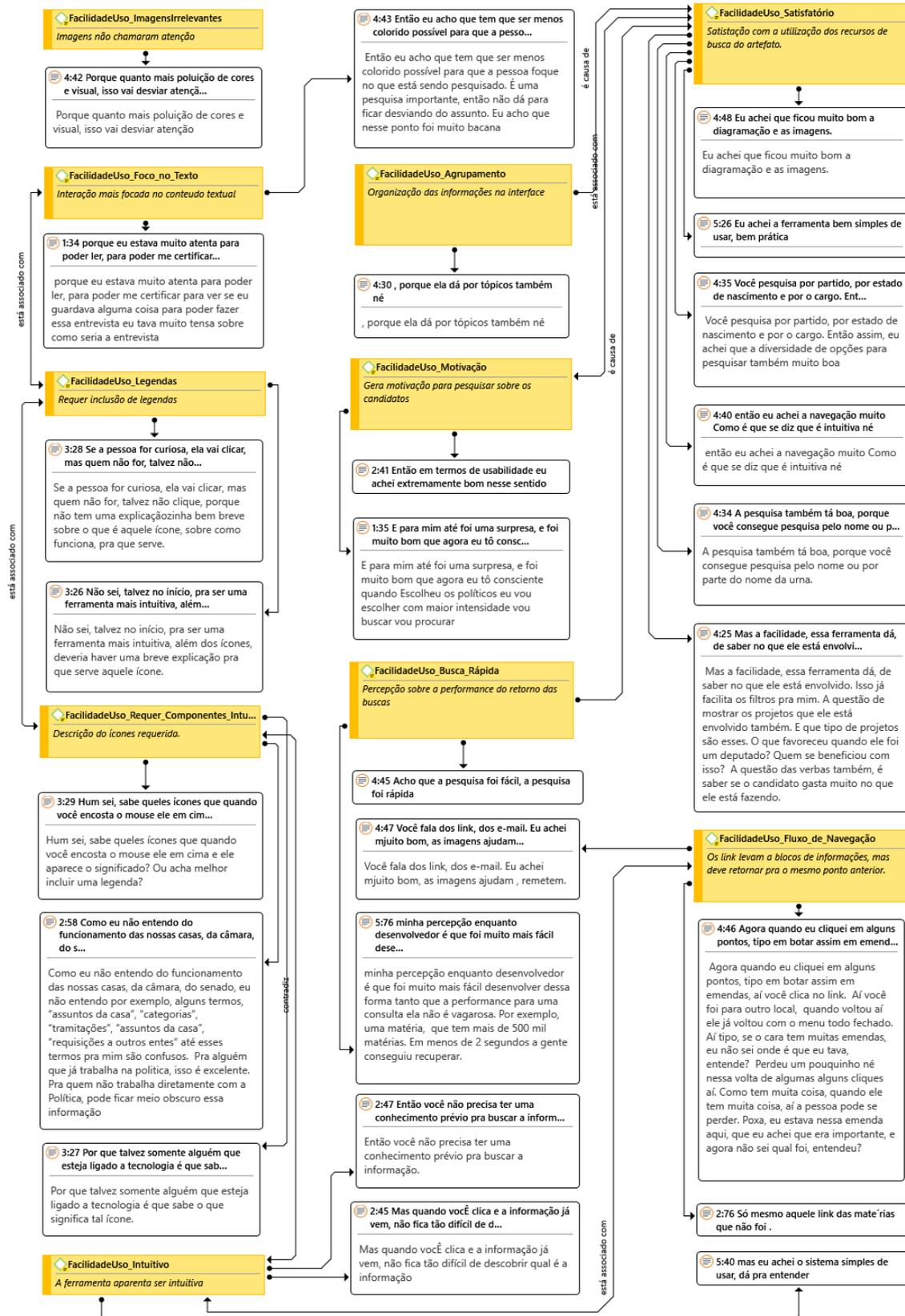


Figura 7.8: Percepções sobre a facilidade de uso do sistema. Fonte: do próprio autor.

modo inédito e compartilhada nas redes e tiver uma aceitação satisfatória, sua utilização terá mais potencial para ocorrer. Dessa forma as categorias “Segurança”, “Período Elei-

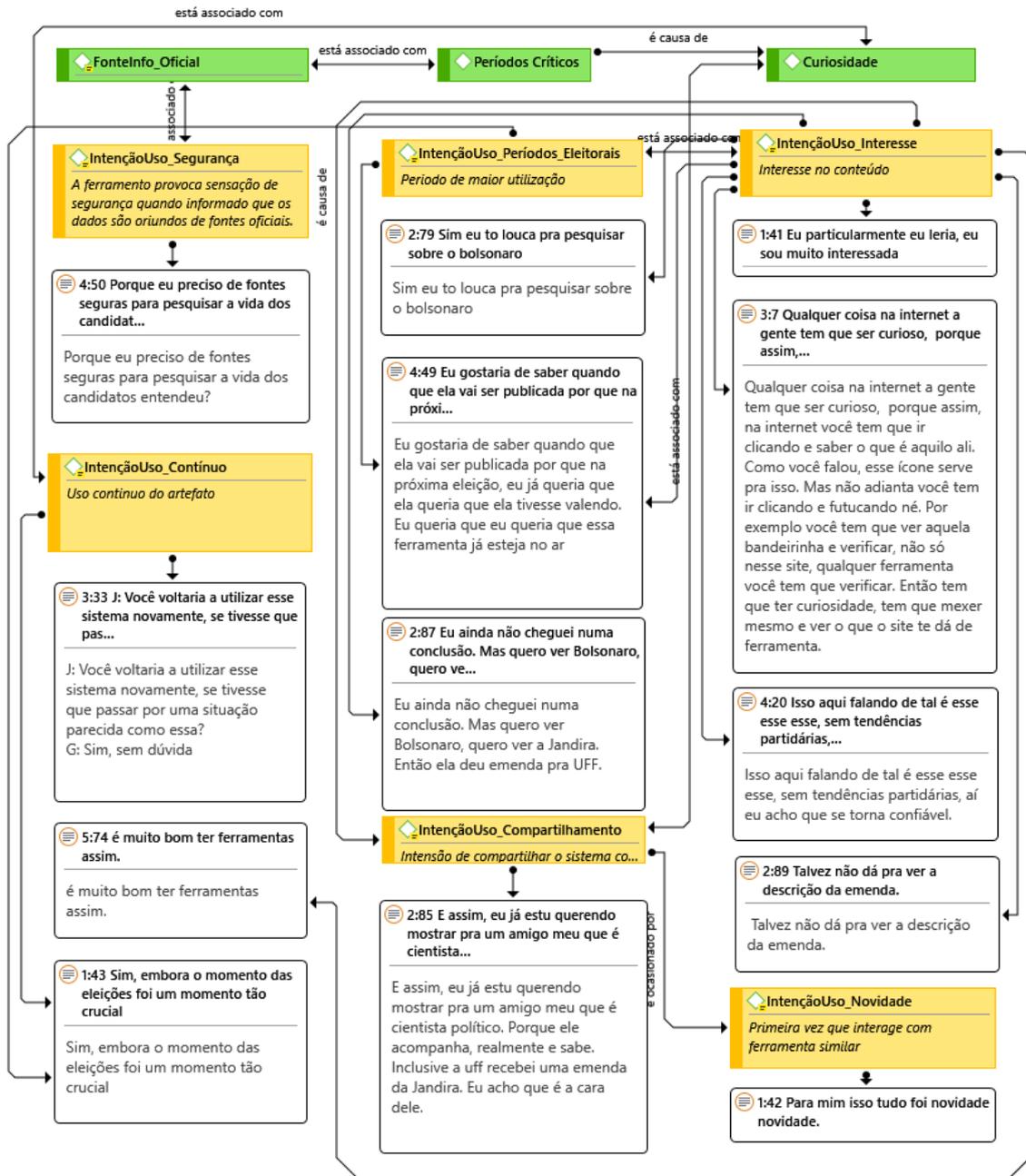


Figura 7.9: Indícios das intenções de uso do sistema. Fonte: do próprio autor,

toral”, “Novidade”, “Compartilhamento” e “Interesse” compõem a rede que mostra uma preocupação pelo uso de fontes seguras em pesquisas sobre a vida do político.

A rede representada na Figura 7.10 expõe as melhorias e mostra algumas contribuições do sistema para a sociedade, segundo relatos dos participantes.

Sobre as categorias que refletem melhorias no sistema, foi citado: “Texto longo” que, de acordo com o participante, torna a busca cansativa. Entendemos que a informação deve ser apresentada de maneira mais objetiva. Por outro lado, outros entrevistados mencionaram sobre textos muito simplificados, com pouca informação. Entendemos que deve

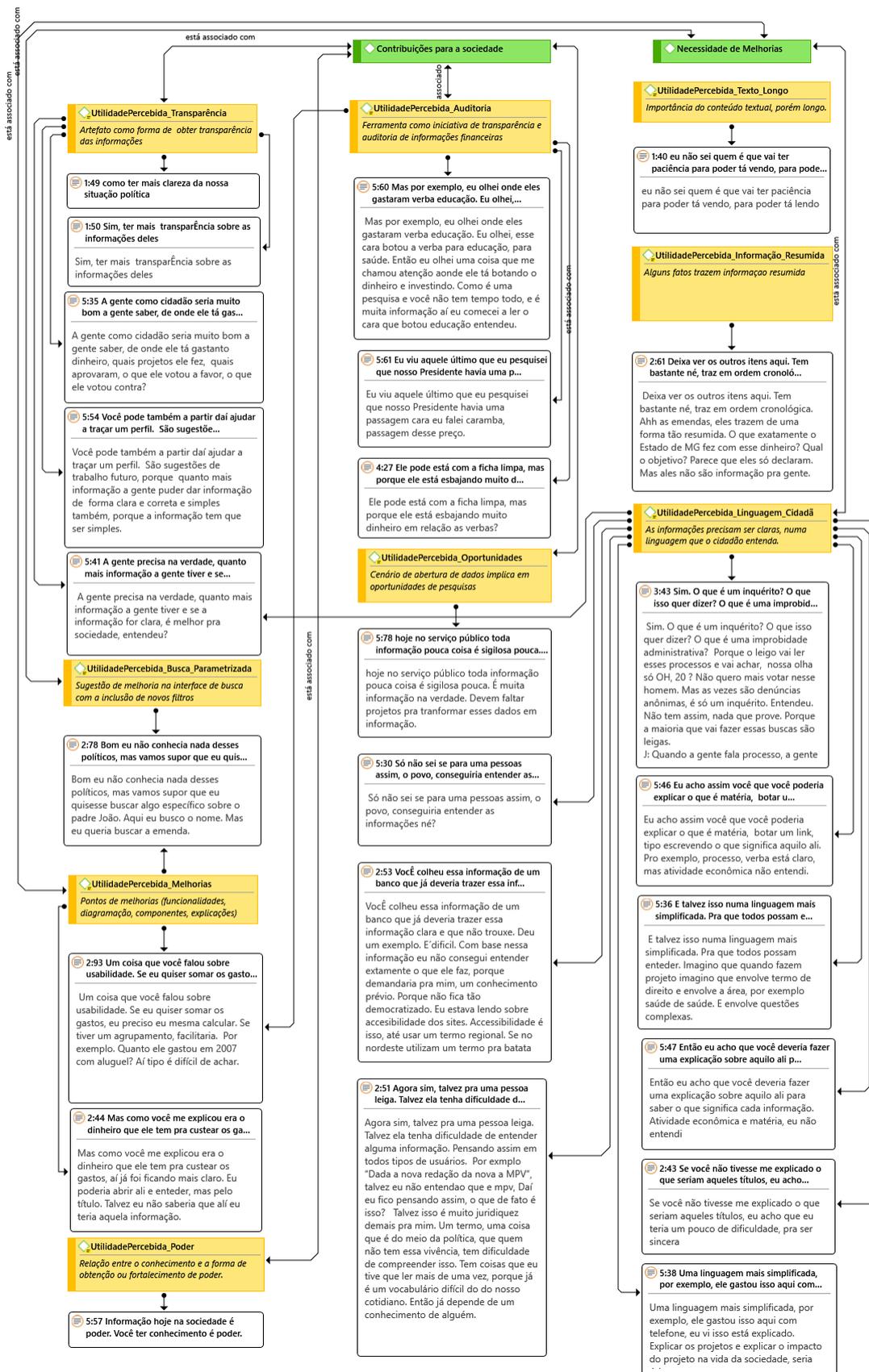


Figura 7.10: Contribuições para sociedade e sugestão de melhorias. Fonte: do próprio autor,

haver um equilíbrio e dependendo do cenário explorado, trazer o necessário para a interface do sistema.

Com o intuito de melhorar as formas de buscas, os participantes sugeriram a inclusão de novos parâmetros de buscas a partir do próprio perfil do candidato. Por isso a categoria “Busca Parametrizada” foi concebida. *Bom, eu não conhecia nada desses políticos, mas vamos supor que eu quisesse buscar algo específico sobre o padre João. Aqui eu busco o nome, mas eu queria buscar a emenda.* (BH).

Por fim, a categoria relacionada a “Melhorias” que mais se destacou durante as entrevistas, com 23 citações vinculadas, foi a “Linguagem Cidadã”. Os participantes enfatizaram que o governo deve cuidar das informações, dar significado a esses termos, pois ajudam o cidadão a melhor compreender a informação. Munir o sistema de glossários também pode reduzir as dúvidas. Os relatos mostram essa fragilidade e necessidade de melhoria.

*Você colheu essa informação de um banco que já deveria trazer essa informação clara e que não trouxe. ... com base nessa informação eu não consegui entender exatamente o que ele faz, porque demandaria pra mim, um conhecimento prévio. Porque não fica tão democratizado* (BH).

*E talvez isso numa linguagem mais simplificada. Pra que todos possam entender. Imagino que quando fazem projeto, imagino que envolve termo de direito e envolve a área, por exemplo, saúde* (MA).

Como contribuições, os participantes reconheceram que o sistema é uma ferramenta que fomenta a transparência de informações e reforçam a questão da confiança pelo fato dos dados serem oriundos de bases oficiais.

A categoria “Auditoria”, na opinião dos participantes, revela que o cidadão pode acompanhar, com mais rigor, os gastos por área de governo em que o candidato está aplicando os recursos e identificar desvios. *“Ele pode estar com a ficha limpa, mas porque ele está esbanjando muito dinheiro em relação às verbas?”* (JS). *“Eu olhei onde eles gastaram verba educação. Eu olhei, esse cara botou a verba para educação, para saúde. Então eu olhei uma coisa que me chamou atenção aonde ele tá botando o dinheiro e investindo”* (MA).

Em relação a categoria “Oportunidades”, os participantes identificaram espaço para novos projetos para tratar os dados públicos e transformá-los em informações úteis, transparentes e de interesse do cidadão.

*Hoje, no serviço público, pouca informação é sigilosa. É muita informação, na verdade.*

*Devem faltar projetos pra transformar esses dados em informação (MA).*

A “Transparência” foi uma das categorias com maior número de citações vinculadas. Os participantes mencionaram que o sistema poderá contribuir com a transparência de informações sobre a situação política dos candidatos, sobre como eles administram as verbas e as emendas. Acrescentaram também que a transparência de dados poderá ajudar a traçar o perfil do político, sendo esse conhecimento considerado importante para a sociedade.

A partir das entrevistas e da elaboração do *set* inicial de categorias, esta pesquisa se empenhou em manter o exercício de comparação com o intuito de obter um nível de coerência entre as categorias. Os resultados mostraram que a transparência obteve a maior incidência de percepções positivas. Desse modo, reconhecemos que a relevância do sistema está fortemente relacionada à transparência de informações oficiais conectadas, por isso a identificamos como *core category*, conforme os preceitos da *Grounded Theory* (TA-ROZZI, 2011).

Embora as percepções de transparência sejam marcantes, por outro lado, as categorias “EducaçãoPolítica\_Distanciamento” e “UtilidadePercebida\_liguagem\_Cidadã” também o são. Isso mostra que apesar da informação prosopográfica ser transparente, concentrada e de recuperação rápida provida pelo sistema de informações PROSOPOLÍTICOS, é necessário um formato ou padrão de linguagem que possibilite ao cidadão ter maior grau de entendimento das informações obtidas.

Por esse motivo, a categoria central passou a ser denominada de **Transparência Moderada de Informações Oficiais**

### **7.2.1 Rigor da Interpretação dos Resultados**

No esforço de realizarmos, de maneira sistemática, algumas das diretrizes da técnica de codificação da *Grounded Theory*, ressaltamos alguns procedimentos que mostram a credibilidade dos resultados dessa pesquisa.

Através dos procedimentos explanados, esta pesquisa está em conformidade com as orientações dadas por (CHARMAZ, 2006), quando a autora descreve os critérios para avaliação dos estudos em GT sob a ótica da credibilidade e originalidade.

Em relação à credibilidade, esta pesquisa alcançou familiaridade com o tópico em questão. Desde a aplicação da GT no estudo exploratório até a fase de análise dos dados, foi mantido envolvimento e observação intensos, além da relação com os participantes, pois todos fazem parte do meio social do pesquisador. A forma como as consultas foram rea-

lizadas, considerando entrevistas semiestruturadas e fazendo o uso de uma linguagem em sua forma natural, fizeram com que o conhecimento emergisse de maneira imprevisível, o suficiente para as reivindicações apontadas.

Considerando a profundidade contida nos dados, foi possível fazer interpretações dos dados nos momentos das suas transcrições, durante a codificação inicial e focalizada, entre os dados/citações com as categorias teóricas bem como no processo de formação das redes. Também aumenta a confiabilidade dos achados da pesquisa em relação a forma como foram realizados: o processo de coleta de dados, as transcrições, os memorandos, os comentários nos diferentes controles e instrumentos de registro, além da gravação em tempo real das interações e feições dos usuários (imagem e som).

Quanto à originalidade, as categorias ou códigos construídos oferecem novas ideias e provocações acerca dos aspectos comportamentais e de necessidade de políticas educacionais e de gestão de estado, à medida que a utilidade percebida e a intensão de uso emergiram durante o estudo empírico.

### **7.2.2 Discutindo os resultados obtidos em comparação com os trabalhos relacionados**

O presente trabalho também parte de estudos sobre temas semelhantes em campos de pesquisa como ontologia, prosopografia, arquivologia e história. A ontologia FACTOID, de acordo com (PASIN; BRADLEY, 2013), trabalha com informação prosopográfica e foi o ponto de partida para repensar e melhorar a definição de um ato de interpretação. A pesquisa trata de bases de dados prosopográficas estruturadas usando um modelo geral de estrutura “orientada para fatos” que conecta as pessoas às informações sobre elas através de pontos em fontes primárias que afirmam essa informação. Com o objetivo de disponibilizar os bancos de dados prosopográficos na Web Semântica, tal pesquisa revisa os princípios por trás de um modelo baseado em factóides e, reformula usando uma abordagem mais interoperável, baseada em princípios de representação de conhecimento e ontologias formais.

O trabalho de (HYVÖNEN et al., 2018), trata a publicação e utilização de biografias textuais na Web. O estudo de caso mostrou a manipulação de metadados com o SPARql através do projeto Link Data Finland (LDF). Neste projeto, as aplicações incluem pesquisa sobre histórias pessoais individuais, bem como pesquisas históricas de grupos de pessoas usando métodos de prosopografia. Os dados biográficos são enriquecidos pela extração de eventos de dados não estruturados e semiestruturados, e pela vinculação de entidades internas à fontes externas de dados. Um mecanismo de busca semântico é for-

necido para filtrar grupos de pessoas para pesquisa prosopográfica, mas não se preocupa com a questão da transparência e entendimento da informação ao devolver a informação.

Em (DAQUINO et al., 2017), o trabalho é voltado para construção de ontologias no contexto de gestão do patrimônio cultural e reforça os desafios de criar modelos que representem a vasta dimensão das obras existentes. O objetivo do trabalho é de sustentar uma plataforma de pesquisa de imagens. O objetivo do projeto é evitar que o mesmo objeto seja catalogado mais de uma vez, evitando assim informações contraditórias ou redundantes, pois as bases fazem parte de um consórcio de organizações que utilizam modelos conceituais distintos.

O Trabalho de (TUOMINEN; HYVÖNEN; LESKINEN, 2017) segue a mesma linha do anterior, pois propõe um modelo para harmonizar e interligar dados biográficos heterogêneos de diferentes fontes. O modelo faz distinção dos atores em relação aos papéis desempenhados e eventos ocorridos.

## 8. Conclusão

Foi uma longa jornada até aqui. Mais importante do que a entrega do artefato, foi toda a trajetória percorrida. Através dessa pesquisa, tive a oportunidade de me conectar a outros pesquisadores, outros grupos de estudo e compartilhar as experiências e os desafios da pesquisa. Desafios esses que, por meio do paradigma epistemológico-metodológico *Design Science Research*, foi possível projetar um sistema de informações como meio para construir conhecimento científico.

A pesquisa teórica demonstrou que, à luz da Teoria do Voto, o entendimento dos modelos cognitivos que explicam o comportamento do eleitor, foi relevante para fundamentar as conjecturas comportamentais que, por sua vez, direcionaram o desenvolvimento do artefato. O desenvolvimento dessa pesquisa resultou em um modelo multidimensional de dados prosopográficos sobre políticos brasileiros como elemento central utilizado por um Sistema de Informação. Esse modelo foi pensado considerando os fenômenos contemporâneos de propagação da informação e o comportamento do cidadão enquanto influenciado, influenciador e tomador de decisão.

Em relação aos objetivos específicos, pode-se afirmar que o Sistema de Informação foi concebido e contemplou o que foi proposto de acordo com cada item que segue:

- Fonte de dados existentes: a cooperação foi firmada e a identificação das bases, bem como a documentação, foi realizada e atualizada durante o projeto. O conjunto de dados é composto pelas bases e dados do Senado, Câmara dos deputados, Departamento de Ciência da Computação da UFS e Portal de Dados Abertos do TSE.
- Autenticidade das bases: A liberação das bases do Senado e da Câmara foi dada pela diretoria do Grupo Transparência Brasil. A Base da UFS foi liberada pelo grupo de pesquisa e APIS publicadas, e o repositório de dados abertos do TSE compartilha os dados em seu sítio oficial conforme regulamentado na Lei Acesso

a Informação(LAI, 2011) e a Política de Dados Abertos do Poder Executivo Federal(BRASIL, 2016)

- Pesquisar sobre o comportamento do eleitor e identificação dos fatores que influenciam o voto: para esse objetivo, foi realizado um estudo exploratório considerando a revisão sobre modelos cognitivos de Michigan (HIMMELWEIT; HUMPHREYS; JAEGER, 1985), Downs (DOWNS, 1957) e Holbrook (LODGE; MCGRAW; STROH, 1989), o que resultou na concepção do modelo FIV.Br2018, fruto de uma pesquisa inicial para compreensão do fenômeno investigado.
- Especificação do modelo conceitual de conhecimento: este modelo foi criado a partir da compreensão do modelo FIV.Br2018 e em conformidade com as bases de dados selecionadas, a fim de determinar os elementos a serem representados pelo modelo multidimensional.
- Composição do modelo dimensional: os requisitos de negócio foram definidos com base nos seguintes fatores oriundos do FIV.Br2018: experiência e comportamento do candidato, propostas de governo e identificação ideológica. A modelagem foi realizada e o *data warehouse* foi criado com os modelos materializados.
- Sobre a escolha das ferramentas de ETL: os *softwares* utilizados fazem parte da suíte Pentaho BI por se tratar de *software* livre e atender aos requisitos técnicos para extração, transformação e carga dos dados selecionados.
- Publicação dos modelos/cubos dimensionais: esse foi um dos objetivos mais desafiadores, pois as bases de origem possuíam formatos diferentes. As atividades de pré-processamento foram realizadas. Os processos de migração dos dados foram executados, juntamente com os *jobs* de integração.
- Criação da interface de consulta dos dados prosopográficos: um sistema de informação foi construído com o objetivo de possibilitar ao cidadão o acesso a informações organizadas sobre a biografia do político, sob os aspectos da transparência de informações oficiais e conectadas. Oficiais porque vem de bases do governo, e conectadas porque o modelo multidimensional permite o desdobramento das informações a medida que a interação acontece. O sistema foi desenvolvido, testado e liberado em ambiente de pesquisa no Data Center da UFF. A versão atual está disponível em [www.prosopoliticos.uff.br](http://www.prosopoliticos.uff.br).

Em resposta a questão de pesquisa: **como possibilitar a melhoria do entendimento do cidadão e trazer informações relevantes sobre a vida dos políticos, através de um modelo multidimensional de dados prosopográficos?**

Buscamos responder essa questão por meio da avaliação do uso do sistema desenvolvido nesta pesquisa. Investigamos, junto ao cidadão, a utilidade percebida, a facilidade de uso e intenção de uso do sistema. Para a avaliação dos resultados, foi utilizada a técnica de codificação da *Grounded Theory* Construtivista (CHARMAZ, 2006).

Pela interpretação dos depoimentos que os usuários deram sobre o sistema experienciado, interpretamos que eles reconhecem que o sistema é uma potencial fonte de informação confiável e fomenta a busca pelo conhecimento sobre política e políticos. Além disso, concentra informações sobre vários fatos. Expõe o desempenho político através dos processos, matérias publicadas, projetos criados, verbas utilizadas, e permite verificar as conexões/redes de pessoas a partir dos fatos e das informações profissionais. Por outro lado, merece atenção em relação ao significado da informação, pois as informações precisam ser mais claras ao olhos do cidadão.

## 8.1 Principais Contribuições Científicas e Tecnológicas

Cumprindo os requisitos do processo do conhecimento disciplinar, a metodologia DSR descreve a comunicação como atividade a ser desempenhada pelo pesquisador ao final do desenvolvimento do artefato (PEFFERS et al., 2007). Nesse sentido, este trabalho tem como contribuição científica:

- Barroso Júnior, J.S., Pimentel, M., Nunes, V., and Cappelli, C. (2018). Comportamento Eleitoral: quais fatores influenciam a decisão de voto em tempos de cibercultura? In ABCIBER XI-Simpósio Nacional da Associação Brasileira de Cibercultura
- Barroso Júnior, J. S., Pimentel, M., Nunes, V., and Cappelli, C. (2019). Design Science Research to design a conceptual model about prosopographic information related to politicians. In Proceedings of the XV Brazilian Symposium on Information Systems (p. 24). ACM.
- Barroso Júnior, J.S., Diirr, B., Nunes, V., and Cappelli, C. (2019). Where to go next in Brazilian research and practice on transparency through software? In 6th Global Conference on Transparency Research.

Esta última contribuição, apesar do problema de pesquisa ser diferente, grande parte da revisão bibliográfica foi realizada durante esse trabalho a partir das bases Scopus, ScienceDirect, ACM, El Compedex, IEEE, Springer e WebOfScience.

Ademais, o artefato projetado contribui para um processo mental de identificação de informações relevantes sobre os políticos. Isso evita a criação de estruturas no banco de dados a revelia, que no primeiro momento poderia ser confiável, mas irrelevantes perante o FIV.Br2018.

Com relação à relevância e ao impacto da pesquisa na sociedade, esperamos que o resultado possa ser utilizado em iniciativas de pesquisa/investigação ou processos eleitorais, tanto na esfera parlamentar quanto em outras esferas eleitorais.

Para a Universidade, nosso objetivo é promover a pesquisa de serviço público, no ambiente político e democrático. Para os Programas de Pós-Graduação, fomentar o desenvolvimento de pesquisas em sistemas de informação nas áreas de participação cidadã, linguagem clara, democracia digital, cibercultura, transparência organizacional e informação.

## **8.2 Limitações da Pesquisa**

Algumas limitações técnicas foram identificadas nessa pesquisa ao recuperar os dados nas plataformas abertas do TSE. Observamos algumas mudanças nos padrões ou estruturas das fontes em função da temporalidade das atualizações no repositório aberto. Desse modo, para que o processo de recuperação seja automatizado, é necessário que haja um mapeamento e a tratativas de todos os tipos de estruturas, para as candidaturas de 1945 até 2016.

Uma dificuldade encontrada foi a nomenclatura das tabelas, mantidas pelo grupo Transparência Brasil, pois não condizia com o conteúdo das mesmas. Isso demandou mais tempo para entendimento de toda a estrutura e do negócio, pois não foi possível obter o suporte exclusivo de um analista de dados ou especialista do domínio. O formato com que os dados eram armazenados não seguiam as estruturas formais de normalização para bases de dados relacionais e isso tornou o processo de transformação mais custoso.

Sobre a avaliação da pesquisa, pelos resultados coletados, percebe-se que alguns participantes têm maior conhecimento sobre política e democracia. Verificou-se que para essas pessoas, as redes construídas a partir dos códigos, demonstraram uma percepção mais positiva em relação a utilidade percebida. Desse modo a pesquisa poderia envolver um número maior de pessoas que não possuem um conhecimento amplo sobre as ciências políticas, a fim de identificarmos mais categorias/códigos, e associações entre eles, pois (CHARMAZ, 2006) reconhece que podemos iniciar nossos estudos a partir dessas pers-

pectivas privilegiadas, mas precisamos permanecer o mais aberto possível a tudo o que vemos e sentimos nas etapas iniciais da pesquisa.

### 8.3 Trabalhos Futuros e Oportunidades

Como trabalhos futuros, almejamos avaliar o artefato através da metodologia do *Focus Group* (Grupo Focal). De acordo com (BARBOUR; KITZINGER, 1998) são grupos de discussão que dialogam sobre um tema em particular ao receberem estímulos apropriados para o debate. Essa técnica distingue-se por suas características próprias, principalmente pelo processo de interação grupal, que é uma resultante da procura de dados. Já temos uma proposta de grupo focal na Universidade Federal Fluminense, que inicialmente possuirá o seguinte perfil demográfico selecionado: alunos de graduação, de um mesmo curso ou mesma área de conhecimento, faixa etária de 20-35 anos, residentes no estado do Rio de Janeiro e que tenham acesso a internet por celular.

Um dos fatos que mereceu atenção foi o posicionamento Político dos Candidatos. Identificamos uma quantidade ainda pequena desses registros na base de dados da UFS. Em (PASSOS et al., 2019), o autor justifica que não foi possível localizar uma estrutura de dados abertos que trate desses elementos e a própria natureza do discurso ambíguo dos políticos brasileiros dificulta o trabalho de algoritmos de aprendizagem conseguirem tirar conclusões precisas. Desse modo é um desafio e fica anotado como escopo de possíveis projetos de pesquisa, o desenvolvimento de um método automatizado para se obter o posicionamento de candidato(a), individualmente, em jornais, revistas, sessões técnicas, debates e demais audiência, redes sociais, programas de televisão ou outras fontes de dados não estruturados.

Outra recomendação para modelo multidimensional é a introdução de novas tabelas, a fim de realizar análises de dados mais ricas e posterior publicação/conversão de cubos na linguagem MDX, que permitem ser usados em ferramentas de BI e gerar *dash boards* ou painéis de dados em formatos gráficos.

Em relação a integração e engajamento com o Governo, sugerimos que, para eventuais atualizações, este modelo faça uso das APIs das Plataformas de dados Abertos da Câmara, Senado e TSE, e que os métodos de extração sejam refatorados para serem mais automatizados. Outrossim, levantar juntos ao Protocolo Integrado do Governo Federal a possibilidade de integração aos processos abertos dos políticos.

Em contrapartida, o sistema poderia disponibilizar APIs de acordo com os padrões abertos

de interoperabilidade (BRASIL, 2015), para que outros sistemas também possam usar os dados. Pode-se ainda verificar como os dados sensíveis podem ser tratados de acordo com a regulamentação da Lei Geral de Proteção de Dados Pessoais (LGPD) (BRASIL, 2018)

Finalmente, esperamos que esse modelo multidimensional evolua e possa contribuir no contexto da compreensão e organização de informações prosopográficas e provoque soluções de apoio para combater a ignorância, promover a transparência e a confiabilidade desse tipo de informação.

## Referências Bibliográficas

ABEL, M.; FIORINI, S. R. Uma revisão da engenharia do conhecimento: Evolução, paradigmas e aplicações. *International Journal of Knowledge Engineering and Management (IJKEM)*, v. 2, n. 2, p. 1–35, 2013.

ABELLÓ, A.; SAMOS, J.; SALTOR, F. Yam2: a multidimensional conceptual model extending uml. *Information Systems*, Elsevier, v. 31, n. 6, p. 541–567, 2006.

ABRAMO, C. W.; MONDO, B. V.; ALPENDRE, G.; ISSA, M.; VIANNA, R.; ANGÉLICO, F. *Como são nossos parlamentares*. 2015. Disponível em: <https://www.transparencia.org.br/docs/excelencias.pdf>, Acessado: 12 Nov. 2017. Disponível em: <<https://www.transparencia.org.br/docs/excelencias.pdf>>. Acesso em: 12 Nov. 2017.

AKEN, J. E. v. Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules. *Journal of management studies*, Wiley Online Library, v. 41, n. 2, p. 219–246, 2004.

ALAN, R. H. V.; MARCH, S. T.; PARK, J.; RAM, S. Design science in information systems research. *MIS quarterly*, Springer, v. 28, n. 1, p. 75–105, 2004.

ALLCOTT, H.; GENTZKOW, M. Social media and fake news in the 2016 election. *Journal of economic perspectives*, v. 31, n. 2, p. 211–36, 2017.

ANGÉLICO, F. Estudo analítico sobre transparência e legitimidade das organizações da sociedade civil brasileira. *ABONG–Associação Brasileira de Organizações Governamentais*, 2011.

ARAÚJO, P. M. Recrutamento parlamentar para o senado e o perfil dos senadores brasileiros, 1989-2006. ”, Universidade Federal de Pernambuco, 2011.

ASSOCIATION, C. *Comprehensive Knowledge Archive Network*. 2015. Disponível em: <https://ckan.org>. Acessado: 10 Mai. 2018. Disponível em: <<https://ckan.org/>>. Acesso em: 10 Mai. 2018.

ASSOCIATION, C. *Dados Abertos da Câmara dos Deputados*. 2018. 2018. Disponível em: <https://dadosabertos.camara.leg.br>. Acessado: 10 Mai. 2018. Disponível em: <<https://dadosabertos.camara.leg.br/swagger/api.html>>. Acesso em: 10 Mai. 2018.

- BALLARD, C.; FARRELL, D. M.; GUPTA, A.; MAZUELA, C.; VOHNIK, S. et al. *Dimensional Modeling: In a Business Intelligence Environment*. [S.l.]: IBM Redbooks, 2012.
- BARBOUR, R.; KITZINGER, J. *Developing focus group research: politics, theory and practice*. [S.l.]: Sage, 1998.
- BAX, M. P. Design science: filosofia da pesquisa em ciência da informação e tecnologia. 2017.
- BLACKMORE, S. *The meme machine*. [S.l.]: Oxford Paperbacks, 2000. v. 25.
- BOEHM, B. A view of 20th and 21st century software engineering. In: ACM. *Proceedings of the 28th international conference on Software engineering*. [S.l.], 2006. p. 12–29.
- BORBA, F. O impacto da propaganda negativa na decisão do voto. *Encontro da Associação Brasileira de Ciência Política ABCP*, 2012.
- BRAGA, S.; NICOLÁS, M. A. Prosopografia a partir da web: avaliando e mensurando as fontes para o estudo das elites parlamentares brasileiras na internet. *Revista de Sociologia e Política*, Universidade Federal do Paraná, v. 16, n. 30, 2008.
- BRASIL. *Guia de Interoperabilidade do Governo*. 2015. Disponível em: <http://eping.governoeletronico.gov.br>, Acessado: 12 Set. 2019. Disponível em: <http://eping.governoeletronico.gov.br>. Acesso em: 12 Set. 2019.
- BRASIL. *Lei Geral de Proteção de Acesso à Dados*. 2018. Lei nº 13.709, de 14 de agosto de 2018, dispõe sobre a proteção de dados pessoais. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/L13709.htm](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm). Acesso em: 10 Nov. 2018.
- BRASIL, G. do. *Lei Anticorrupção*. 2013. Lei no 12.846, de 01 de agosto de 2013: Dispõe sobre a responsabilização administrativa e civil de pessoas jurídicas pela prática de atos contra a administração pública, nacional ou estrangeira, e dá outras providências. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2013/lei/112846.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2013/lei/112846.htm). Acesso em: 25 mai. 2018.
- BRASIL, G. do. *Lei do Conflito de Interesses no exercício de cargo ou emprego do Poder Executivo Federal*. 2013. Lei no 12.813, de 16 de maio de 2013: Dispõe sobre o conflito de interesses no exercício de cargo ou emprego do Poder Executivo federal e impedimentos posteriores ao exercício do cargo ou emprego; e revoga dispositivos da Lei nº 9.986, de 18 de julho de 2000, e das Medidas Provisórias nºs 2.216-37, de 31 de agosto de 2001, e 2.225-45, de 4 de setembro de 2001. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2013/lei/112813.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2013/lei/112813.htm). Acesso em: 02 jun. 2018.
- BRASIL, G. do. *Política de Dados Abertos do Poder Executivo Federal*. 2016. Política de Dados Abertos do Poder Executivo Federal. Decreto n. 8.777, de 11 de maio de 2016. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2016/decreto/D8777.htm](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/D8777.htm). Acesso em: 30 mai. 2018.

- BRASIL, G. do. *Portal Brasileiro de Dados Abertos*. 2017. Portal Brasileiro de Dados Abertos, Resolução nº 02, de 24/3/2017, anexo1. Disponível em: <<http://dados.gov.br/>>. Acesso em: 20 mai. 2018.
- BRICKLEY, D.; MILLER, L. *FOAF vocabulary specification 0.91*. [S.l.]: Citeseer, 2007.
- BUNGE, M. *Treatise on basic philosophy: Ontology II: A world of systems*. [S.l.]: Springer Science & Business Media, 2012. v. 4.
- BUTLER, D.; STOKES, D. *Political change in britain*. new york: St. Martin's, 1969.
- CAMPBELL, A.; CONVERSE, P. E.; MILLER, W. E.; STOKES, D. E. *The american voter*. [S.l.]: University of Chicago Press, 1980.
- CAPPELLI, C. *Uma abordagem para transparência em processos organizacionais utilizando aspectos*. Tese (Doutorado) — Pontifícia Universidade Católica do Rio de Janeiro - PUC-Rio, 2009.
- CAPPELLI, C.; LEITE, J. *Transparência de processos organizacionais. II Simpósio Internacional de Transparência nos Negócios, Universidade Federal Fluminense, LATEC, Niterói, RJ, Brasil*, 2008.
- CASTRO, L. *Metodologia para Modelagem Conceitual de Dados*. V Simpósio Brasileiro de Sistemas de Informação, 2009. Disponível em: <http://www.lbd.dcc.ufmg.br/colecoes/sbsi/2009/0025.pdf>, Acessado: 03 Jan. 2019. Disponível em: <<http://www.lbd.dcc.ufmg.br/colecoes/sbsi/2009/0025.pdf>>. Acesso em: 03 Jan. 2019.
- CHARMAZ, K. *Constructing grounded theory: A practical guide through qualitative analysis*. [S.l.]: sage, 2006.
- CHARMAZ, K. *A construção da teoria fundamentada: guia prático para análise qualitativa*. [S.l.]: Bookman Editora, 2009.
- CHAUDHURI, S.; DAYAL, U. An overview of data warehousing and olap technology. *ACM Sigmod record*, ACM, v. 26, n. 1, p. 65–74, 1997.
- CHUNG, L.; LEITE, J. C. S. do P. On non-functional requirements in software engineering. In: *Conceptual modeling: Foundations and applications*. [S.l.]: Springer, 2009. p. 363–379.
- CONVERSE, P. E. The nature of belief systems in mass publics (1964). *Critical review*, Taylor & Francis, v. 18, n. 1-3, p. 1–74, 2006.
- CORBIN, J.; STRAUSS, A. *Basics of qualitative research: Techniques and procedures for developing grounded theory*. [S.l.]: Sage publications, 2014.
- COSTA, D. R. Componentes determinantes na tomada de decisão na escolha do candidato: Um estudo empírico nas eleições majoritárias. *Projetos, dissertações e teses do Programa de Doutorado e Mestrado em Administração*, v. 6, n. 1, 2016.
- COSTA, L. D.; CODATO, A. Profissionalização ou popularização da classe política brasileira? um perfil dos senadores da república. *Os eleitos: representação e carreiras políticas em democracias*. Porto Alegre: Editora da UFRGS, p. 107–134, 2013.

CRESWELL, J. W.; POTH, C. N. *Qualitative inquiry and research design: Choosing among five approaches*. [S.l.]: Sage publications, 2016.

DAQUINO, M.; MAMBELLI, F.; PERONI, S.; TOMASI, F.; VITALI, F. Enhancing semantic expressivity in the cultural heritage domain: exposing the zero photo archive as linked open data. *Journal on Computing and Cultural Heritage (JOCCH)*, ACM, v. 10, n. 4, p. 21, 2017.

DAVIS, F. D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, JSTOR, p. 319–340, 1989.

DIAS, M.; MATTOS, M.; ROMÃO, W.; TODESCO, J.; PACHECO, R. Data warehouse—presente e futuro. *Revista Tecnológica*, v. 7, p. 59–73, 1998.

DJANG, S. F. *Django: The web framework for perfectionists with deadlines*. 2005. Disponível em: <https://docs.djangoproject.com/en/2.2/>. Acessado: 9 Sep. 2019. Disponível em: [<https://docs.djangoproject.com/en/2.2/>](https://docs.djangoproject.com/en/2.2/). Acesso em: 9 Sep. 2019.

DOCUMENTATION, D. *System check framework*. 2019. Disponível em: <https://docs.djangoproject.com/en/2.2/topics/checks/>. Acessado: 9 Sep. 2019. Disponível em: [<https://docs.djangoproject.com/en/2.2/topics/checks/>](https://docs.djangoproject.com/en/2.2/topics/checks/). Acesso em: 9 Sep. 2019.

DOWNS, A. An economic theory of political action in a democracy. *Journal of political economy*, The University of Chicago Press, v. 65, n. 2, p. 135–150, 1957.

DOWNS, A. An economic theory of political action in a democracy. *The Political Economy*, p. 12–26, 1984.

DRESCH, A. Design science e design science research como artefatos metodológicos para engenharia de produção. Universidade do Vale do Rio dos Sinos, 2013.

ECONOMIA, M. da. *VLibras - Tradutor Libras em Software Livre*. 2018. Disponível em: <https://www.gov.br/governodigital/pt-br/transformacao-digital/ferramentas/vlibras>, Acessado: 10 Mai. 2018. Disponível em: [<https://www.gov.br/governodigital/pt-br/transformacao-digital/ferramentas/vlibras>](https://www.gov.br/governodigital/pt-br/transformacao-digital/ferramentas/vlibras). Acesso em: 10 Mai. 2018.

EISENBERG, J. Internet, democracy, and the republic. *Dados*, SciELO Brasil, v. 46, n. 3, p. 491–511, 2003.

ELEITORAL, T. S. *Repositório de Dados Eleitorais*. 2018. Repositório de Dados Eleitoral. Disponível em: <http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais-1/repositorio-de-dados-eleitorais>. Acesso em: 10 mai. 2018.

ENELOW, J. M.; HINICH, M. J. *The spatial theory of voting: an introduction*. [S.l.]: Cambridge University Press, 1984.

ENGIEL, P.; ARAÚJO, R.; CAPPELLI, C. 0004/2012-catálogo de características de entendimento de modelos de processo de prestação de serviços públicos. *RelaTe-DIA*, v. 6, n. 1, 2012.

FEDERAL, B. R. *Classificação Nacional de Atividades Econômicas*. 2019. Disponível em: <http://receita.economia.gov.br/orientacao/tributaria/cadastrados/cadastro-nacional-de-pessoas-juridicas-cnpj/classificacao-nacional-de-atividades-economicas-2013-cnae/apresentacao>. Acessado: 12 Mai. 2019. Disponível em: <<http://receita.economia.gov.br/orientacao/tributaria/cadastrados/cadastro-nacional-de-pessoas-juridicas-cnpj/classificacao-nacional-de-atividades-economicas-2013-cnae/apresentacao>>. Acesso em: 17 Jan. 2019.

FERREIRA, A.; CALDEIRA, C. P.; OLIVAL, F. Dados qualitativos, prosopografia e análise de redes. Instituto de Ciências Sociais-Universidade de Lisboa, 2012.

FONSECA, M. d. J. Leonardo da vinci: um gênio universal. *Millenium*, Instituto Politécnico de Viseu, 1997.

FUNG, A.; GRAHAM, M.; WEIL, D. *Full disclosure: The perils and promise of transparency*. [S.l.]: Cambridge University Press, 2007.

GLASER, B. G.; STRAUSS, A. L. *Discovery of grounded theory: Strategies for qualitative research*. [S.l.]: Routledge, 2017.

GONCALVES, S. *Information and accountability: Evidence from Brazil*. Tese (Doutorado) — London School of Economics and Political Science (United Kingdom), 2010.

GROUP, O. G. W. et al. The annotated 8 principles of open government data. *Available at (Jan. 13, 2017): <http://opengovdata.org>*, 2007.

GUIZZARDI, G. On ontology, ontologies, conceptualizations, modeling languages, and (meta) models. *Frontiers in artificial intelligence and applications*, IOS Press, v. 155, p. 18, 2007.

HEVNER, A.; MARCH, S.; PARK, J.; RAM, S. *Design science in information systems research*. *MIS Q* 28 (1): 75–105. 2004.

HIMMELWEIT, H. T.; HUMPHREYS, P.; JAEGER, M. *How Voters Decide: A Model of Vote Choice Based on a Special Longitudinal Study Extending over Fifteen Years and the British Election Surveys of 1970–1983*. [S.l.]: Open University Press, Milton Keynes, 1985.

HOLZNER, B.; HOLZNER, L. *Transparency in global change: The vanguard of the open society*. *Pittsburgh*. [S.l.]: PA: University of Pittsburgh Press, 2006.

HUFF, A.; TRANFIELD, D.; AKEN, J. E. van. Management as a design science mindful of art and surprise a conversation between anne huff, david tranfield, and joan ernst van aken. *Journal of Management Inquiry*, Sage, v. 15, n. 4, p. 413–424, 2006.

HYVÖNEN, E.; LESKINEN, P.; TAMPER, M.; TUOMINEN, J.; KERAVUORI, K. Semantic national biography of finland. In: *Proceedings of the Digital Humanities in the Nordic Countries 3rd Conference (DHN 2018), Helsinki, Finland*. [S.l.: s.n.], 2018.

IBGE, F. I. B. d. G. e. E. *Estatísticas de Gênero. Indicadores sociais das mulheres no Brasil*. 2018. Disponível em: <http://bit.ly/2ODhLmP>, Acessado: 20 Set. 2018. Disponível em: <<https://biblioteca.ibge.gov.br/>>. Acesso em: 20 Set. 2018.

- IBM, K. C. *Modelos Dimensionais*. 2018. Disponível em: <https://ibm.co/35XrOLo>, Acessado: 01 fev. 2019. Disponível em: <<https://ibm.co/35XrOLo>>. Acesso em: 01 fev. 2019.
- INMON, W. H. *Building the data warehouse*. [S.l.]: John Wiley & sons, 2005.
- JARKE, M.; LENZERINI, M.; VASSILIOU, Y.; VASSILIADIS, P. *Fundamentals of data warehouses*. [S.l.]: Springer Science & Business Media, 2013.
- JÄRVINEN, P. Action research is similar to design science. *Quality & Quantity*, Springer, v. 41, n. 1, p. 37–54, 2007.
- JR, J. F. N.; CHEN, M.; PURDIN, T. D. Systems development in information systems research. *Journal of management information systems*, Taylor & Francis, v. 7, n. 3, p. 89–106, 1990.
- JUNIOR, J. P. M. Modelo sociológico de decisão de voto presidencial no Brasil 1994-2006. *Revista Debates*, Universidade Federal do Rio Grande do Sul, IFCH-PPG em Ciência Política, v. 3, n. 2, p. 68, 2009.
- KIMBALL, R.; ROSS, M. *The data warehouse toolkit: the complete guide to dimensional modeling*. [S.l.]: John Wiley & Sons, 2011.
- LAGOZE, C.; HUNTER, J. The abc ontology and model. In: *International Conference on Dublin Core and Metadata Applications*. [S.l.: s.n.], 2001. p. 160–176.
- LAI. *Lei de Acesso à Informação*. 2011. Lei no 12.527, de 18 de novembro de 2011: Regula o acesso a informações previsto no Inciso XXXIII do Art. 5º, no Inciso II do § 3º do Art. 37 e no § 2º do Art. 216 da Constituição Federal; altera a Lei no 8112, de 11 de dezembro de 1990; revoga a Lei no 11.111, de 5 de maio de 2005, e dispositivos da Lei no 8.159, de 8 de janeiro de 1991; e dá outras providências. Diário Oficial da União, Brasília, 18 de novembro de 2011. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/lei/112527.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm)>. Acesso em: 10 jan. 2018.
- LASSWELL, H. D.; LERNER, D. *World revolutionary elites: Studies in coercive ideological movements*. [S.l.]: MIT Press, 1965. v. 60.
- LAUDON, K. C.; LAUDON, J. P. *Management information system*. [S.l.]: Pearson Education India, 2016.
- LEITE, F. Raciocínio e procedimentos da grounded theory construtivista. *Questões Transversais*, v. 3, n. 6, 2016.
- LIMEIRA, T. M. V.; MAIA, T. Comunicação política e decisão de voto: o que as pesquisas revelam. *Ponto-e-Vírgula: Revista de Ciências Sociais*, v. 1, n. 8, p. 42–55, 2010.
- LODGE, M.; MCGRAW, K. M.; STROH, P. An impression-driven model of candidate evaluation. *American Political Science Review*, Cambridge University Press, v. 83, n. 2, p. 399–419, 1989.
- LOPES, A. B. P. *Aplicação de técnicas de business intelligence a base de dados prospectivas*. Dissertação (Mestrado) — Universidade de Évora, 2017.

LUZ, M. G. d. *Pragmatismo Eleitoral e Sistemas de Crenças Ideológicas*. Seminário Nacional de Ciência Política: Democracia em Debate, 2008. Disponível em: <http://www.ufrgs.br/sncp/edicoes-antiores/2008/arquivos>, Acessado: 25 Out. 2018. Disponível em: <http://www.ufrgs.br/sncp/edicoes-antiores/2008/arquivos>. Acesso em: 25 Out. 2018.

MANSANO, S. R. V. Sujeito, subjetividade e modos de subjetivação na contemporaneidade. *Revista de Psicologia da UNESP*, v. 8, n. 2, 2018.

MARCH, S. T.; SMITH, G. F. Design and natural science research on information technology. *Decision support systems*, Elsevier, v. 15, n. 4, p. 251–266, 1995.

MATOS, D. *Por que Cientistas de Dados escolhem Python?* 2017. Disponível em: <http://www.cienciaedados.com/por-que-cientistas-de-dados-escolhem-python/>. Acesso em: 15 mai. 2018.

MATOS, D. *Organograma do Poder Judiciário*. jfal. Disponível em: <http://www.jfal.jus.br/institucional/organograma>. Acesso em: 15 Jan. 2018.

MCCABE, M. C.; LEE, J.; CHOWDHURY, A.; GROSSMAN, D.; FRIEDER, O. On the design and evaluation of a multi-dimensional approach to information retrieval. In: ACM. *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*. [S.l.], 2000. p. 363–365.

MIGUEL, L. F.; MOTA, F. F. *Mídia, eleições e pesquisa de opinião no Brasil (1989-2010): um mapeamento da presença das pesquisas na cobertura eleitoral*. Associação Brasileira de Pesquisadores em Comunicação e Política, 2011.

MOODY, D. L.; KORTINK, M. A. From enterprise models to dimensional models: a methodology for data warehouse and data mart design. In: *DMDW*. [S.l.: s.n.], 2000. p. 5.

OGP. *Open Government Partnership. Open Government Declaration*. 2011. Disponível em: <https://www.opengovpartnership.org/open-government-declaration>, Acessado: 25 Marc. 2018. Disponível em: <https://www.opengovpartnership.org/open-government-declaration>. Acesso em: 25 Marc. 2018.

ORACLE. *Oracle Business Intelligence Discoverer Plus User's Guide*. Oracle Documentation Library, 2018. Accessed: December 15, 2018. Disponível em: [https://docs.oracle.com/cd/B14098\\_01/bi.1012/b13915/i\\_olap\\_chapter005.htm](https://docs.oracle.com/cd/B14098_01/bi.1012/b13915/i_olap_chapter005.htm). Acesso em: 10 Dec. 2018.

PARISER, E. *O filtro invisível: o que a internet está escondendo de você*. [S.l.]: Zahar, 2012.

PASIN, M.; BRADLEY, J. Factoid-based prosopography and computer ontologies: towards an integrated approach. *Digital Scholarship in the Humanities*, Oxford University Press, v. 30, n. 1, p. 86–97, 2013.

PASSOS, N. R.; RODRIGUES, A. F.; MACEDO, H. T.; PRADO, B. O.; SILVA, G. J. da; MATOS, L. N. Open data extraction, transformation, and loading as a tool for supporting 2018 elections' voters. In: ACM. *Proceedings of the XV Brazilian Symposium on Information Systems*. [S.l.], 2019. p. 23.

- PEFFERS, K.; TUUNANEN, T.; ROTHENBERGER, M. A.; CHATTERJEE, S. A design science research methodology for information systems research. *Journal of management information systems*, Taylor & Francis, v. 24, n. 3, p. 45–77, 2007.
- PERFEITO, J. Mercado político: componentes determinantes na tomada de decisão do eleitor. *Revista de Negócios*, v. 8, n. 1, 2007.
- PERISSINOTO, R.; COSTA, L. D.; TRIBESS, C. Origem social dos parlamentares paranaenses (1995-2006): alguns achados e algumas questões de pesquisa. *Sociologias*, SciELO Brasil, v. 11, n. 22, 2009.
- PIMENTEL, M. *DSR-Model: modelo de pesquisa em Design Science Research*. 2019. Disponível em: <https://www.slideshare.net/pimentelmario/dsr-model>, Acessado: 25 Out. 2019. Disponível em: <https://www.slideshare.net/pimentelmario/dsr-model>. Acesso em: 25 Out. 2019.
- PIMENTEL, M.; FILIPPE, D.; SANTORO, F. M. Metodologia de pesquisa científica em informática na educação: Concepção de pesquisa. In: \_\_\_\_\_. [s.n.], 2018. v. 1, cap. 5, p. 1–29. Disponível em: <http://metodologia.ceie-br.org/livro-1/>>>.
- POE, V.; BROBST, S.; KLAUER, P. *Building a data warehouse for decision support*. [S.l.]: Prentice-Hall, Inc., 1997.
- POLSBY, N. W. The institutionalization of the us house of representatives. *American political science review*, Cambridge University Press, v. 62, n. 1, p. 144–168, 1968.
- PRZEYBILOVICZ, E. A e-participação e a e-transparência na execução do orçamento público brasileiro para educação: uma investigação sob o paradigma de design science. In: *Congreso Iberoamericano de Ciencia, Tecnología, Innovación y Educación*. [S.l.: s.n.], 2014. p. 1–23.
- PYTHON, S. F. *Python 3.7.6 documentation*. 2001. Disponível em: <https://docs.python.org/3.7/>, Acessado: 10 Jun. 2019. Disponível em: <https://docs.python.org/3.7/>>. Acesso em: 10 Jun. 2019.
- RAINARDI, V. *Building a data warehouse: with examples in SQL Server*. [S.l.]: John Wiley & Sons, 2008.
- RECLAME, A. *Sobre o Reclame Aqui*. 2018. Disponível em: <https://www.reclameaqui.com.br/institucional>. Acessado: 12 Jul. 2018. Disponível em: <https://www.reclameaqui.com.br/institucional/>>. Acesso em: 12 Jul. 2018.
- RECLAMEAQUI, I. *O plugin que vai mudar a política brasileira*. 2018. Disponível em: <http://www.vigieaqui.com.br/>. Acessado: 12 Jul. 2018. Disponível em: <http://www.vigieaqui.com.br/>>. Acesso em: 12 Jul. 2018.
- REIS, E.; TEIXEIRA, F.; ARAÚJO, M. Implementando uma solução de business intelligence com o microsoft sql server 2005–parte 1. *SQL Magazine, Rio de Janeiro*, n. 59, p. 52–66, 2009.
- RENNÓ, L. A estrutura das crenças de massa e seu impacto na decisão de voto. *BIB–Revista Brasileira de Informação Bibliográfica em Ciências Sociais*, v. 51, p. 85–103, 2001.

- RIBEIRO, V. B. As possibilidades e desafios do relato biográfico: biografia & história: hagiografia, trajetórias e prosopografia. *Protestantismo em Revista*, v. 44, n. 1, p. 140–152, 2018.
- ROLNIK, S.; GUATTARI, F. Micropolítica. *Cartografias del deseo*, 2006.
- ROMME, A. G. L. Making a difference: Organization as design. *Organization science, INFORMS*, v. 14, n. 5, p. 558–573, 2003.
- RUMBAUGH, J.; JACOBSON, I.; BOOCH, G. *Unified modeling language reference manual, the*. [S.l.]: Pearson Higher Education, 2004.
- SALVADOR, V. F. M.; BRITTO, M.; JUNIOR, L. M.; JUNIOR, J. A. Qualidade de dados para gestão de conhecimento na área de saúde. In: *Anais do X Congresso Brasileiro de Informática em Saúde. Florianópolis, Brasil*. [S.l.: s.n.], 2006.
- SANTAELLA, L. *A Pós-Verdade é Verdadeira ou Falsa?* [S.l.]: Editora estação das letras e cores, 2018.
- SERRANO, M.; LEITE, J. C. S. do P. Capturing transparency-related requirements patterns through argumentation. In: IEEE. *2011 First International Workshop On Requirements Patterns*. [S.l.], 2011. p. 32–41.
- SHAW, R.; TRONCY, R.; HARDMAN, L. Lode: Linking open descriptions of events. In: SPRINGER. *Asian semantic web conference*. [S.l.], 2009. p. 153–167.
- SIMON, H. A. *The sciences of the artificial*. [S.l.]: MIT press, 1996.
- SIRINELLI, J.-F. Os intelectuais. *Por uma história política*, Editora UFRJ Rio de Janeiro, v. 2, p. 231–269, 1996.
- SOUSA, H. P. de S.; LEAL, A. L. de C.; LEITE, J. C. S. do P. Alinhamento de operacionalizações entre transparência e mps. br. *iSys-Revista Brasileira de Sistemas de Informação*, v. 8, n. 4, p. 109–141, 2016.
- SOX. *Weekly Compilation of Presidential Documents. Presidential remarks and Statement*. 2002. Disponível em: <http://www.gpo.gov>, Acessado: 14 Marc. 2018. Disponível em: <<http://www.gpo.gov>>. Acesso em: 14 Marc. 2018.
- STACKOVERFLOW. *Newest MySql Questions*. 2018. Disponível em: <https://stackoverflow.com/questions/tagged/mysql>, Acessado: 10 Marc. 2018. Disponível em: <<https://stackoverflow.com/questions/tagged/mysql>>. Acesso em: 10 Marc. 2018.
- STF. *S. T. Sistema Judiciário Brasileiro: organização e competências*. 2018. Disponível em: <https://stf.jusbrasil.com.br/noticias/2535347/sistema-judiciario-brasileiro-organizacao-e-competencias>, Acessado: 20 Fev. 2018. Disponível em: <<https://stf.jusbrasil.com.br/noticias/2535347/sistema-judiciario-brasileiro-organizacao-e-competencias>>. Acesso em: 20 Fev. 2018.
- STONE, L. Prosopography. *Daedalus*, JSTOR, p. 46–79, 1971.

STONE, L. Prosopography. *Revista de Sociologia e Política*, SciELO Brasil, v. 19, n. 39, p. 115–137, 2011.

TAKEDA, H.; VEERKAMP, P.; YOSHIKAWA, H. Modeling design process. *AI magazine*, v. 11, n. 4, p. 37–37, 1990.

TAROZZI, M. O que é a grounded theory. *Metodologia de pesquisa e de teoria fundamentada nos dados*. Petrópolis–RJ: Vozes, 2011.

TAVARES, E. L. H. Avaliação em larga escala e qualidade da educação: um estudo a partir da visão dos sujeitos da rede escolar municipal de cachoeirinha/rs. Universidade do Vale do Rio dos Sinos, 2013.

THOMSEN, E. *OLAP solutions: building multidimensional information systems*. [S.l.]: John Wiley & Sons, 2002.

TRANSPARENCIA, B. *Estatuto da Associação Transparência Brasil*. 2015. Disponível em: <https://www.transparencia.org.br/quem-somos>. Acessado: 12 Nov. 2017. Disponível em: <https://www.transparencia.org.br/quem-somos>. Acesso em: 12 Nov. 2017.

TRIBUNAL, d. C. d. U. *5 motivos para a abertura de dados na Administração Pública*. 2015. Disponível em: <https://portal.tcu.gov.br/biblioteca-digital/cinco-motivos-para-a-abertura-de-dados-na-administracao-publica.htm>. Acessado: 06 jun. 2018. Disponível em: <https://portal.tcu.gov.br/biblioteca-digital/cinco-motivos-para-a-abertura-de-dados-na-administracao-publica.htm>. Acesso em: 06 Jun. 2018.

TUOMINEN, J.; HYVÖNEN, E.; LESKINEN, P. Bio crm: A data model for representing biographical data for prosopographical research. *Biographical Data in a Digital World (BD2017)(November 2017)*, <https://doi.org/10.5281/zenodo>, v. 1040712, 2017.

ULLMANN, S. *Semantics: an introduction to the science of meaning*. Barnes & Noble, 1979.

UNIÃO, C. G. da. *Programa Brasil Transparente*. 2013. Portaria nº 277, de 7 de Fevereiro de 2013: Institui o Programa Brasil Transparente que tem por objetivo geral apoiar Estados e Municípios na implementação da Lei de Acesso à Informação, Lei nº 12.527, de 18 de novembro de 2011, no incremento da transparência pública e na adoção de medidas de governo aberto. Disponível em: [http://www.cgu.gov.br/sobre/legislacao/arquivos/portarias/portaria\\_cgu\\_277\\_2013-1.pdf](http://www.cgu.gov.br/sobre/legislacao/arquivos/portarias/portaria_cgu_277_2013-1.pdf). Acesso em: 16 jun. 2018.

UNIÃO, C. G. da. *Escala Brasil Transparente*. 2017. A EBT foi desenvolvida para fornecer os subsídios necessários à Controladoria-Geral da União (CGU) para o exercício das competências que lhe atribuem os artigos 59 da Lei Complementar nº 101/2000 e 41 (I) da Lei de Acesso à Informação, assim como os artigos 68 (II) do Decreto nº 7.724/2012 e 18 (III), do Decreto nº 8.910/2016. Disponível em: <http://www.cgu.gov.br/assuntos/transparencia-publica/escala-brasil-transparente>. Acesso em: 16 jun. 2018.

URQUHART, C.; LEHMANN, H.; MYERS, M. D. Putting the ‘theory’ back into grounded theory: guidelines for grounded theory studies in information systems. *Information systems journal*, Wiley Online Library, v. 20, n. 4, p. 357–381, 2010.

VAISHNAVI, V.; KUECHLER, W. Design research in information systems. 2004.

VIANNA, M. *Design thinking: inovação em negócios*. [S.l.]: Design Thinking, 2012.

ZIULKOSKI, L. Coleta de requisitos e modelagem de dados para data warehouse: um estudo de caso utilizando técnicas de aquisição de conhecimento. *Bacharelado em Ciência da Computação pelo Instituto de Informática da Universidade Federal do Rio Grande do Sul*, 2003.



## **A. Estudo Exploratório - Questionário Fase 2**

# Fase 2 - Comportamento Eleitoral: Fatores que influenciam a decisão de voto.

Esta pesquisa faz parte de um estudo exploratório que visa a identificar quais são os fatores que mais influenciam os(as) eleitores(as) na escolha de um(a) candidato(a). Para isso, basta indicar quais dos fatores até agora identificados mais influenciam a sua escolha, quais influenciam um pouco, e quais não influenciam a sua escolha. E se tivermos esquecido de identificar algum(ns) fator(es) que você considera importante, por favor, nos indique no final da listagem.

Sua participação é voluntária, e ao aceitar participar, o suas respostas serão analisadas e transcritas em documentos (dissertação, artigos, relatórios técnicos etc.) que serão publicados em literatura científica especializada. Os dados serão coletados sem a necessidade de identificação, garantindo assim o sigilo e a privacidade dos participantes.

**\*Obrigatório**

## 1- Gênero \*

- Prefiro não declarar
- Homem
- Mulher
- Não binário
- Outro:

## 2 - Idade \*

Sua resposta



### 3- Orientação sexual \*

- Prefiro não declarar
- Heterossexual
- Homossexual
- Bissexual
- Transsesual
- Intersexual
- Outro:

### 4- Concordância entre o gênero e o sexo biológico \*

- Prefiro não dizer
- Cisgênero
- Transgênero (homem-trans ou mulher-trans)
- Outro:

### 5- Cor/Etnia \*

Escolher ▼

### 6- Estado \*

Escolher ▼

### 7- Localização \*



Escolher ▼

**8- Grau de Escolaridade \***

Escolher ▼

**9- Profissão \***

Sua resposta

**10- Renda Mensal (Bruta) \***

Escolher ▼

**11- Religião \***

- Prefiro não informar
- Católica
- Evangélica
- Espírita
- Umbanda
- Sem religião
- Outro:



No processo para escolher um(a) candidato(a), dentre os fatores listados abaixo, em que nível ou aproximação, eles influenciam a sua decisão de voto? \*

	Não influencia	Influencia pouco	Influencia muito	Não quero responder
A expressividade do candidato	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Acesso às informações dos candidatos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Acompanhamento dos debates	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Acredita na mudança e em novas formas de governar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Alta rejeição a um candidato que se apresenta na frente das pesquisas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Alta rejeição a um candidato que se apresenta na frente das pesquisas, escolhendo um candidato com potencial enfrentamento ao rejeitado	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Amadurecimento em relação ao movimento político	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aproximação dos valores do serviço público	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Atributos ou qualidades pessoais que se aproximam dos interesses	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Atuação dentro e fora do período eleitoral	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Capacidade de articular mudanças	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



anterior. Indique

Competência comportamental	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Comportamento político	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Confiança no candidato	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Conjuntura atual	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Conjuntura atual e expectativas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Conjuntura econômica	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Conjuntura econômica e política	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Conjuntura política	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Consciência Política	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Decepção com atual governo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Discurso proferido	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Engajamento com a política	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Entendimento e interesse pelo prestação de serviços públicos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Envolvimento em casos de escândalos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Envolvimento em processos administrativos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Envolvimento em processos de corrupção	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Estilo de vida do candidato	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Estratégia de voto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Exercer o direito de votar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Experiencia politica	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Experiências de vida	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Exposição sobre os candidatos na mídia e internet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fatos, ocorrências, resultados de pesquisas: cenário das eleições	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ficha Limpa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fontes de Informações - Postagens nas Redes Sociais (Whatsapp, Twitter, Facebook, Instagram) e as oriundas do perfil do candidato	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fontes de InformaçõesTV e Radio	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fontes de Informações - Reportagem publicadas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fontes de Informações - Resultados de pesquisas publicados	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fontes de Informações - Resultados de pesquisas publicados na internet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fontes de Informações - Resultados de pesquisas publicados no site do TSE	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Grau de instrução do Candidato	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Histórico da atuação do partido	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Histórico da atuação política e avaliação positiva em algumas áreas de atuação do	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



-----  
candidato

Histórico da  
atuação política

Histórico de  
atuação política em  
relação a  
elaboração e  
aprovação de  
projetos

Histórico dos  
Candidatos

Histórico e  
experiência do  
candidato

Histórico pessoal

Ideias contrarias ao  
candidato mais  
bem votado nas  
pesquisas de  
intenção de voto

Identificação com  
as ideias e  
propostas do  
candidato  
relacionadas ao  
combate à  
corrupção

Identificação com  
as ideias e  
propostas do  
candidato  
relacionadas às  
relações  
comerciais

Identificação com  
as ideias e  
propostas  
inovadoras

Identificação com  
as ideias e  
propostas para  
áreas consideradas  
fundamentais na  
formação do  
cidadão

Identificação com  
as ideias e valores

Identificação com  
as ideias e valores  
do partido

Identificação com  
as ideias e valores  
em áreas



específicas de atuação

Identificação com as ideias e valores em defesa da reformulações da

política de segurança pública de encarceramento e genocídio dos grupos socialmente vulneráveis

Identificação com as ideias e valores em defesa das minorias e grupos socialmente vulneráveis

Identificação com as ideias e valores relacionadas a controladoria e punição no combate a corrupção

Identificação com as ideias e valores relacionadas a educação e ciência

Identificação com as ideias e valores relacionadas a segurança

Identificação com as ideias e valores relacionados aos movimentos políticos

Identificação com as ideias em defesa ao meio ambiente

Identificação com as ideias em defesa ao patrimônio público

Identificação com as ideias em defesa das questões trabalhistas

Identificação de ideia e valores em áreas específicas como educação e cultura

identificação ideológica



Ideologia	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Improbidade Administrativa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
identificação com ideias e valores do partido	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Influência de conflitos políticos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Influência política na família	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Influência Regional	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inicia sua análise por eliminação	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inovação	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Insatisfação com a atuação do governo atual	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Insatisfação com a realidade da educação, com os desafios de ser professor e suas condições de trabalho.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Insatisfação com o a situação atual	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Insatisfação com o atual governo que tenta reeleição	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Insatisfação em relação as políticas atuais da educação.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Insegurança	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interesse e idéias contrárias aos candidatos mais bem colocados, influenciou na decisão pelo terceiro mais vem votados nas pesquisa, visando sua aprovação para o segundo turno	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interesse em				



assuntos sobre distribuição de renda	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interesse em propostas inovadoras	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interesse em propostas que fortaleçam o Estado	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interesse em propostas que tratam a educação, como prioridade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interesse em propostas que tratam a saúde, como prioridade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interesses em defesa da sociedade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Investigações	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Linhas de atuação de governo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Maturidade intelectual	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Maturidade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Não aprovação do governo atual que tenta reeleição	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Não concordância com as articulações políticas dos candidatos para permanência no poder	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Necessidade da população	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Neutralidade ideológica partidária	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O perfil do candidato que se mostra como defensor dos interesses da sociedade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Opta por candidatos que não	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



estão atuando no atual governo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Opta por mudanças, e exclui candidatos que já estão no poder	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
poder ou declaram apoio direto				
Parcerias políticas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Participação da sociedade nas decisões dentro e fora do período eleitoral	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Partidos e candidatos novos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pensamento liberal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Perfil do comportamental do candidato	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Política voltada para a maioria	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Política voltada para minorias e grupos socialmente vulneráveis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Posicionamento do candidato frente a questões políticas relevantes para você	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Posicionamento do candidato frente aos eventos considerados anti-democráticos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Posicionamento do partido ou candidato frente a questões ligadas aos direitos fundamentais.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Posicionamento político	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Preferência de voto ou opinião de amigos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Preferência de voto ou opinião de colegas de grupos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



específicos como, cursos, escolas, academias	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Preferência de voto ou opinião de colegas de trabalho.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Preferência de voto ou opinião de parentes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Programa de governo proposto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Programa de governo proposto especificamente para as áreas de saúde, educação	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Programa de governo proposto, especificamente para a população da região onde reside	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Propostas de governo a favor do estatuto do desarmamento	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Propostas de governos relacionadas redução da maioridade penal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Rejeição	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Segurança para tomada de decisão	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Situação atual	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sociedade patriarcal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tentativa de diminuir as probabilidade do candidato não desejado e que está a frente das pesquisas ser eleito	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tentativa de diminuir as probabilidade do candidato não desejado e que está tentando reeleger	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Tentativa de  
diminuir as  
probabilidade do  
candidato não  
desejado ser eleito

Trajatória Política

Vida Pessoal

Vota-se por ordem  
de outrem

Voto estratégico

Voto parental

Voto por  
eliminação.

Voto protesto

**Outros fatores:**

Sua resposta

ENVIAR

Nunca envie senhas pelo Formulários Google.

Este formulário foi criado em Centro de Ciências Exatas e Tecnologia da UNIRIO. [Denunciar abuso](#)

Google Formulários



## **B. Estudo Exploratório - Questionário Fase 3**

# Fase 3 - Comportamento eleitoral: quais fatores influenciam a sua decisão de voto?

Esta pesquisa está sendo realizada pelo Programa de Pós Graduação de Informática (PPGI) da UNIRIO através do mestrando José Barroso Júnior, sob a orientação do Profº Mariano Pimentel, Dra Vanessa Nunes e Dra. Claudia Cappelli.

Como parte de um estudo exploratório, esta pesquisa visa identificar quais são os fatores que mais influenciam os(as) eleitores(as) na escolha de um(a) candidato(a). Para isso, basta indicar quais dos fatores até agora identificados influenciam decisivamente, quais influenciam muito, quais influenciam um pouco, e quais não influenciam a sua escolha. E se tivermos esquecido de identificar algum(ns) fator(es) que você considera importante, por favor, nos indique no final da listagem.

Sua participação é voluntária, e ao aceitar participar, suas respostas serão analisadas e transcritas em documentos (dissertação, artigos, relatórios técnicos etc.) que serão publicados em literatura científica especializada. Os dados serão coletados sem a necessidade de identificação, garantindo assim o sigilo e a privacidade dos participantes. Caso haja alguma dúvida em relação as questões, entre em contato através do e-mail [jose.junior@uniriotec.br](mailto:jose.junior@uniriotec.br).

## 1- Gênero

- Prefiro não declarar
- Homem
- Mulher
- Não binário
- Outro:

## 2 - Idade

Sua resposta



### 3- Orientação sexual

- Prefiro não declarar
- Heterossexual
- Homossexual
- Bissexual
- Transsexual
- Interssexual
- Outro:

### 4- Concordância entre o gênero e o sexo biológico

- Prefiro não dizer
- Cisgênero
- Transgênero (homem-trans ou mulher-trans)
- Outro:

### 5- Cor/Etnia

Escolher ▼

### 6- Estado

Escolher ▼

### 7- Localização



Escolher

## 8- Grau de Escolaridade

Escolher

## 9- Profissão

Sua resposta

## 10- Renda Mensal (Bruta)

Escolher

## 11- Religião

- Prefiro não informar
- Catolico
- Evangélico
- Espírita
- Umbanda
- Sem religião
- Outro:

No processo para escolher um(a) candidato(a), quais os fatores que mais influenciam a sua decisão de voto? Informe o grau de influência que cada um dos fatores a seguir tem na sua decisão de voto.



12- Os interesses pelas ideias e valores do candidato ou partido, exercem qual nível influência em sua decisão de voto? Estes fatores estão relacionados à Identificação Ideológica

- Não influencia
- Influencia pouco
- Influencia muito
- Influencia totalmente (Decisivo)
- Não quero responder

13- Você considera que as opiniões de terceiros podem intervir em sua decisão de voto em que grau de influência? Estes fatores estão relacionados à Influência Social.

- Não influencia
- Influencia pouco
- Influencia muito
- Influencia totalmente (Decisivo)
- Não quero responder



14- Com que grau de influência os sentimentos que você tem pelo(a) candidato(a) ou pelo partido podem intervir em sua decisão de voto?

- Não influencia
- Influencia pouco
- Influencia muito
- Influencia totalmente (Decisivo)
- Não quero responder

15- Em que grau, os resultados das pesquisas de intenção de voto publicados, influenciam em sua decisão de voto?

- Não influencia
- Influencia pouco
- Influencia muito
- Influencia totalmente (Decisivo)
- Não quero responder



16- Qual nível de influência você considera que as fontes de informações sobre os candidatos, como a mídia (rádio e TV) e internet através das redes sociais (Facebook, Twitter, Instagram, Whatsapp), exercem em sua decisão de voto? Estes fatores estão relacionados a Fontes de Informação.

- Não influencia
- Influencia pouco
- Influencia muito
- Influencia totalmente (Decisivo)
- Não quero responder

17- Você considera que o perfil comportamental do(a) candidato(a), a expressividade, os discursos proferidos, o envolvimento em processos de corrupção, comportamento político e suas atitudes com suas parcerias podem interferir na sua escolha em que nível de influência? Estes fatores estão relacionados ao Comportamento do Candidato.

- Não influencia
- Influencia pouco
- Influencia muito
- Influencia totalmente (Decisivo)
- Não quero responder



18- Qual nível de influência, a experiência política do candidato(a), a experiência de vida, o grau de instrução, a atuação dentro e fora do período eleitoral e o histórico de atuação no partido exercem sobre sua intenção de voto? Este fatores estão relacionados à Experiência do Candidato.

- Não influencia
- Influencia pouco
- Influencia muito
- Influencia totalmente (Decisivo)
- Não quero responder

19- Você considera que exercendo seu direito de voto, fazendo suas análises, seja ela por eliminação, estratégia de voto ou voto protesto, essas ações influenciam em que nível, a sua decisão de voto? Esses fatores estão relacionados à Participação do Eleitor

- Não influencia
- Influencia pouco
- Influencia muito
- Influencia totalmente (Decisivo)
- Não quero responder



20- As propostas de governo apresentadas pelos(as) candidatos(as) aos eleitores exercem que grau de influência em sua intenção de voto?

- Não influencia
- Influencia pouco
- Influencia muito
- Influencia totalmente (Decisivo)
- Não quero responder

21- Qual nível de influência as reflexões em relação ao seu posicionamento no espectro político, econômico e social, se direita ou esquerda, podem exercer na escolha do seu candidato? Esses fatores estão relacionados ao Posicionamento Político do Eleitor.

- Não Influencia
- Influencia pouco
- Influencia muito
- Influencia totalmente (Decisivo)
- Não quero responder

22- Informe aqui outros fatores:

Sua resposta



Nunca envie senhas pelo Formulários Google.

Este formulário foi criado em Centro de Ciências Exatas e Tecnologia da UNIRIO. [Denunciar abuso](#)

Google Formulários



## **C. Dicionário de Dados do Esquema Transparência Brasil**

# Schema Report for database: Lista de Tabelas - Transparência Brasil

Table: agregacoes						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
endereco	VARCHAR(70)	Yes	No	No		
legenda	VARCHAR(70)	Yes	No	No		
prioridade	TINYINT(3)	Yes	No	No	'0'	

Table: anos						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
ano	YEAR(4)	Yes	No	No		

Table: antigo_comissoes_alemg						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	No		
comissoes	TEXT	No	No	No	NULL	

Table: asclaras						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	Yes	No		
outroId	MEDIUMINT(9)	Yes	No	No	'0'	
partidoId	TINYINT(3)	Yes	No	No		
estadoId	TINYINT(3)	Yes	No	No		
municipioId	SMALLINT(5)	No	No	No	NULL	
deferimentoId	TINYINT(4)	No	No	No	NULL	
resultadoId	TINYINT(10)	Yes	No	No	'1'	
cargoId	TINYINT(3)	Yes	No	No		
cargoPai	TINYINT(3)	Yes	No	No	'0'	
eleito	TINYINT(3)	Yes	No	No	'0'	

bens	DECIMAL(15,0)	Yes	No	No	'0'	
recursosCand	DECIMAL(10,2)	Yes	No	No	'0.00'	
recursosCom	DECIMAL(11,2)	Yes	No	No	'0.00'	
recursosPF	DECIMAL(10,0)	Yes	No	No	'0'	
recursosComPF	DECIMAL(10,2)	Yes	No	No	'0.00'	
autoadoacoes	DECIMAL(10,2)	Yes	No	No	'0.00'	
alodoacoes	DECIMAL(10,2)	Yes	No	No	'0.00'	
doacoesCom	DECIMAL(10,2)	Yes	No	No	'0.00'	
desconto	DECIMAL(11,2)	Yes	No	No	'0.00'	
desconto1	DECIMAL(11,2)	Yes	No	No	'0.00'	
numero	MEDIUMINT(9)	Yes	No	No		
votos	INT(11)	No	No	No	NULL	
turno	TINYINT(3)	Yes	No	No	'1'	
nome	VARCHAR(255)	Yes	No	No		
apelido	VARCHAR(255)	No	No	No	NULL	
cgc	VARCHAR(20)	No	No	No	NULL	
sequencial	INT(11)	Yes	No	No	'0'	
matriculaTSE	BIGINT(20)	No	No	No	'0'	
tituloEleitoral	VARCHAR(15)	No	No	No	NULL	
ocupacao	INT(10)	Yes	No	No	'0'	
ano	YEAR(4)	Yes	No	No		

**Table: calendario**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
data	DATE	No	No	No	NULL	

**Table: cargos**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	TINYINT(3)	Yes	Yes	No		
descricao	VARCHAR(100)	Yes	No	No		

**Table: casas**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	TINYINT(4)	Yes	Yes	No		
nome	VARCHAR(150)	No	No	No	NULL	

abrev	VARCHAR(2)	No	No	No	NULL	
uf	TINYINT(4)	No	No	No	NULL	
municipio	SMALLINT(6)	Yes	No	No		

**Table: casas\_atributos**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	TINYINT(4)	Yes	Yes	Yes		
quantidade	SMALLINT(6)	Yes	No	No	'0'	
dados_abertos	TINYINT(1)	Yes	No	No	'0'	
pres_plenario	TINYINT(1)	Yes	No	No	'0'	
pres_comissoes	TINYINT(1)	Yes	No	No	'0'	
indenizatoria	TINYINT(1)	Yes	No	No	'0'	
viagens	TINYINT(1)	Yes	No	No	'0'	
diarias	TINYINT(1)	Yes	No	No	'0'	
leis	TINYINT(1)	No	No	No	'0'	
votacoes	TINYINT(1)	No	No	No	'0'	
emendas	TINYINT(1)	Yes	No	No	'0'	
endereco	VARCHAR(250)	No	No	No	NULL	
hash	VARCHAR(32)	No	No	No	NULL	
campo	VARCHAR(4)	No	No	No	NULL	
publicada	TINYINT(1)	Yes	No	No	'0'	

**Table: casas\_atualizacoes**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
casa	TINYINT(4)	No	No	Yes	NULL	
tipo	VARCHAR(20)	No	No	No	NULL	
data	INT(11)	No	No	No	NULL	
url	TEXT	No	No	No	NULL	
ultima	TEXT	No	No	No	NULL	
ativa	TINYINT(1)	Yes	No	No	'1'	

**Table: casas\_orca**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	TINYINT(4)	Yes	No	Yes		

total	DECIMAL(15,2)	No	No	No	NULL	Atenção: este total é na verdade o total dividido pelo número de cadeiras
per_capita	DECIMAL(10,2)	No	No	No	NULL	
ano	YEAR(4)	Yes	No	No		

Table: casas_resumo						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
uf	TINYINT(3)	Yes	No	No	'0'	
casa	TINYINT(4)	Yes	No	Yes		
qtd_2002	INT(3)	Yes	No	No	'0'	
qtd_2004	INT(3)	Yes	No	No	'0'	
qtd_2006	INT(3)	Yes	No	No	'0'	
qtd_2008	INT(3)	Yes	No	No	'0'	
qtd_2010	INT(3)	Yes	No	No	'0'	
qtd_2012	INT(3)	Yes	No	No	'0'	
fin_2002	INT(12)	Yes	No	No	'0'	
fin_2004	INT(12)	Yes	No	No	'0'	
fin_2006	INT(12)	Yes	No	No	'0'	
fin_2008	INT(12)	Yes	No	No	'0'	
fin_2010	INT(12)	Yes	No	No	'0'	
fin_2012	INT(12)	Yes	No	No	'0'	
fis_2002	INT(11)	Yes	No	No	'0'	
fis_2004	INT(11)	Yes	No	No	'0'	
fis_2006	INT(11)	Yes	No	No	'0'	
fis_2008	INT(11)	Yes	No	No	'0'	
fis_2010	INT(11)	Yes	No	No	'0'	
fis_2012	INT(11)	Yes	No	No	'0'	
fin_cam_2002	BIGINT(12)	Yes	No	No	'0'	
fin_cam_2006	BIGINT(12)	Yes	No	No	'0'	
fin_cam_2010	BIGINT(12)	Yes	No	No	'0'	
fin_sen_2002	BIGINT(12)	Yes	No	No	'0'	
fin_sen_2006	BIGINT(12)	Yes	No	No	'0'	
fin_sen_2010	BIGINT(12)	Yes	No	No	'0'	
votos_cam_2002	INT(10)	Yes	No	No	'0'	
votos_cam_2006	INT(10)	Yes	No	No	'0'	
votos_cam_2010	INT(10)	Yes	No	No	'0'	
votos_sen_2002	INT(10)	Yes	No	No	'0'	
votos_sen_2006	INT(10)	Yes	No	No	'0'	
votos_sen_2010	INT(10)	Yes	No	No	'0'	
bens_2006	INT(12)	Yes	No	No	'0'	

bens_2008	INT(12)	Yes	No	No	'0'	
bens_2010	INT(12)	Yes	No	No	'0'	
bens_2012	INT(12)	Yes	No	No	'0'	
verbas	INT(12)	Yes	No	No	'0'	
diarias	INT(12)	Yes	No	No	'0'	
faltas	FLOAT	Yes	No	No	'0'	
faltas_com	FLOAT	Yes	No	No	'0'	
citados	INT(4)	Yes	No	No	'0'	
materias	INT(10)	Yes	No	No	'0'	
irrelevantes	FLOAT	Yes	No	No	'0'	
situacao	TINYINT(3)	Yes	No	No	'0'	
oposicao	TINYINT(3)	Yes	No	No	'0'	
independentes	TINYINT(3)	Yes	No	No	'0'	
religioso	INT(3)	Yes	No	No	'0'	
ruralista	INT(3)	Yes	No	No	'0'	
concessoes	INT(3)	Yes	No	No	'0'	
faculdade	INT(3)	Yes	No	No	'0'	
sindicalista	INT(3)	Yes	No	No	'0'	
policia	INT(3)	Yes	No	No	'0'	
ong	INT(3)	Yes	No	No	'0'	
servidor	INT(3)	Yes	No	No	'0'	
advogado	INT(3)	Yes	No	No	'0'	
empresario	INT(3)	Yes	No	No	'0'	
total	INT(3)	Yes	No	No	'0'	

**Table: categorias**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	SMALLINT(6)	Yes	Yes	No		
categoria	VARCHAR(100)	No	No	No	NULL	
relevancia	TINYINT(1)	Yes	No	No	'1'	

**Table: cnae\_divisoes**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
divisao	VARCHAR(3)	No	No	No	NULL	
denomina	VARCHAR(128)	No	No	No	NULL	
sae	TINYINT(3)	Yes	No	No	'0'	

**Table: comissoes**

Table: comissoes						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	No	No	Yes	NULL	
tipo	VARCHAR(150)	No	No	No	NULL	
nome	VARCHAR(250)	No	No	No	NULL	
data	DATE	No	No	No	NULL	
tipo_reuniao	VARCHAR(50)	No	No	No	NULL	
status	VARCHAR(50)	No	No	No	NULL	

**Table: desfaz\_materias**

Table: desfaz_materias						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	INT(11)	Yes	Yes	No		
mats	TEXT	No	No	No	NULL	
cat_orig	INT(11)	No	No	No	NULL	
cat_dest	INT(11)	No	No	No	NULL	

**Table: diarias**

Table: diarias						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	Yes		
data	DATE	No	No	No	NULL	
montante	DECIMAL(10,2)	Yes	No	No	'0.00'	
moeda	ENUM('R', 'U')	No	No	No	NULL	
casa	TINYINT(4)	Yes	No	No		

**Table: emails\_coletivos\_log**

Table: emails_coletivos_log						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	INT(11)	Yes	Yes	No		
nome	VARCHAR(250)	No	No	No	NULL	
email	VARCHAR(250)	No	No	No	NULL	
assunto	VARCHAR(128)	No	No	No	NULL	
mensagem	TEXT	No	No	No	NULL	
data	TIMESTAMP	No	No	No	CURRENT_TIMESTAMP	

ip	VARCHAR(15)	No	No	No	NULL	
----	-------------	----	----	----	------	--

**Table: emails\_prontos**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	SMALLINT(6)	Yes	Yes	No		
dest	VARCHAR(150)	No	No	No	NULL	
assunto	VARCHAR(50)	No	No	No	NULL	
conteudo	TEXT	No	No	No	NULL	
apresentacao	VARCHAR(255)	No	No	No	NULL	
ativo	TINYINT(1)	Yes	No	No	'0'	

**Table: emails\_prontos\_log**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	SMALLINT(5)	Yes	No	Yes		
nome	VARCHAR(250)	No	No	No	NULL	
email	VARCHAR(250)	No	No	No	NULL	
data	TIMESTAMP	No	No	No	CURRENT_TIMESTAMP	
ip	VARCHAR(15)	No	No	No	NULL	

**Table: emendas**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	Yes		
codigo	INT(10)	Yes	No	No		
funcao_programatica	VARCHAR(22)	No	No	No	NULL	
montante_emenda	DECIMAL(18,2)	Yes	No	No	'0.00'	
dot_inic	DECIMAL(18,2)	Yes	No	No	'0.00'	
autorizado	DECIMAL(18,2)	Yes	No	No	'0.00'	
empenhado	DECIMAL(18,2)	Yes	No	No	'0.00'	
liquidado	DECIMAL(18,2)	Yes	No	No	'0.00'	
pago	DECIMAL(18,2)	Yes	No	No	'0.00'	
ano	YEAR(4)	Yes	No	No		

**Table: emendas\_bruto\_2011**

Table Comments						
Columns						

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
nada1	VARCHAR(1)	Yes	No	No		
nada2	VARCHAR(1)	Yes	No	No		
autor	VARCHAR(255)	Yes	No	No		
partido	VARCHAR(255)	Yes	No	No		
uf	VARCHAR(255)	Yes	No	No		
codigo	VARCHAR(255)	Yes	No	No		
unidade_orcamentaria	VARCHAR(255)	Yes	No	No		
funcao_programatica	VARCHAR(255)	Yes	No	No		
titulo	VARCHAR(255)	Yes	No	No		
saldo	VARCHAR(255)	Yes	No	No		
cred_adic	VARCHAR(255)	Yes	No	No		
dot_inic	VARCHAR(255)	Yes	No	No		
autorizado	VARCHAR(255)	Yes	No	No		
empenhado	VARCHAR(255)	Yes	No	No		
liquidado	VARCHAR(255)	Yes	No	No		
pago	VARCHAR(255)	Yes	No	No		
porct	VARCHAR(255)	Yes	No	No		

**Table: emendas\_bruto\_2012**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
nada1	VARCHAR(1)	Yes	No	No		
nada2	VARCHAR(1)	Yes	No	No		
autor	VARCHAR(255)	Yes	No	No		
partido	VARCHAR(255)	Yes	No	No		
uf	VARCHAR(255)	Yes	No	No		
codigo	VARCHAR(255)	Yes	No	No		
unidade_orcamentaria	VARCHAR(255)	Yes	No	No		
funcao_programatica	VARCHAR(255)	Yes	No	No		
titulo	VARCHAR(255)	Yes	No	No		
saldo	VARCHAR(255)	Yes	No	No		
cred_adic	VARCHAR(255)	Yes	No	No		
dot_inic	VARCHAR(255)	Yes	No	No		
autorizado	VARCHAR(255)	Yes	No	No		
empenhado	VARCHAR(255)	Yes	No	No		
liquidado	VARCHAR(255)	Yes	No	No		
pago	VARCHAR(255)	Yes	No	No		
porct	VARCHAR(255)	Yes	No	No		

<b>Table: emendas_bruto_2013</b>						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
nada1	VARCHAR(1)	Yes	No	No		
nada2	VARCHAR(1)	Yes	No	No		
autor	VARCHAR(255)	Yes	No	No		
partido	VARCHAR(255)	Yes	No	No		
uf	VARCHAR(255)	Yes	No	No		
codigo	VARCHAR(255)	Yes	No	No		
unidade_orcamentaria	VARCHAR(255)	Yes	No	No		
funcao_programatica	VARCHAR(255)	Yes	No	No		
titulo	VARCHAR(255)	Yes	No	No		
saldo	VARCHAR(255)	Yes	No	No		
cred_adic	VARCHAR(255)	Yes	No	No		
dot_inic	VARCHAR(255)	Yes	No	No		
autorizado	VARCHAR(255)	Yes	No	No		
empenhado	VARCHAR(255)	Yes	No	No		
liquidado	VARCHAR(255)	Yes	No	No		
pago	VARCHAR(255)	Yes	No	No		
porct	VARCHAR(255)	Yes	No	No		

<b>Table: emendas_bruto_2014</b>						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
nada1	VARCHAR(1)	Yes	No	No		
nada2	VARCHAR(1)	Yes	No	No		
autor	VARCHAR(255)	Yes	No	No		
partido	VARCHAR(255)	Yes	No	No		
uf	VARCHAR(255)	Yes	No	No		
codigo	VARCHAR(255)	Yes	No	No		
unidade_orcamentaria	VARCHAR(255)	Yes	No	No		
funcao_programatica	VARCHAR(255)	Yes	No	No		
titulo	VARCHAR(255)	Yes	No	No		
saldo	VARCHAR(255)	Yes	No	No		
cred_adic	VARCHAR(255)	Yes	No	No		
dot_inic	VARCHAR(255)	Yes	No	No		
autorizado	VARCHAR(255)	Yes	No	No		
empenhado	VARCHAR(255)	Yes	No	No		

liquidado	VARCHAR(255)	Yes	No	No		
pago	VARCHAR(255)	Yes	No	No		
porct	VARCHAR(255)	Yes	No	No		

**Table: emendas\_porct**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
codigo	INT(11)	Yes	No	No		
porct	DECIMAL(5,4)	Yes	No	No		
ano	YEAR(4)	Yes	No	No	'2011'	

**Table: emendas\_tipos**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
funcao_programatica	VARCHAR(22)	No	No	No	NULL	
titulo	VARCHAR(255)	Yes	No	No		
ano	YEAR(4)	Yes	No	No		

**Table: estados**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	TINYINT(3)	Yes	Yes	No		
nome	VARCHAR(255)	No	No	No	NULL	
sigla	VARCHAR(10)	No	No	No	NULL	
artigo	VARCHAR(10)	Yes	No	No	'do'	
prioridade	TINYINT(3)	Yes	No	No	'0'	

**Table: filtros**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	SMALLINT(6)	Yes	Yes	No		
cat	SMALLINT(6)	No	No	Yes	NULL	
texto	TEXT	No	No	No	NULL	
tipo	TINYINT(4)	No	No	No	NULL	

**Table: ipca**

Table: ipca						
Table Comments	julho					
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
ano	YEAR(4)	Yes	No	No		
indice	DECIMAL(6,2)	Yes	No	No		

**Table: jornalistas**

Table: jornalistas						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	SMALLINT(5)	Yes	Yes	No		
nome	VARCHAR(125)	No	No	No	NULL	
funcao	VARCHAR(125)	No	No	No	NULL	
veiculo	VARCHAR(125)	No	No	No	NULL	
email	VARCHAR(125)	No	No	No	NULL	
telefone	VARCHAR(125)	No	No	No	NULL	
celular	VARCHAR(125)	No	No	No	NULL	

**Table: log\_edicao**

Table: log_edicao						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
parlamentar	INT(11)	No	No	No	NULL	
usuario	INT(11)	No	No	No	NULL	
data	DATETIME	No	No	No	'2001-01-01 00:00:00'	

**Table: materias**

Table: materias						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	INT(11)	Yes	Yes	No		
antigo	VARCHAR(200)	No	No	No	NULL	
nome	VARCHAR(200)	No	No	No	NULL	
numero	BIGINT(20)	No	No	No	NULL	
ano	YEAR(4)	No	No	No	NULL	
promulgado	TINYINT(1)	Yes	No	No	'0'	
casa	SMALLINT(6)	No	No	No	NULL	
ementa	TEXT	No	No	No	NULL	
categoria	TINYINT(3)	Yes	No	No		

atualizacao	TIMESTAMP	No	No	No	ON UPDATE CURRENT_TIMESTAMP	
temporario	TINYINT(1)	Yes	No	No	'-1'	

**Table: municipios**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	SMALLINT(5)	Yes	Yes	No		
codigo	VARCHAR(5)	Yes	No	No		
estado	TINYINT(3)	Yes	No	No		
capital	TINYINT(1)	Yes	No	No	'0'	
nome	VARCHAR(255)	Yes	No	No		

**Table: ocupacoes**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
codigo	INT(11)	Yes	No	No		
descricao	VARCHAR(128)	No	No	No	NULL	

**Table: parl\_admin**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	Yes		
situacao	TINYINT(1)	Yes	No	No	'0'	
status	ENUM('Edição', 'Supervisão', 'Aprovação', 'Publicação', 'Publicado')	Yes	No	No	'Edição'	
mat1	INT(10)	Yes	No	No	'0'	antigo pca_matricula
mat2	INT(10)	Yes	No	No	'0'	antigo pca_matricula_2006
mat3	INT(10)	Yes	No	No	'0'	antigo pca_depid
mat4	VARCHAR(255)	No	No	No	NULL	antigo pca_matricula_str
mat5	INT(10)	Yes	No	No	'0'	
cnae	TINYINT(3)	Yes	No	No	'0'	
data_processos	DATE	Yes	No	No		

data_dnj	DATE	Yes	No	No		
data_salva	DATE	Yes	No	No		

**Table: parl\_bancadas**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	Yes		
religioso	TINYINT(1)	Yes	No	No	'0'	
concessoes	TINYINT(1)	Yes	No	No	'0'	
faculdade	TINYINT(1)	Yes	No	No	'0'	
ruralista	TINYINT(1)	Yes	No	No	'0'	
sindicalista	TINYINT(1)	Yes	No	No	'0'	
policia	TINYINT(1)	Yes	No	No	'0'	
ong	TINYINT(1)	Yes	No	No	'0'	
servidor	TINYINT(1)	Yes	No	No	'0'	
advogado	TINYINT(1)	Yes	No	No	'0'	
empresario	TINYINT(1)	Yes	No	No	'0'	
parente	TINYINT(1)	Yes	No	No	'0'	

**Table: parl\_cnae**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	No		
divisao	VARCHAR(2)	No	No	No	NULL	
montante	DECIMAL(12,0)	Yes	No	No	'0'	
ano	YEAR(4)	Yes	No	No		

**Table: parl\_dnj**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	Yes		
dnj	TEXT	No	No	No	NULL	

**Table: parl\_materias**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment

id	MEDIUMINT(8)	Yes	No	No		
mat_id	INT(11)	Yes	No	Yes		
deletar	TINYINT(1)	Yes	No	No	'0'	

**Table: parl\_plenario**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	No	'0'	
data	DATE	Yes	No	No	'2001-01-01'	
status	VARCHAR(51)	No	No	No	NULL	

**Table: parl\_processos**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	Yes		
processos	TEXT	No	No	No	NULL	

**Table: parl\_votacoes**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(9)	Yes	Yes	Yes		
materia	VARCHAR(255)	Yes	Yes	No	"	
data	DATE	Yes	Yes	No		
codigo	VARCHAR(255)	Yes	Yes	No	"	
presenca	VARCHAR(50)	Yes	Yes	No	"	
voto	VARCHAR(250)	Yes	Yes	No	"	
sessao	INT(11)	Yes	Yes	No		

**Table: parlamentares**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	No		
nome	VARCHAR(150)	No	No	No	NULL	
apelido	VARCHAR(250)	No	No	No	NULL	
partido	INT(11)	No	No	No	NULL	
partido_eleito	INT(11)	No	No	No	NULL	

cpf	VARCHAR(14)	No	No	No	NULL	
titulo	VARCHAR(14)	No	No	No	NULL	
cargos	TEXT	No	No	No	NULL	
dados	TEXT	No	No	No	NULL	
comentario	TEXT	No	No	No	NULL	
uf	TINYINT(4)	Yes	No	No		
hits	BIGINT(20)	Yes	No	No	'0'	
hits_mail	BIGINT(20)	Yes	No	No	'0'	
email	VARCHAR(250)	No	No	No	NULL	
casa	INT(11)	No	No	No	'0'	
historico_partidos	TEXT	No	No	No	NULL	
legislatura	INT(11)	Yes	No	No	'0'	
url	VARCHAR(250)	No	No	No	NULL	
apelido2	VARCHAR(100)	No	No	No	NULL	
suplente	MEDIUMINT(8)	Yes	No	No	'0'	
cand_2014	TINYINT(3)	Yes	No	No	'0'	

**Table: partidos**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	TINYINT(3)	Yes	Yes	No		
sigla	VARCHAR(11)	Yes	No	No		
nome	VARCHAR(255)	No	No	No	NULL	
numero	TINYINT(4)	Yes	No	No		
primeiro_ano	YEAR(4)	Yes	No	No	'2002'	
ultimo_ano	YEAR(4)	Yes	No	No	'2012'	

**Table: partidos\_bancada**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	TINYINT(3)	Yes	No	Yes		
sigla	VARCHAR(11)	Yes	No	No		
ano	YEAR(4)	Yes	No	No	'2013'	
situacao	ENUM('Situação', 'Oposição', 'Independente')	Yes	No	No		

**Table: plenario\_ant**

Table Comments	
----------------	--

Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	Yes	'0'	
data	DATE	Yes	No	No	'2001-01-01'	
status	VARCHAR(51)	No	No	No	NULL	

Table: quemquer						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	BIGINT(20)	Yes	Yes	No		
estado	TINYINT(3)	Yes	No	No	'0'	
cargo	TINYINT(3)	Yes	No	No	'0'	
partido	TINYINT(3)	Yes	No	No	'0'	
nome	VARCHAR(128)	No	No	No	NULL	
cand	MEDIUMINT(9)	Yes	No	No	'0'	
origem	ENUM('quemquer', 'cel', 'api_busca', 'api_exc', 'api_compara', 'api_bens', 'api_doacoes', 'api_passadas', 'busca_parte', 'google')	Yes	No	No	'quemquer'	
data	TIMESTAMP	Yes	No	No	CURRENT_TIMESTAMP	
filtro	TINYINT(3)	Yes	No	No	'0'	
tempo	FLOAT	Yes	No	No	'0'	
ip	VARCHAR(45)	No	No	No	NULL	

Table: quemquer cola						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
tipo	ENUM('imprime', 'envia')	Yes	No	No	'imprime'	
estado	TINYINT(3)	Yes	No	No	'0'	
Presidente	TINYINT(3)	Yes	No	No	'0'	
Governador	TINYINT(3)	Yes	No	No	'0'	
Senador	INT(10)	Yes	No	No	'0'	
Federal	INT(10)	Yes	No	No	'0'	
Estadual	INT(10)	Yes	No	No	'0'	
ip	VARCHAR(45)	No	No	No	NULL	

<b>Table: quemquer_testa</b>						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
server	TEXT	Yes	No	No		
id	BIGINT(20)	Yes	No	No	'0'	

<b>Table: recados</b>						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	Yes		
quem	MEDIUMINT(8)	Yes	No	Yes		
texto	TEXT	Yes	No	No		
data	TIMESTAMP	Yes	No	No	CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP	

<b>Table: registros</b>						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	Yes		
quem	MEDIUMINT(8)	Yes	No	Yes		
data	DATE	Yes	No	No		

<b>Table: relatorios</b>						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
arquivo	VARCHAR(40)	Yes	No	No		
legenda	VARCHAR(70)	Yes	No	No		
resumo	TEXT	No	No	No	NULL	
ano	YEAR(4)	Yes	No	No		
data	TIMESTAMP	Yes	No	No	CURRENT_TIMESTAMP	
ativo	TINYINT(1)	Yes	No	No	'1'	

**Table: resultados**

Table: resultados						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	TINYINT(3)	Yes	Yes	No		
codigo	TINYINT(4)	Yes	No	No		
descricao	VARCHAR(255)	Yes	No	No		

**Table: sae**

Table: sae						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
cod	TINYINT(3)	Yes	No	No	'0'	
descricao	VARCHAR(64)	No	No	No	NULL	

**Table: temp**

Table: temp						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	No	No		
cpf	VARCHAR(20)	No	No	No	NULL	
titulo	VARCHAR(20)	No	No	No	NULL	

**Table: textos**

Table: textos						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	TINYINT(4)	Yes	Yes	No		
codigo	VARCHAR(20)	No	No	No	NULL	
titulo	VARCHAR(124)	No	No	No	NULL	
texto	TEXT	No	No	No	NULL	

**Table: usuarios**

Table: usuarios						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	Yes	Yes	No		
nome	VARCHAR(100)	No	No	No	NULL	
sobrenome	VARCHAR(100)	No	No	No	NULL	

usuario	VARCHAR(100)	No	No	No	NULL	
email	VARCHAR(100)	No	No	No	NULL	
senha	VARCHAR(32)	No	No	No	NULL	
categoria	ENUM('administrador', 'coordenador', 'editor', 'redator', 'visitante')	No	No	No	NULL	
ativo	ENUM('false', 'true')	No	No	No	NULL	

**Table: verbas**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	No	No	Yes	NULL	
ano	INT(11)	No	No	No	NULL	
mes	INT(11)	No	No	No	NULL	
codigo	SMALLINT(5)	Yes	No	No	'0'	
despesa	VARCHAR(255)	No	No	No	NULL	
montante	FLOAT	No	No	No	NULL	
casa	TINYINT(2)	Yes	No	No	'0'	
dt_atualiza	TIMESTAMP	Yes	No	No	CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP	

**Table: verbas\_tipos**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
original	VARCHAR(255)	No	No	No	NULL	
codigo	SMALLINT(5)	Yes	No	No	'0'	
categoria	ENUM('Aluguel/Despesas diversas', 'Divulgação/Consultorias', 'Transportes/Estadias', 'Desconto')	Yes	No	No		
casa	INT(11)	No	No	No	NULL	

**Table: viagens**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	MEDIUMINT(8)	No	No	Yes	NULL	
data_i	DATE	No	No	No	NULL	

data_f	DATE	No	No	No	NULL	
assunto	VARCHAR(254)	No	No	No	NULL	
destino	VARCHAR(254)	No	No	No	NULL	
qtd	DECIMAL(10,2)	No	No	No	NULL	
montante	DECIMAL(10,2)	No	No	No	NULL	
moeda	ENUM('R', 'U')	No	No	No	NULL	
dt_atualiza	TIMESTAMP	Yes	No	No	CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP	

## **D. Dicionário de Dados do Esquema da UFS**

# Schema Report for database: Lista de Tabelas - Base UFS

Table: candidato						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	INT(11)	Yes	Yes	No		
nome	VARCHAR(255)	Yes	No	No		
foto	VARCHAR(145)	No	No	No	NULL	
data_nascimento	DATETIME	No	No	No	NULL	
cpf	CHAR(11)	No	No	No	NULL	
titulo_eleitoral	CHAR(12)	No	No	No	NULL	
email	VARCHAR(255)	No	No	No	NULL	
cidade_id	INT(11)	No	No	Yes	NULL	
ocupacao_id	INT(11)	No	No	Yes	NULL	
nacionalidade_id	INT(11)	No	No	Yes	NULL	
grau_instrucao_id	INT(11)	No	No	Yes	NULL	
genero_id	INT(11)	No	No	Yes	NULL	
cor_raca_id	INT(11)	No	No	Yes	NULL	

Table: deputado						
Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	INT(11)	Yes	Yes	No		
nome	VARCHAR(255)	No	No	No	NULL	
nome_civil	VARCHAR(255)	Yes	No	No		
data_nascimento	DATE	Yes	No	No		
data_falecimento	DATE	No	No	No	NULL	
sexo	VARCHAR(2)	No	No	No	NULL	
sigla_partido	VARCHAR(20)	Yes	No	No		
sigla_uf	VARCHAR(2)	Yes	No	No		
municipio_nascimento	VARCHAR(100)	No	No	No	NULL	
uf_nascimento	VARCHAR(2)	No	No	No	NULL	
data_ultimo_status	DATE	No	No	No	NULL	
gabinete_nome	VARCHAR(45)	No	No	No	NULL	
gabinete_predio	VARCHAR(45)	No	No	No	NULL	
gabinete_sala	VARCHAR(45)	No	No	No	NULL	

gabinete_andar	VARCHAR(45)	No	No	No	NULL	
gabinete_telefone	VARCHAR(45)	No	No	No	NULL	
gabinete_email	VARCHAR(100)	No	No	No	NULL	
website	VARCHAR(100)	No	No	No	NULL	
candidato_id	INT(11)	No	No	Yes	NULL	

**Table: hot\_posicionamento\_candidato**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id_candidato	INT(11)	Yes	No	No		
id_tema	INT(11)	Yes	No	No		
titulo	VARCHAR(255)	No	No	No	NULL	
descricao	TEXT	No	No	No	NULL	
id_categoria	INT(11)	No	No	No	NULL	
categoria	VARCHAR(45)	No	No	No	NULL	
id_posicao	INT(11)	Yes	No	No		
posicao	VARCHAR(20)	No	No	No	NULL	
fonte_informacao	VARCHAR(255)	Yes	No	No		Link para fonte
sigla_projeto	VARCHAR(10)	No	No	No	NULL	
numero_projeto	INT(11)	No	No	No	NULL	
ano_projeto	INT(11)	No	No	No	NULL	
referencia_1_id	BIGINT(11)	No	No	No	NULL	
referencia_1_titulo	VARCHAR(255)	No	No	No	NULL	
referencia_1_url	VARCHAR(45)	No	No	No	NULL	
referencia_2_id	BIGINT(11)	No	No	No	NULL	
referencia_2_titulo	VARCHAR(255)	No	No	No	NULL	
referencia_2_url	VARCHAR(45)	No	No	No	NULL	
referencia_3_id	BIGINT(11)	No	No	No	NULL	
referencia_3_titulo	VARCHAR(255)	No	No	No	NULL	
referencia_3_url	VARCHAR(45)	No	No	No	NULL	

**Table: hot\_vigiaqui\_politico\_processado**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id_candidato	INT(11)	No	No	No	NULL	
processado	TINYINT(4)	Yes	No	No		

**Table: hot\_vigiaqui\_processos**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id_candidato	INT(11)	No	No	No	NULL	
id_processo	INT(11)	Yes	No	No		
numero	VARCHAR(255)	Yes	No	No		
tipo	VARCHAR(255)	Yes	No	No		
tribunal	VARCHAR(255)	Yes	No	No		
descricao	TEXT	No	No	No	NULL	
link	TEXT	Yes	No	No		

**Table: processo\_judicial**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	INT(11)	Yes	Yes	No		
numero	VARCHAR(255)	Yes	No	No		
tipo	VARCHAR(255)	Yes	No	No		
tribunal	VARCHAR(255)	Yes	No	No		
descricao	TEXT	No	No	No	NULL	
link	TEXT	Yes	No	No		
politico_vigieaqui_id	INT(11)	Yes	No	Yes		

**Table: tema\_posicao\_candidato**

Table Comments						
Columns						
Name	Data Type	Nullable	PK	FK	Default	Comment
id	INT(11)	Yes	Yes	No		
candidato_id	INT(11)	Yes	No	Yes		
tema_id	INT(11)	Yes	No	Yes		
tema_posicao_id	INT(11)	Yes	No	Yes		
url_fonte	VARCHAR(255)	Yes	No	No		Link para fonte

## **E. Plataforma de Dados Abertos - Layout Arquivo Candidatura**

## 2 CONSULTA\_CAND

NOTAÇÃO:

consulta\_cand\_<ANO>\_<UF>

consulta\_cand\_<ANO>\_BRASIL

### 2.1 Observação

Para o caso de candidatos não divulgáveis, seus dados pessoais são preenchidos com:

- -4, em caso de campos numéricos, exceto campo de idade;
- "NÃO DIVULGÁVEL", em caso de campos textuais;
- "", no caso de campos relativos a idade do candidato e sua data de nascimento.

Variável	Descrição
DT_GERACAO	Data de geração do arquivo (data da extração dos dados) .
HH_GERACAO	Hora de geração do arquivo (hora da extração) - Horário de Brasília.
ANO_ELEICAO	Ano da eleição (referente ao ano eleitoral de pesquisa) .
CD_TIPO_ELEICAO	Código do tipo de eleição.
NM_TIPO_ELEICAO	Nome do tipo de eleição.
NR_TURNO	Número do turno da eleição.
CD_ELEICAO	Código da eleição.
DS_ELEICAO	Descrição da eleição.
DT_ELEICAO	Data em que ocorreu a eleição.
TP_ABRANGENCIA	Tipo de abrangência da eleição.
SG_UF	Sigla da Unidade da Federação em que ocorreu a eleição.
SG_UE	Sigla da Unidade Eleitoral do candidato (Em caso de eleição majoritária é a sigla da UF que o candidato concorre e em caso de eleição municipal é o código TSE do município) .
NM_UE	Nome de Unidade Eleitoral do candidato (Em caso de eleição majoritária é o nome da UF que o candidato concorre e em caso de eleição municipal é o nome do município) .
CD_CARGO	Código do cargo ao qual o candidato concorre.
DS_CARGO	Descrição do cargo ao qual o candidato concorre.
SQ_CANDIDATO	Número sequencial do candidato gerado internamente pelos sistemas eleitorais. Não é o número de campanha do candidato.
NR_CANDIDATO	Número do candidato na urna.
NM_CANDIDATO	Nome completo do candidato.
NM_URNA_CANDIDATO	Nome de urna do candidato.
NM_SOCIAL_CANDIDATO	Nome social do candidato.

Variável	Descrição
NR_CPF_CANDIDATO	Número do CPF do candidato.
NM_EMAIL	Endereço de e-mail do candidato.
CD_SITUACAO_CANDIDATURA	Código da situação do registro de candidatura do candidato.
DS_SITUACAO_CANDIDATURA	Descrição da situação do registro de candidatura do candidato. Pode assumir os valores 'Apto' (candidato apto para ir para urna) ; 'Inapto' (candidato inapto para ir para urna) ; 'Cadastrado' (registro de candidatura realizado, mas ainda não julgado) .
CD_DETALHE_SITUACAO_CAND	Código do detalhe da situação do registro de candidatura do candidato.
DS_DETALHE_SITUACAO_CAND	Descrição do detalhe da situação do registro de candidatura do candidato.
TP_AGREMIACAO	Tipo de agremiação. Pode assumir os valores 'Coligação', quando o candidato concorre por coligação; 'Partido isolado', quando o candidato concorre somente pelo partido.
NR_PARTIDO	Número do partido do candidato.
SG_PARTIDO	Sigla do partido do candidato.
NM_PARTIDO	Nome do partido do candidato.
SG_LEGENDA	Sigla da coligação gerado pela Justiça Eleitoral, da qual o candidato pertence.
SQ_COLIGACAO	Sequencial da coligação gerado pela Justiça Eleitoral, da qual o candidato pertence.
NM_COLIGACAO	Nome da coligação da qual o candidato pertence.
DS_COMPOSICAO_COLIGACAO	Descrição da composição da coligação da qual o candidato pertence.
CD_NACIONALIDADE	Código da nacionalidade do candidato.
DS_NACIONALIDADE	Descrição da nacionalidade do candidato.
SG_UF_NASCIMENTO	Sigla da UF de nascimento do candidato.
CD_MUNICIPIO_NASCIMENTO	Código TSE do município de nascimento do candidato.
NM_MUNICIPIO_NASCIMENTO	Nome do município de nascimento do candidato.
DT_NASCIMENTO	Data de nascimento do candidato.
NR_IDADE_DATA_POSSE	Idade do candidato na data da posse. Consultar a data da posse para cada cargo e unidade eleitoral nos arquivos de vagas.
NR_TITULO_ELEITORAL_CANDIDATO	Número do título eleitoral do candidato.
CD_GENERO	Código do gênero do candidato.
DS_GENERO	Descrição do gênero do candidato.
CD_GRAU_INSTRUCAO	Código do grau de instrução do candidato.
DS_GRAU_INSTRUCAO	Descrição do grau de instrução do candidato.

## **F. Relatório de Codificação - Categorias e Grupos**

<b>Código/Categoria</b>	<b>Significado</b>	<b>Grupos de Códigos</b>
Características do Sistema	Funcionalidades	
Comportamento	Reações ai lidar com o sistema	
Conhecimento da Utilidade	Percepção da utilidade	
Contribuições para a sociedade	Reconhecimento das contribuições	
Curiosidade	Curiosidade em realizar buscas sobre os candidatos	
Necessidade de Melhorias	Sugestões de Melhorias	
Períodos Críticos	Períodos críticos potencializarm o uso	
EducaçãoPolítica_Distanciamento	Distanciamento das questões ou educação política	Educação_Politica
EducaçãoPolítica_Engajamento	Participa de discussões e pesquisa informações em fontes mais sólidas.	Educação_Politica
EducaçãoPolítica_Pesquisa_não_usual	Pesquisar sobre políticos não é um hábito frequente	Educação_Politica
FacilidadeUso_Agrupamento	Organização das informações na interface	Facilidade_Uso
FacilidadeUso_Busca_Rápida	Percepção sobre a performance do retorno das buscas	Facilidade_Uso
FacilidadeUso_Fluxo_de_Navegação	Os link levam a blocos de informações, mas deve retornar pra o mesmo ponto anterior.	Facilidade_Uso
FacilidadeUso_Foco_no_Texto	Interação mais focada no conteúdo textual	Facilidade_Uso
FacilidadeUso_ImagensIrrelevantes	Imagens não chamaram atenção	Facilidade_Uso
FacilidadeUso_Intuitivo	A ferramenta aparenta ser intuitiva	Facilidade_Uso
FacilidadeUso_Legendas	Requer inclusão de legendas	Facilidade_Uso
FacilidadeUso_Motivação	Gera motivação para pesquisar sobre os candidatos	Facilidade_Uso
FacilidadeUso_Requer_Componentes_Intuitivos	Descrição do ícones requerida.	Facilidade_Uso
FacilidadeUso_Satisfatório	Satisfação com a utilização dos recursos de busca do artefato.	Facilidade_Uso
Fasiv_Comportamento_Político	Comportamento do candidato como fator de influência.	Fator_Influência
Fasiv_Experiência_Política	Experiência Política como fator de influência.	Fator_Influência
Fasiv_Fonte_de_Informação_Mídia_Internet	A mídia e internet como forma de influência a(o) cidadã(o)	Fator_Influência
Fasiv_Identificação_Ideológica	Identificação ideológica como fator de influência	Fator_Influência
Fasiv_Influência_Social	Meio social como fator de influência	Fator_Influência
FontelInfo_Acesso	Acesso a informação	Fonte_Informação
FontelInfo_Alienação	Não conhece ou não compreende devido ao distanciamento da realidade	Fonte_Informação
FontelInfo_Atenção	Atenção aos conteúdos, acontecimentos e atualização	Fonte_Informação
FontelInfo_Conhecimento	Conhecimento sobre os candidatos	Fonte_Informação
FontelInfo_Decisão	Tomada de decisão de voto	Fonte_Informação
FontelInfo_Desconfiança	Desconfiança da informação	Fonte_Informação
FontelInfo_FakeNews	Ocorrências de FakeNews	Fonte_Informação
FontelInfo_Histórico_Político	Histórico do político como fonte de informação	Fonte_Informação
FontelInfo_Ignorância	Não sabe onde encontrar a informação	Fonte_Informação
FontelInfo_Indisponibilidade	Falta de difusão da informação dificulda ao acesso.	Fonte_Informação
FontelInfo_Informação_Política	Conhecimento sobre a informação politico-partidária	Fonte_Informação
FontelInfo_Informação_Tendenciosa	Informação disponibilizada em prol da promoção de alguém ou do próprio dono (político) do site.	Fonte_Informação
FontelInfo_Internet	Internet como fonte de informação	Fonte_Informação
FontelInfo_Jornais	Jornais da mídia e internet e impressos como fonte de informação.	Fonte_Informação
FontelInfo_Meio_Social	Meio social a que pertence e atualiza-se.	Fonte_Informação
FontelInfo_Necessidade	Necessidade de informação e conhecimento	Fonte_Informação
FontelInfo_Observação	Observa a opinião alheia.	Fonte_Informação
FontelInfo_Oficial	Origem das fontes	Fonte_Informação
FontelInfo_Oportunidade	Oportunidade de pesquisa, de explorar as informações	Fonte_Informação
FontelInfo_Propostas	Propostas dos candidatos	Fonte_Informação
FontelInfo_Questionamento_Origem	Curiosidade em relação a origem dos dados	Fonte_Informação
FontelInfo_Redes_Sociais	Redes Sociais como fonte de informação	Fonte_Informação
IntensãoUso_Compartilhamento	Intensão de compartilhar o sistema com outras pessoas ou grupos.	Intensão_Uso
IntensãoUso_Contínuo	Uso contínuo do artefato	Intensão_Uso
IntensãoUso_Interesse	Interesse no conteúdo	Intensão_Uso
IntensãoUso_Novidade	Primeira vez que interage com ferramenta similar	Intensão_Uso
IntensãoUso_Períodos_Eleitorais	Período de maior utilização	Intensão_Uso
IntensãoUso_Segurança	Segurança quando informado que os dados são oriundos de fontes oficiais.	Intensão_Uso
RedesSociais_Conflitos	Conflitos em redes sociais causados por discussões sobre política	Redes_Sociais_Digitais
RedesSociais_Evasão_Grupos	Movimento de saída dos grupos nas redes sociais	Redes_Sociais_Digitais
RedesSociais_Fact_Cheking	Hábito de verificar a autenticidade da informação	Redes_Sociais_Digitais
RedesSociais_Fake_News	Ocorrência de mensagens falsas nas redes sociais	Redes_Sociais_Digitais
RedesSociais_Recomendação	Informações são direcionadas pelas redes sociais como uma recomendação filtrada.	Redes_Sociais_Digitais
UtilidadePercebida_Aprofundamento	Necessidade de um leitura mais aprofundada sobre candidato.	Utilidade_Percebida
UtilidadePercebida_Auditoria	Ferramenta como iniciativa de transparência e auditoria de informações	Utilidade_Percebida

UtilidadePercebida_Busca_Parametrizada	Sugestão de melhorias na interface com a inclusão de novos filtros de busca	Utilidade_Percebida
UtilidadePercebida_Conexões	O artefato apresenta conexões/relações entre os políticos	Utilidade_Percebida
UtilidadePercebida_Espanto	Ao utilizar a ferramenta teve a sensação de espanto, devido quantidade de informações registradas.	Utilidade_Percebida
UtilidadePercebida_Estranheza	Desconfiança das informações encontradas em relação as informações/opiniões do meio social.	Utilidade_Percebida
UtilidadePercebida_Expectativa	Expectativas em relação a participação na entrevista e também com as possíveis funcionalidades do sistema	Utilidade_Percebida
UtilidadePercebida_Informação_Concentrada	Artefato aparenta ter informações completas	Utilidade_Percebida
UtilidadePercebida_Informação_Confiável	Conexões levam a sítios oficiais de informação e aumenta confiabilidade	Utilidade_Percebida
UtilidadePercebida_Informação_Resumida	Alguns fatos trazem informação resumida	Utilidade_Percebida
UtilidadePercebida_Linguagem_Cidadã	As informações precisam ser claras, numa linguagem que o cidadão entenda.	Utilidade_Percebida
UtilidadePercebida_Melhorias	Pontos de melhorias (funcionalidades, diagramação, componentes, explicações)	Utilidade_Percebida
UtilidadePercebida_Oportunidades	Cenário de abertura de dados implica em oportunidades de pesquisas	Utilidade_Percebida
UtilidadePercebida_Poder	Relação entre o conhecimento e a forma de obtenção ou fortalecimento de poder.	Utilidade_Percebida
UtilidadePercebida_Processos	Conteúdo de processos destacou-se como relevante	Utilidade_Percebida
UtilidadePercebida_Progressivo	Reconhecimento gradativo da utilidade do artefato	Utilidade_Percebida
UtilidadePercebida_Publicos_Especificos	Publico alvo, usuários do sistema.	Utilidade_Percebida
UtilidadePercebida_Relevância	Conteúdo relevante	Utilidade_Percebida
UtilidadePercebida_Texto_Longo	Importância do conteúdo textual, porém longo.	Utilidade_Percebida
UtilidadePercebida_Transparência	Artefato como forma de obter transparência das informações	Utilidade_Percebida

## **G. Implementação da Extração dos Dados do Mandato**

/home/junior/extrairMandatos.java

```

1 /*
2  * To change this license header, choose License Headers in Project Properties.
3  * To change this template file, choose Tools | Templates
4  * and open the template in the editor.
5  /**
6   *
7   * @author José Soares Barroso Júnior
8   */
9 //Extracao de dados dos mandatos
10 var i, j;
11 var mandato_politico = "";
12
13 if (historico_cargos != null) {
14 var periodos = historico_cargos.match(/(\d{4}-?)+/g);
15
16 if (periodos != null) {
17     var num = (periodos[0].substr()).length;
18     if (num > 4) {
19         var tamanho = periodos.length;
20         for (i = 0; i < tamanho; i++){
21             var inicio = periodos[i].substr(0,4);
22             var fim = periodos[i].substr(5,4);
23             for (j = abs(inicio); j <= abs(fim); j++) {
24                 if ( mandato_politico.indexOf(j.toString()) == -1 ) {
25                     mandato_politico += (j >= abs(inicio) && i == 0)?
26                     (j).toString() + "-" :j == abs(fim) ? (j).toString() :
27                     (j).toString() + "-";
28                 }
29             }
30         }
31     }
32     else
33         mandato_politico = periodos[0];
34 }
35 var cargo = "";
36 var presidente = historico_cargos.match(/([A-z])residente/);
37 if (presidente) cargo = "Presidente";
38 var vicepresidente = historico_cargos.match(/([A-z])ice-?([A-z])residente/);
39 if (vicepresidente) cargo = "Vice-Presidente";
40 var deputadofederal = historico_cargos.match(/([A-z])eputad([a-z])
41     \s([A-z])ederal/);
42 if (deputadofederal) cargo = "Deputado Federal";
43 var deputadoestadual = historico_cargos.match(/([A-z])eputad([a-z])
44     \s([A-z])stadual/);
45 if (deputadoestadual) cargo = "Deputado Estadual";
46 var senador = historico_cargos.match(/([A-z])enador([a-z])?/);
47 if (senador) cargo = "Senador";
48 var governador = historico_cargos.match(/([A-z])overnador([a-z])?/);
49 if (governador) cargo = "Governador";
50 var prefeito = historico_cargos.match(/([A-z])refeit([A-z])/);
51 if (prefeito) cargo = "Prefeito";
52 var viceprefeito = historico_cargos.match(/([A-z])ice-?([A-z])
53     refeit([a-z])+/);
54 if (viceprefeito) cargo = "Vice-Prefeito";
55 var ministro = historico_cargos.match(/([A-z])inistr([A-z])/);
56 if (ministro) cargo = "Ministro";
57 var suplente = historico_cargos.match(/([A-z])uplente/);
58 if (suplente) cargo = "Suplente";
59 var secretario = historico_cargos.match(/([A-z])ecret([A-Za-zA-u])
60     ri([a-z])/);
61 if (secretario) cargo = "Secretario";
62 var vereador = historico_cargos.match(/([A-z])ereador([a-z])?/);
63 if (vereador) cargo = "Vereador";
64 var presidentedaassembleia = historico_cargos.match(/([A-z])residente\
65     sd([a-z])\s([A-z])ssembl([A-Za-zA-u])ia/);
66 if (presidentedaassembleia) cargo = "Presidente da Assembléia

```