



JOSIR CARDOSO GOMES

**Financiamento de Campanha Política da Bancada Evangélica da
Câmara Federal: uma abordagem com o uso de Mineração de
Dados**

Defesa de Doutorado

**RIO DE JANEIRO
2022**

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO – UFRJ
ESCOLA DE COMUNICAÇÃO – ECO
INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA –
IBICT
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO – PPGCI

JOSIR CARDOSO GOMES

FINANCIAMENTO DE CAMPANHA POLÍTICA DA BANCADA EVANGÉLICA DA
CÂMARA FEDERAL: UMA ABORDAGEM COM O USO DE MINERAÇÃO DE
DADOS

Tese de Doutorado apresentada à Banca Examinadora do curso de Doutorado em Ciência da Informação da Universidade Federal do Rio de Janeiro / IBICT como parte dos requisitos para obtenção do grau de Doutor.

Orientador: Prof. MARCO SCHNEIDER

RIO DE JANEIRO - RJ
2022

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

Catálogo na Publicação

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a), sob a responsabilidade de Miguel Romeu Amorim Neto - CRB-7/6283.

C633f CARDOSO GOMES, JOSIR
Financiamento de Campanha Política da Bancada
Evangélica da Câmara Federal: uma abordagem com o
uso de Mineração de Dados / JOSIR CARDOSO GOMES. --
Rio de Janeiro, 2022.
176 f.

Orientador: Marco André Feldman Schneider.
Tese (doutorado) - Universidade Federal do Rio
de Janeiro, Escola da Comunicação, Instituto
Brasileiro de Informação em Ciência e Tecnologia,
Programa de Pós-Graduação em Ciência da Informação,
2022.

1. bancada evangélica. 2. financiamento de
campanha eleitoral. 3. mineração de dados. 4.
clusterização. I. Feldman Schneider, Marco André,
orient. II. Título.

Financiamento de Campanha Política da Bancada Evangélica da Câmara Federal: :
uma abordagem com o uso de Mineração de Dados.

Defesa de Doutorado apresentada ao Programa de Pós-graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e Universidade Federal do Rio de Janeiro (UFRJ)/Escola de Comunicação (ECO), como requisito parcial para a obtenção do título de doutor em Ciência da Informação.

Rio de Janeiro, 10 de fevereiro de 2022.

BANCA EXAMINADORA

Prof. Dr. Marco André Feldman Schneider (Orientador)
PPGCI – IBICT/UFRJ

Prof. Dr. Arthur Coelho Bezerra (Membro Interno)
PPGCI – IBICT/UFRJ

Prof. Dr. Gustavo Silva Saldanha (Membro Interno)
PPGCI – IBICT/UFRJ

Prof. Dr. Ricardo Medeiros Pimenta (Membro Interno)
PPGCI – IBICT/UFRJ

Prof. Dra. Magali do Nascimento Cunha (Membro externo)
ISER

Prof. Dr. Francisco Josué Medeiros de Freitas (Membro externo)
PPGTI – UFRJ/IFCS

Prof. Dr. Marcelo Fornazin (Suplente interno)
PPGCI – IBICT/UFRJ

AGRADECIMENTOS

Dedico este trabalho à minha melhor amiga e também esposa Patricia Saldanha que me acompanhou amorosamente nessa jornada. Seu incentivo, conselhos e, principalmente, o carinho foram essenciais para que essa pesquisa pudesse ser realizada. Sem ela ao meu lado, provavelmente nenhuma linha aqui seria escrita, ou se fosse, não teria o perfume de sua atenção e dedicação.

Não posso dizer que esse carinho e compreensão me surpreenderam. Ao longo de 37 primaveras juntos, desde os meus 18 anos, essa foi a nossa rotina e nosso projeto de vida: apoiar um ao outro incondicionalmente. E mesmo nos momentos em que eu não estive tão presente, ela sempre esteve comigo.

Patricia já orientou centenas de monografias, dissertações e teses e acompanhou com um olhar atento os rumos do meu trabalho. Me deu dicas valiosas e várias vezes revisou aquele parágrafo que não saía de jeito nenhum. Entretanto, sempre soube manter uma distância segura para não atravessar uma linha tênue que pudesse invadir o meu espaço como pesquisador ou as recomendações do meu orientador. Enfim, seguiu me apoiando atentamente e sempre com muito carinho.

Às minhas filhas Bruna e Carolina e ao meu filho João Pedro, frutos de muito amor, agradeço pelo carinho, paciência e incentivo. São sementes que Deus me emprestou para que eu as cultivasse. Floresceram e hoje são motivo de muito orgulho para mim. Suas ações e atitudes foram um motivo extra para que eu seguisse em frente. Se tornaram pessoas fortes, solidárias e felizes. Felizes por terem também companheiros e companheiras que pensam como eles. Agradeço ao Luis, ao Eduardo e a Isabela por terem entrado na família.

Ao meu pai Josir e minha mãe Fernanda, agradeço o apoio e, principalmente, o exemplo de vida de vocês. A dedicação dos dois com a nossa família, mesmo em momentos tão adversos, guiou o meu comportamento e conduta. Sei que voltar a estudar na UFRJ, universidade onde meu pai trabalhou por tantos anos, foi um motivo de uma alegria especial para eles.

Meu pai é um exemplo de perseverança e, sempre que o cansaço batia, eu lembrava de como foi difícil para ele estudar, em plena ditadura, e sem o apoio que eu sempre tive. Ao terminar a faculdade apenas aos 31 anos de idade, ele e minha mãe se desdoblaram para pagar uma boa escola para mim. Assim, espero que esta

etapa que acabo de cumprir os deixem muito orgulhosos e com a sensação de dever cumprido.

À minha irmã querida Camila, ao Daniel, Nuno e Maitê que mesmo distantes, sempre mandaram boas energias e boas notícias, vídeos e fotos para animar o meu dia. Espero que esta moça talentosa siga também um caminho acadêmico assim que as coisas se acalmarem.

Agradeço também a Maria Saldanha, minha sogra que nunca foi sogra, e ao meu sogro Raymundo Saldanha (in memoriam) que me receberam como um filho e me apoiaram desde o início do nosso namoro. Lembro como se fosse hoje, quando “Seu” Raymundo vinha sempre no escritório e falava comigo: “essa tese sai ou não sai ?!” Saiu Seu Raymundo!

À minha cunhada Raquel Saldanha e seus filhotes Rafaela e Pedro pela parceria nos momentos difíceis e pela alegria de ter vocês próximos.

Ao meu médico e amigo Álcio Braz que me ajudou demais neste último ano, me ensinou um pouco sobre o zen-budismo e foi fundamental para que eu mantivesse a sanidade durante a pandemia e neste caos que o nosso país se transformou.

Agradecimento especial ao meu orientador, Marco Schneider pela atenção dedicada, pelas correções minuciosas e pela amizade que foi se formando ao longo desse projeto de pesquisa. As orientações regadas a filosofia, Marx e também a cerveja foram essenciais para que a pesquisa tenha se tornado prazerosa.

Aos professores e funcionários do PPGCI/IBICT, especialmente ao Professor Ricardo Pimenta, que abriu as portas do LARHUD para as pesquisas que desenvolvi ao longo do Doutorado e que abriu também uma grande janela para o universo das Humanidades Digitais.

Aos professores Gustavo Saldanha, Arthur Bezerra e Fabio Gouveia não só pelas aulas excelentes mas também pela amizade que formamos ao longo deste tempo na Praia Vermelha.

Ao professor Felipe Figueiredo pela consultoria estatística que foi muito importante para o fechamento da tese e aos pesquisadores Fábio Lacerda, Vitor de Moraes Peixoto e Dirceu Gerardi pelo apoio, principalmente em relação à base de financiamento de campanha.

À equipe do IBICT Brasília, em especial ao Tiago Braga e Washington Segundo que acreditaram no meu trabalho.

Aos professores Gerson Lachtermacher e Paulo Sérgio Coelho que me apresentaram o mundo da Mineração de Dados ainda no curso de Mestrado e pelo suporte que eles me proporcionaram na FGV.

À Media Systems, e em especial ao Elias, Marion e D.Eugênia (in memorian) pelo apoio ao permitirem que eu trabalhasse em um horário reduzido para que pudesse fazer o Doutorado. Agradeço também à minha equipe, Isnard, Fabricio, Fábio e Samuel que tiveram muita paciência com as minhas ausências. Findo o Doutorado, espero poder voltar a ajudar mais nos nossos projetos.

E principalmente a Deus, que me mostrou, através da fé, que vale a pena amar e que todo trabalho feito com amor é um lampejo de sua existência.

GOMES, Josir C. Financiamento de Campanha Política da Bancada Evangélica da Câmara Federal: uma abordagem com o uso de Mineração de Dados. Doutorado em Ciência da Informação no PPGCI IBICT/UFRJ– ECO. Rio de Janeiro, 2022.

RESUMO

A atual turbulência política no Brasil tem sido fonte de grande inquietação e o rumo da nossa incipiente democracia tem se tornado incerto. Em especial, o crescimento das candidaturas evangélicas ao longo dos últimos 20 anos trouxe pautas conservadoras que podem limitar ou cercear a liberdade de minorias que não compartilham os mesmos valores morais dos diferentes grupos religiosos evangélicos. Como um dos principais fatores de sucesso eleitoral é o financiamento das campanhas políticas, estudar como se dá o financiamento de campanha desses candidatos pode trazer à tona um novo conhecimento sobre esse fenômeno. A hipótese central é que novas categorias não triviais emergirão deste processamento e que seja possível categorizar os deputados federais evangélicos a partir do padrão de financiamento dos mesmos. A partir dos dados de financiamento de campanha política e a base de resultados eleitorais do TSE, foram utilizadas técnicas de mineração de dados e segmentação para encontrar padrões de financiamento dentre os deputados federais eleitos em 2018 que se autodenominam evangélicos com o intuito de comparar as classificações das denominações evangélicas com as classificações geradas por máquina, fazendo assim uma análise crítica dos resultados encontrados. A pesquisa concluiu que o padrão de financiamento de deputados evangélicos é distinto dos demais deputados e que foi possível classificá-los em grupos distintos. Sendo o foco da pesquisa empírica a influência dos evangélicos na política brasileira, o desenvolvimento de um arcabouço metodológico que facilite a classificação de agentes políticos em função do tipo de financiamento e do sucesso (ou insucesso) eleitoral, abre novas possibilidades na área da ciência política e também para a CI, que tem entre um dos seus pilares os sistemas de recuperação e organização do conhecimento.

Palavra-chave: bancada evangélica, financiamento de campanha eleitoral, mineração de dados, clusterização

ABSTRACT

The current political turbulence in Brazil has been a source of great concern and the path of our incipient democracy has become uncertain. In particular, the growth of evangelical candidacies over the last 20 years has brought conservative agendas that may limit or curtail the freedom of minorities who do not share the same moral values as different evangelical religious groups. As one of the main factors of electoral success is the financing of political campaigns, studying the campaign financing of these candidates may bring up new knowledge about this phenomenon. The central hypothesis is that new non-trivial categories will emerge from this processing and that it will be possible to categorize evangelical federal deputies based on their financing pattern. Political campaign finance data and the TSE election results database were used to find patterns and classifications among federal deputies elected in 2018 who are evangelicals. From the collected data, clustering techniques were used in order to compare the existing evangelical denominations classification with the machine-generated classification, thus making a critical analysis of the results. The research concluded that the financing pattern of evangelical deputies is distinct from that of other deputies, and it was possible to classify them into distinct groups. As the focus of the empirical research was the presence of evangelicals in Brazilian politics, the development of a methodological framework that facilitates the classification of political agents according to the type of funding and electoral success (or failure) opens new possibilities in the area of political science and also for information science, which has information retrieval and knowledge organization as one of its pillars.

Keywords: evangelical bench, electoral financing, datamining, clustering

LISTA DE FIGURAS

Figura 1 – Processo de Mineração de Dados Proposto.....	34
Figura 2 – Tecnologias utilizadas pela Mineração de Dados.....	47
Figura 3 – Tipos de SOC (KOS).....	63
Figura 4 – Fluxo de Doações nas Eleições Presidenciais em 2010.....	118

LISTA DE GRÁFICOS

Gráfico 1 – Percentual da população residente, por grupos de religião.....	88
Gráfico 2 – Total de Deputados Federais Evangélicos de 1946 a 2019.....	102
Gráfico 3 – Deputados Evangélicos por Bloco Ideológico (%)......	110
Gráfico 4 – Exemplo de Fluxo Financeiro das doações de uma empresa.....	123
Gráfico 5 – Fluxo Financeiro das doações em 2014.....	125
Gráfico 6 – Fluxo Financeiro de doações em 2018.....	126
Gráfico 7 – Doações aos candidatos por Tipo de doador.....	127
Gráfico 8 – Doações em 2014 por Setor Econômico.....	130
Gráfico 9 – Doações por Setor Econômico – Somente para Deputados Federais. .	131
Gráfico 10 – Doações por Setor Econômico – Somente para Deputados Federais	135
Gráfico 11 – Doação por Setor Econômico e por Bloco Ideológico em 2018.....	136
Gráfico 12 – Doação por Setor Econômico e por Bloco Ideológico em 2002.....	137
Gráfico 13 – Doação por Setor Econômico e por Bloco Ideológico em 2006.....	137
Gráfico 14 – Doação por Setor Econômico e por Bloco Ideológico em 2010.....	138
Gráfico 15 – Doação por Setor Econômico e por Bloco Ideológico em 2014.....	138
Gráfico 16 – Conjunto de Dados a Segmentar.....	146
Gráfico 17 – Resultado da segmentação.....	147
Gráfico 18 – Exemplo de segmentação Complexa.....	147
Gráfico 19 – Exemplo de segmentação Complexa.....	148
Gráfico 20 – Histograma da primeira tentativa de segmentação.....	150
Gráfico 21 – Exemplo de Outlier.....	151
Gráfico 22 – Histograma do segundo processamento.....	151
Gráfico 23 – Outliers por Partido.....	152
Gráfico 24 – Outliers por Denominação.....	153
Gráfico 25 – Comparação dos Outliers.....	153
Gráfico 26 – Clusterização Final.....	154

LISTA DE TABELAS

Tabela 1 - Total de Artigos na BRAPCI de 2010 a 2018 sobre Mineração de Dados.....	52
Tabela 2 - Classificação das Religiões na CDD.....	70
Tabela 3 - Classificação das Religiões na CDU.....	70
Tabela 4 - Classificação das Religiões na Colon Classification.....	71
Tabela 5 - Classificação das Religiões Cristãs na CDD, CDU e CC.....	71
Tabela 6 - Religiões no Mundo em 2010.....	73
Tabela 7 - Cristianismo no Mundo em 2010.....	74
Tabela 8 - Países com maior número de protestantes em 2010.....	76
Tabela 9 - População de Evangélicos – CENSO 2010.....	89
Tabela 10 - Denominações cristãs não católicas em 2010.....	89
Tabela 11 - Deputados Federais evangélicos por Denominação Religiosa.....	112
Tabela 12 - Financiamento Eleitoral Privado entre 2010 e 2018.....	116
Tabela 13 - Setores Econômicos.....	129
Tabela 14 - Doadores por Setores Econômicos em 2018.....	133
Tabela 15 - Faixas de Doação em 2018.....	134
Tabela 16 - Características dos deputados federais eleitos em 2018.....	144
Tabela 17 - Efeito da receita total recebida por deputados federais.....	145
Tabela 18 - Distribuição das Denominações na clusterização.....	156
Tabela 19 - Distribuição das Denominações por Ideologia.....	156
Tabela 20 - Distribuição por Corporações Pentecostais.....	156
Tabela 21 - Distribuição das Denominações por Partido.....	157

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
AD	Assembleia de Deus
AGP	Agente Partidário
CC	Colon Classification
CDD	Classificação Decimal de Dewey
CDU	Classificação Decimal Universal
CGU	Controladoria-Geral da União
CI	Ciência da Informação
CNPJ	Cadastro Nacional de Pessoa Jurídica
CSV	<i>Comma Separated Values</i> : Formato de Planilha que utiliza vírgulas como separador entre as colunas.
FEFC	Fundo Especial de Financiamento de Campanha
IA	Inteligência Artificial
IECLB	Igreja Evangélica de Confissão Luterana no Brasil
IEQ	Igreja do Evangelho Quadrangular
IIGD	Igreja Internacional da Graça de Deus
INDA	Infraestrutura Nacional de Dados Abertos
IPB	Igreja Presbiteriana do Brasil
IURD	Igreja Universal do Reino de Deus
KDD	do Inglês, <i>Knowledge Discovery on Database</i> , Descoberta de conhecimento em banco de dados.
LAI	Lei de Acesso à Informação
MD	Mineração de Dados
NBR	Prefixo utilizado para Normas Brasileiras, emitidas pela ABNT
ODS	<i>Open Document Sheet</i> : Formato de planilha utilizado em

formato aberto.

OKF	<i>Open Knowledge Foundation</i> : ONG que fomenta o uso de dados abertos e conhecimento livre para a humanidade.
OWL	<i>Web Ontology Language</i> : Linguagem para a construção de ontologias.
PF	Pessoa Física: ente contábil que representa um ser humano considerado como sujeito de direitos e obrigações.
PJ	Pessoa Jurídica: ente contábil que indica uma organização que é detentora de direitos e obrigações.
PRC	Pew Research Center
RDF	<i>Resource Description Framework</i> : padrão para descrever conceitos e modelos de metadados baseado em XML
SOC	Sistema de Organização do Conhecimento
SPARQL	<i>RDF Query Language</i> : Linguagem para consultar o formato RDF
TP	Teologia da Prosperidade
TSE	Tribunal Superior Eleitoral
XLS	Excel Sheet: Formato de planilha utilizado pelo software MS-Excel de propriedade da Microsoft.

SUMÁRIO

1	Introdução.....	23
1.1	O PROBLEMA.....	25
1.2	OBJETIVO GERAL E OBJETIVOS ESPECÍFICOS.....	29
1.3	JUSTIFICATIVA.....	30
1.4	FUNDAMENTAÇÃO TEÓRICA E METODOLÓGICA.....	32
1.5	ESTRUTURA DO TRABALHO.....	35
2	Dados, Mineração de Dados e a Ciência da Informação.....	38
2.1	O CONCEITO DE DADO, INFORMAÇÃO E CONHECIMENTO.....	38
2.2	MINERAÇÃO DE DADOS.....	44
2.3	MINERAÇÃO DE DADOS NA CIÊNCIA DA INFORMAÇÃO.....	51
2.4	DADOS ABERTOS.....	52
3	Classificação das Religiões.....	59
3.1	SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO.....	60
3.2	AS RELIGIÕES NOS SISTEMAS BIBLIOGRÁFICOS.....	66
3.3	O PROTESTANTISMO NO MUNDO.....	72
3.4	O PROTESTANTISMO NO BRASIL.....	80
3.5	A CLASSIFICAÇÃO DOS EVANGÉLICOS NO CENSO DO IBGE.....	87
4	O Espírito Evangélico na Política Brasileira.....	93
4.1	HISTÓRIA DOS EVANGÉLICOS NA POLÍTICA.....	95
4.2	AS CORPORAÇÕES PENTECOSTAIS E OS PARTIDOS POLÍTICOS.....	106
5	Financiamento de Campanha Política.....	114
5.1	HISTÓRICO.....	114
5.2	SISTEMA ATUAL.....	115
5.3	BASES DE DADOS ABERTOS DO TSE.....	120
5.4	FINANCIAMENTO POR SETORES ECONÔMICOS.....	127
6	Resultados.....	140
6.1	MONTAGEM DO BANCO DE DADOS.....	140
6.2	VARIÁVEIS UTILIZADAS.....	141
6.3	ANÁLISE ESTATÍSTICA E REGRESSÃO LOGÍSTICA.....	143
6.4	SEGMENTAÇÃO.....	146
6.5	PROCESSAMENTO.....	149

7 Conclusões.....	158
8 Referências Bibliográficas.....	163
9 Objetos Digitais.....	175
10 ANEXO 1.....	176
11 Glossário.....	177

1 INTRODUÇÃO

Em nenhum outro momento da história, o ser humano pôde obter tantas informações sobre o universo que o cerca tão facilmente e de forma tão barata como na atualidade. Pela Internet e pelas mídias digitais, é possível obter de informações do cotidiano da cidade, informações sobre políticas públicas e artigos científicos capazes de impactar na subsistência de toda uma comunidade.

Entretanto, o acesso a essas informações não se dá de maneira uniforme ou universal. Ainda existe uma divisão muito clara entre quem tem ou não acesso à essa multiplicidade de dados e à chamada utopia tecnoliberal (BEZERRA; SCHNEIDER; SALDANHA, 2013). Desde a Grécia antiga até o início do século XX, a tradição da ética em informação era caracterizada por duas ideias: a liberdade de expressão (desde a Grécia antiga) e a liberdade de imprensa (desde Gutemberg). Um terceiro desafio surgiu então com o novo mundo conectado em rede: a liberdade de acesso. (CAPURRO, 2001)

Este acesso não envolve somente a obtenção do dado físico, do documento ou de uma base de dados, mas também requer que se consiga interpretar e manipular os dados disponíveis de forma prática. Como a quantidade de dados disponíveis tem crescido exponencialmente, o processo de “garimpar” informações relevantes tem se tornado cada vez mais complexo. E a expectativa é que o volume cresça ainda mais: com o advento da Internet das Coisas (IoT) que permite coletar e disponibilizar medições em tempo real de objetos físicos espalhados dentro de nossas casas, na cidade ou até mesmo em satélites orbitais, os dados coletados já ultrapassam o limite do administrável.

A Ciência também já se encontra mergulhada nesse “tsunami” de dados e, em especial para as Ciências Humanas e Sociais, a demanda se evidencia, por exemplo, pelo crescimento do emprego de ferramentas e métodos digitais que permitam o uso de tal volume de dados. Pesquisadores e estudantes têm buscado o letramento necessário para se capacitar em torno da produção de novas metodologias capazes de produzir os meios adequados para analisar seus objetos de pesquisa em meio ao universo digital pelo qual estamos inseridos.

Pode-se verificar tal interesse a partir do crescimento de congressos e publicações nas chamadas Humanidades Digitais. “As humanidades digitais

designam uma transdisciplina, portadora dos métodos, dos dispositivos e das perspectivas heurísticas ligadas ao digital no domínio das Ciências Humanas e Sociais” (THATCAMP, 2011). E talvez, das ciências humanas e sociais, a Ciência da Informação (CI) é a que mais faz interseção com as práticas digitais. Segundo Pimenta,

as ciências humanas são hoje o grande campo no qual a inserção da computação e de recursos diversos marcados pelo Big Data mais destacam tal inflexão multidisciplinar e que, talvez, para a Ciência da Informação mais aponte para a necessidade de ser por ela discutida uma vez que informação, seu acesso, suportes, sistemas, usuários, tanto em aspecto público como privado tornam-se elementos chave para compreender o que convencionou-se chamar de Humanidades Digitais. (PIMENTA, 2016, p.22)

Assim, ao buscar novas metodologias que façam uso intensivo da tecnologia de informação, a Ciência da Informação pode abrir novos horizontes de pesquisa e novas formas de tratar problemas complexos das Ciências Sociais Aplicadas. De fato, a Ciência da Informação dispõe de um amplo ferramental na área de Recuperação da Informação, na área de Organização do Conhecimento, na análise crítica sobre o uso de Dados Abertos Governamentais e no próprio gerenciamento de dados de pesquisa científica. Assim, a interseção destas disciplinas na pesquisa pode colaborar com o surgimento de novas formas de estudar antigos problemas.

No início de 2017, o grupo de pesquisa Perfil-I (Perspectivas Filosóficas da Informação) iniciou o projeto de pesquisa FARMi (Financiamento de Campanha, Atuação Política e Repercussão Midiática e da Informação) para construir um banco de dados que permitisse agregar diversas variáveis ao conjunto de dados de financiamento de campanha disponibilizado em regime de acesso público pelo Tribunal Superior Eleitoral no Brasil. O banco de dados serviria não só para a pesquisa aqui relatada mas para que outros pesquisadores pudessem utilizar os dados para outros temas.

Outra ação do projeto FARMi é o mapeamento as diversas iniciativas construídas pela sociedade civil, acadêmicas ou não, que busquem contribuir para a promoção de transparência dos poderes políticos, competência crítica em informação e cidadania ampliada. Até Dezembro de 2018, o mapeamento contava com 45 projetos que foram divididos em 3 principais categorias: financiamento de campanha, atuação política e repercussões midiáticas e, dentro da categoria de Atuação Política, foram catalogadas iniciativas que investigam ações do Legislativo,

Executivo e do Judiciário. A divulgação dos trabalhos tem sido feita paulatinamente através de um site na web¹ e, a cada novo projeto encontrado, é feita uma resenha divulgada no site e várias conclusões foram obtidas a partir desse mapeamento. (GOMES; SCHNEIDER; BEZERRA, 2018)

Em função das iniciativas que envolveram o uso de tecnologias de processamento e análise de dados, em 2018 o grupo de pesquisa Perfil-I foi convidado para fazer parte do Laboratório em Rede de Humanidades Digitais (LARHUD) que tem como objetivo incentivar a produção científica e tecnológica, além de desenvolver ferramentas e metodologias mediadas no vasto campo das Humanidades intermediada pelas plataformas, ferramentas e registros digitais.

Para atingir tal intento, o primeiro passo foi avaliar a viabilidade de construir um repositório que pudesse abrigar os dados de financiamento de campanha, agregando a este repositório, características específicas dos candidatos que participavam do pleito e outras informações que pudessem servir de insumo para a análise desejada.

A ideia era montar em um único banco de dados, as receitas de campanha política de 1998 a 2018, a base de resultados eleitorais do TSE do mesmo período, a base de dados empresariais da Receita Federal, dados demográficos e econômicos das cidades e os dados pessoais dos candidatos. De posse dessa coleção, seria possível diferenciar os políticos em função do seu padrão de financiamento, e no caso específico desta pesquisa, como veremos à frente, se os deputados federais e estaduais que se autodenominam evangélicos teriam um padrão de financiamento diferente dos demais deputados.

1.1 O PROBLEMA

A motivação desta pesquisa surgiu em função da atual turbulência política no Brasil que tem sido fonte de grande inquietação sobre o rumo que a incipiente democracia tem seguido. E esta turbulência tinha um componente de preocupação em especial: o crescimento das candidaturas evangélicas que traziam em seu discurso pautas conservadoras que poderiam limitar ou cercear a liberdade de minorias que não compartilhassem dos mesmos valores morais que os evangélicos

¹ <http://www.farmi.pro.br>

defendem. Ronaldo de Almeida, professor de Antropologia da Unicamp também compartilha dessa preocupação em recente entrevista:

Apesar de uma maioria evangélica-protestante ainda não ser uma realidade no Brasil, pouco a pouco os líderes pentecostais mais conservadores e combativos começam a bater na tecla da supremacia de uma maioria judaico-cristã², que incluiria os católicos. Quem não estiver de acordo que se sujeite ao desejo da maioria (Fundação FHC, 2019).

Vale ressaltar que o termo genérico “evangélico” utilizado no Brasil é equivalente ao termo “protestante” nos EUA e o que chamamos “protestante” aqui no Brasil é denominado “Main Lane Protestant Church” nos EUA. Segundo Mafra, o termo genérico “Evangélico” não reflete a multiplicidade de grupos sociais, econômicos, crenças religiosas e morais que se encaixam dentro deste segmento religioso. Entretanto, “dada a visibilidade pública que esse segmento ganhou na opinião pública, se forjou um certo consenso referendando o termo ‘evangélico’ como categoria abrangente” (MAFRA, 2001, p.7).

Durante toda a pesquisa, buscou-se definir de forma detalhada os termos utilizados para este segmento religioso e a preocupação em não estigmatizar os evangélicos esteve sempre presente. Ao diferenciar a religiosidade cristã e evangélica dos interesses políticos e econômicos de determinados grupos que operam neste segmento da sociedade, a pesquisa procurou não ser mais um dos instrumentos discriminatórios que tem atingido os evangélicos ao longo dos anos no Brasil.

Ao longo das últimas 3 décadas, o número de políticos evangélicos cresceu vertiginosamente e este número pode ser facilmente contabilizado através da chamada bancada evangélica e também da Frente Parlamentar Evangélica (FPE). Entende-se aqui por bancada evangélica, o conjunto de políticos que se autodenominam evangélicos composta por 13% do total dos deputados eleitos no Congresso Nacional na legislatura 2014-2018. Em 2010, esta mesma bancada já havia crescido em relação à representação de 2006: tomaram posse, em 2 fevereiro de 2011, 74 parlamentares, sendo 71 deputados e 3 senadores evangélicos, ou seja, um aumento considerável dado que em 2006 a mesma bancada tinha apenas 52

² Pelo censo de 2010, existem aproximadamente 300 mil brasileiros que professam a fé judaica. Dessa forma, entende-se que o professor tenha usado o termo “judaico-cristã” em função da origem do cristianismo a partir da religião judaica. Tal termo não é muito preciso, pois não cabe falar, por exemplo, em uma maioria judaico-islâmica na Arábia Saudita mesmo sendo o Islã originário também do Judaísmo.

integrantes. Já a Frente Parlamentar Evangélica é uma organização formal dentro do Congresso formada por deputados que apoiam as causas defendidas pelos evangélicos e nesta mesma legislatura contava com a assinatura de 199 deputados³.

Já nas eleições de 2018, a bancada evangélica cresceu ainda mais. No Senado, a bancada quase quadruplicou, passando de 4 para 14 senadores e na Câmara, o número passou para 116 deputados⁴. E essa ascensão também chegou no Poder Executivo com a eleição à presidência de Jair Bolsonaro. Apesar do candidato não ser evangélico, realizou um culto de batismo no Rio Jordão em Israel, em plena campanha eleitoral, para angariar os votos dos evangélicos e teve apoio da grande parte das lideranças evangélicas.

E a estratégia parece que deu certo: ao analisar a pesquisa amostral do Datafolha que antecedeu o segundo turno das eleições de 2018, o pesquisador José Eustáquio Alves concluiu que os evangélicos foram decisivos na eleição do presidente da República. Na pesquisa, a diferença de votos entre Fernando Haddad e Bolsonaro (11 milhões de votos aproximadamente) equivaleria justamente ao número de eleitores evangélicos que migraram seus votos para o então candidato Bolsonaro⁵. Apesar de não ser possível correlacionar que estes 11 milhões de votos específicos sejam dos evangélicos, o fato é que houve uma grande influência de grupos evangélicos nas eleições de 2018.

O crescimento evangélico está associado também à ascensão da Direita e dos movimentos conservadores e reacionários às pautas identitárias e as que buscam a diminuição da desigualdade social. Por mais que não exista unanimidade dentro deste segmento religioso e que lideranças evangélicas mais progressistas tentem apresentar alternativas a este modelo conservador, existe uma clara hegemonia dos setores de direita no campo evangélico. Cabe estudar em que grau esta hegemonia pode ser contestada.

Efetivamente, a bancada que representa os evangélicos se tornou uma força política com grande representação tanto no Executivo como no Legislativo. Inúmeros pesquisadores têm buscado entender como se deu esse crescimento. No campo

³ <http://www.camara.leg.br/internet/deputado/frenteDetalhe.asp?id=53658>. Acessado em: 01,Jun,2016.

⁴ Estes números já são resultado da pesquisa e divergem dos números indicados pelo DIAP que indicou 84 deputados. <https://www.diap.org.br/index.php/noticias/noticias/88900-eleicoes-2018-bancada-evangelica-cresce-na-camara-e-no-senado> Acessado em: 01/03/2021

⁵ <http://www.ihu.unisinos.br/584304>

das Ciências Sociais, alguns dos mais profícuos são os pesquisadores Paul Freston, Ricardo Mariano e Antônio Pierucci. Recentemente Lacerda (2017) apresentou uma análise das candidaturas evangélicas de 1986 até 2010, demonstrando a preponderância das igrejas neopentecostais. Christina Cunha relatou como se deu a campanha a presidente do Pastor Everaldo e a força do apoio da comunidade evangélica à sua candidatura (CUNHA, LEITE e LUI, 2017) e Magali Cunha já com o enfoque na Comunicação e na cultura Gospel realizou uma análise histórica da ascensão dos evangélicos na obra “Do púlpito às mídias sociais” (CUNHA, 2019).

Entretanto, além das análises qualitativas e sociológicas, um outro caminho também pode ser trilhado para estudar estas candidaturas. Como um dos principais fatores de sucesso eleitoral está associado ao financiamento das campanhas políticas (MANCUSO e SPECK, 2015; PEIXOTO, 2016), estudar como se dá o financiamento de campanha desses candidatos pode trazer à tona novos conhecimentos sobre o fenômeno eleitoral.

A partir de uma perspectiva marxista, talvez seja possível vislumbrar relações entre frações de classes e outros grupos de interesse aos políticos da bancada evangélica. Novas perspectivas podem ser encontradas ao relacionar quais frações de classe e grupos econômicos financiam quais políticos, se tal financiamento se diferencia dos políticos não evangélicos e para quais outros políticos esses capitais tendem a investir no âmbito do financiamento de campanhas.

De antemão, já se sabia que o número de variáveis a explorar seria muito grande e que talvez métodos tradicionais de análise não fossem adequados, além de serem muito custosos em termos de tempo de análise. Nesse contexto, o uso de técnicas de mineração de dados (MD) poderiam ser úteis já que, por definição, tais técnicas são ferramentas apropriadas para tratar um enorme volume de dados e para buscar padrões não triviais nos dados disponíveis (HAN; KAMBER; PEI, 2012).

Além de tentar encontrar dados relevantes sobre a bancada evangélica, a análise do processo de MD a ser realizada em uma base com dados eleitorais, sociais e políticos poderia trazer novas frentes de pesquisa para a Ciência da Informação já que o método investiga o acesso aos dados abertos governamentais e ao processo de classificação dos atores evangélicos (instituições religiosas, personagens relevantes neste contexto e partidos políticos).

Assim, esta pesquisa busca verificar duas hipóteses:

h1: o financiamento dos deputados evangélicos eleitos segue um padrão diferente dos demais deputados;

h2: novas classificações sobre os políticos evangélicos emergirão a partir do uso de técnicas de mineração de dados e tais classificações abrirão caminho para que cientistas e analistas políticos possam gerar novas teorias sobre o impacto e influência da bancada evangélica na política brasileira a partir da classificação encontrada.

1.2 OBJETIVO GERAL E OBJETIVOS ESPECÍFICOS

O presente estudo tem como objetivo criar uma tipologia dos políticos que compõe a bancada evangélica utilizando técnicas de Mineração de Dados.

Já os objetivos específicos são os seguintes:

1. Criar uma base de dados que consolide os dados de financiamento de campanha, resultado eleitoral desde 1998 a 2020 onde também seja possível a inserção de novos atributos que classifiquem os candidatos a cargos eletivos no Brasil. Este objetivo tem como foco o reuso dos dados de pesquisa e está no cerne das discussões do campo da Ciência da Informação sobre Ciência Aberta e de como armazenar e difundir os dados de pesquisa.
2. Avaliar o desempenho de algoritmos de Mineração de Dados neste domínio do conhecimento. Ao realizar o processamento de dados, será possível avaliar se os algoritmos empregados agregam novas formas de “visualizar” os dados eleitorais e se estes resultados fazem sentido dentro do campo da Ciência Política.
3. Desenvolver uma metodologia que facilite a classificação de agentes políticos em função do tipo de financiamento de campanha e do sucesso (ou insucesso) eleitoral para que tal metodologia possa ser utilizada tanto na área da ciência política como também para a CI, que tem entre um dos seus pilares os sistemas de classificação e organização do conhecimento.

1.3 JUSTIFICATIVA

Com a promulgação da Lei de Acesso a Informação em 2011, a oferta de novas fontes de dados governamentais tem crescido e proporcionado novas possibilidades de análises da realidade brasileira. Concomitantemente, novas ferramentas e métodos digitais e a cultura baseada em software livre abriram novas possibilidades metodológicas para o uso de tais volumes de dados para a pesquisa científica, principalmente no chamado Sul Global onde a escassez de recursos computacionais e financeiros sempre foi uma constante.

A Ciência da Informação tem concentrado esforços no estudo dos usos dos Dados Abertos Governamentais, bem como na análise crítica de sua adoção pelos diversos atores da sociedade (BRAGA e TUZZO, 2016; GOMES, SCHNEIDER e BEZERRA, 2018). Já no estudo das ferramentas e métodos digitais, pesquisas sobre Humanidades Digitais têm dissecado novas possibilidades metodológicas no campo das Humanidades e Ciências Sociais. Por fim, pesquisas na área de sistemas de organização do conhecimento tem trabalhado com a questão tanto das melhores práticas de classificação quanto nas novas formas de representação do conhecimento através de ontologias e taxonomias.

Os termos mineração de dados, aprendizado de máquina (*machine learning* ou ML), redes neurais (*neural networks*) e outros correlatos têm invadido não só as manchetes de jornal e as redes sociotécnicas como também tem aparecido com uma maior frequência em periódicos científicos. Apesar de boa parte destes conceitos terem sido criados na ciência da computação desde a década de 60, somente agora seus usos têm se popularizado em função do maior processamento dos computadores pessoais, da disponibilização de serviços na nuvem e do surgimento de produtos e serviços que fazem uso destas tecnologias.

A discussão sobre Inteligência Artificial (IA) também tem aparecido constantemente no noticiário, seja para relatar os últimos testes (e acidentes) de carros autônomos ou para anunciar os novos dispositivos domésticos como o Alexa e o Google Home que já podem ser adquiridos a preços módicos. Entusiastas na área alardeiam que o uso de IA será um tsunami (MANNING, 2015) ou uma revolução (SEJNOWSKI, 2018) que irá alterar profundamente a sociedade, incluindo aí o próprio fazer científico.

Mesmo fazendo um filtro de todo o discurso mercadológico e de marketing de empresas interessadas em fazer novos negócios, percebe-se que o próprio meio científico já aderiu a este discurso. Cursos de “ciência da dados” começam a surgir em várias universidades, tanto na área da Estatística e de Computação que seriam o território natural da chamada Ciência de Dados mas também em programas de CI.

Especificamente na CI já existem inúmeros estudos de como os algoritmos têm impactado os aspectos sociais, econômicos e culturais na sociedade contemporânea. Questões sobre a invasão de privacidade, vigilância, desinformação e os fluxos informacionais e econômicos das redes sociais privadas já fazem parte dos anais e revistas científicas do campo. Entretanto, o campo poderia aproveitar mais essas tecnologias para criar métodos de pesquisa que consigam analisar tal volume e variedade de dados.

Ao longo da revisão bibliográfica, foram encontrados poucos trabalhos acadêmicos que tratam da interseção da CI com o uso de mineração de dados, embora existam algumas poucas pesquisas na área de recuperação da informação e indexação (KOURIS, MAKRIS e TSAKALIDIS, 2005; CARVALHO e TSUNODA, 2018), na criação de taxonomias semi-automatizadas e no uso de mineração que apoiam estudos bibliométricos. Importante citar a pesquisa de Nhacuongue (2015) que trata também de mineração de dados na CI que, apesar de não realizar uma pesquisa empírica, trouxe contribuições importantes para esta tese. Desta forma, uma pesquisa empírica sobre esta interseção pode contribuir para o fortalecimento do campo e sugerir novas metodologias de pesquisa nas Ciências Sociais, em especial na Ciência Política, numa perspectiva transversalizada pelo projeto das Humanidades Digitais.

Apesar do foco da pesquisa empírica ser a classificação dos candidatos evangélicos na política brasileira, o desenvolvimento de um arcabouço metodológico que facilite a classificação de agentes políticos em função do tipo de financiamento e do sucesso eleitoral, abre novas possibilidades na área da ciência política e também para a CI, que tem entre um dos seus pilares os sistemas de classificação e organização do conhecimento.

A aposta metodológica é que uma seleção criteriosa e criativa dos dados aliada ao processo de MD permitirá extrair novos conhecimentos sobre os diversos grupos

políticos que compõe o segmento evangélico e, em última instância, a cena política nacional.

Espera-se que a mesma metodologia que foi utilizada nesta tese, possa ser útil para outros temas, ou seja, aqui se está analisando a bancada evangélica mas o mesmo raciocínio pode ser utilizado para a bancada ruralista, para grupos de interesse na área de infraestrutura ou para políticos que representem o interesse de determinado grupo de trabalhadores.

1.4 FUNDAMENTAÇÃO TEÓRICA E METODOLÓGICA

Para testar a hipótese de que novas classificações sobre os políticos evangélicos possam ser encontradas a partir do uso de técnicas de mineração de dados, será construída uma metodologia utilizando técnicas de segmentação sobre um problema real de classificação comparando as classificações existentes das denominações evangélicas com as classificações geradas por máquina.

A classificação de objetos complexos para encaixá-los em um número menor de categorias é fundamental para a pesquisa científica. As ciências naturais e comportamentais estão repletas de taxonomias, esquemas de categorização e tipologias e esta categorização está intimamente relacionada à definição e medição: os atributos dos objetos de interesse da pesquisa são coletados para, em seguida, formar “tipos” ou categorias, dependendo dos dados coletados e organizados (AHLQUIST e BREUNIG, 2012).

Especificamente, o termo “classificação” já traz uma imediata correlação com o campo das Ciências da Informação que busca não só os mais eficazes métodos para classificar mas também procura analisar criticamente quais são os impactos do processo de classificação na sociedade.

A metodologia desta pesquisa é quantitativa e terá como parâmetro o processo clássico de MD que envolve a montagem da base de dados, a seleção das variáveis que serão tratadas pelo algoritmo, a limpeza dos dados, o processamento em si quando algoritmos são aplicados e por fim, a análise dos resultados obtidos. Além disso, o processo trará algumas inovações com o intuito de acelerar e dar maior qualidade ao desenvolvimento da pesquisa e ao resultado das análises.

Todo o processamento será desenvolvido na linguagem de programação Python e com bancos de dados MySQL. Tanto o Python como o MySQL são softwares livres, ou seja, não necessitam de licença proprietária e tem o código fonte disponível na Internet. Ao não utilizar um software proprietário, a pesquisa fica independente das políticas informacionais impostas pelo proprietário do software. Esse é um elemento importante da metodologia pois muitas vezes, uma pesquisa deixa de ser reproduzível porque o software utilizado não está mais disponível, ou porque os algoritmos utilizados na pesquisa original foram alterados sem aviso prévio ou pelo custo de aquisição do software se tornar proibitivo em função de mudanças nas políticas de preços do fornecedor do software (AQUINO, 2015).

Além disso, ao utilizar uma linguagem de programação livre, todas as etapas do processamento são documentadas na própria linguagem, ou seja, o “caderno laboratorial” é o próprio código fonte que executa as operações nos dados. Assim, um outro pesquisador consegue reproduzir passo a passo todas as transformações que foram realizadas nos dados e validar ou estender o experimento.

Esta etapa de documentação via código faz parte do que se convencionou chamar de *e-science* e tal tema é uma das fronteiras de estudo da CI que tem se envolvido ativamente no estudo de como está sendo feito o gerenciamento e organização da produção científica. Falar em *e-science* é falar de recuperação da informação, de sistemas de organização do conhecimento, de indexação e também de economia política da informação (MEDEIROS e CAREGNATO, 2012; SALES, SAYÃO e SOUZA, 2013).

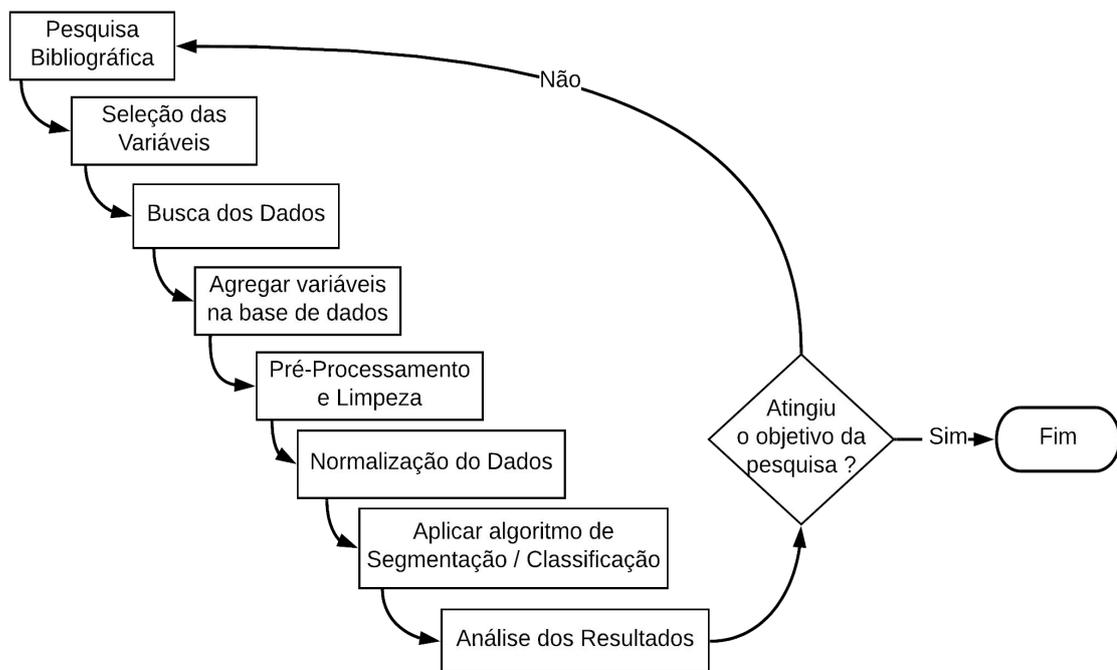
Os modelos de publicação estendida onde o pesquisador precisa enviar não só o artigo científico mas também os dados que permitem a validação da pesquisa em si tem entrado no rol de atividades essenciais de cientistas e pesquisadores. E a necessidade de depositar os dados em repositórios digitais não se limita apenas a fazer o *upload* de arquivos. O detalhamento de como foi realizado o processamento, a descrição dos metadados, da escolha dos formatos de arquivos, a preocupação em anonimizar dados que não podem se tornar públicos passam a ser uma tarefa extra para os pesquisadores.

O ponto de partida da presente pesquisa foi a montagem de um banco de dados a partir de várias bases de dados: a base de financiamento eleitoral, a base

com os resultados eleitorais, a base do CNPJ e a base de candidatos evangélicos que foi elaborada durante a pesquisa.

O processo como um todo pode ser visto na figura abaixo:

Figura 1 – Processo de Mineração de Dados Proposto



Fonte: Próprio autor

Vale ressaltar que o processo de “garimpagem” dos dados passa necessariamente pela experiência do pesquisador em selecionar quais as variáveis devem fazer parte do processamento. Ou seja, o simples fato de utilizar MD não é certeza de que o processo será bem-sucedido. Mancuso (2015) também alerta sobre a importância da seleção das variáveis, ao criticar que:

a inclusão pouco criteriosa de variáveis independentes nos modelos explicativos, sem justificativa teórica suficiente (King, Keohane & Verba 1994, p. 174), o que pode levar à identificação de correlações triviais ou espúrias. Para evitar o problema de *data mining* e, ao mesmo tempo, evitar a omissão de variáveis independentes relevantes, o desafio aqui é construir modelos mais completos que incluam, simultaneamente, os determinantes políticos, econômicos e sociais cuja importância tenha sido explicada por trabalhos anteriores e/ou que possam contar com justificativa teórica sólida (p.175)

Dessa forma, a análise da bibliografia sobre financiamento de campanha, dos resultados eleitorais e da participação de evangélicos na política partidária será um insumo fundamental para a definição de quais variáveis serão utilizadas no

processo. Além disso, a medida que o processamento for realizado, novas variáveis poderão ser adicionadas caso o resultado de tais processamentos não seja satisfatório.

Foram encontrados poucos trabalhos na área de Ciência Política que fizeram uso de técnicas de mineração de dados e apenas um único trabalho que detalhasse o uso de segmentação (técnica específica que será utilizada na pesquisa) foi encontrado. Filho et al (2014) sustentam que “a análise de segmentação é uma ferramenta importante para classificar unidades em grupos. Sua principal vantagem é produzir uma classificação objetiva e replicável que pode aprimorar o nosso conhecimento sobre fenômenos políticos”⁶.

Espera-se, portanto, que o uso de mineração de dados, aliado às teorias de Organização do Conhecimento da Ciência da Informação, possam formular novas metodologias que permitam que pesquisadores consigam recuperar, organizar e analisar grandes volume de dados, capacitando-os a gerar novos conhecimentos e teorias científicas.

1.5 ESTRUTURA DO TRABALHO

Para descrever a pesquisa da tese que busca encontrar novas classificações dos políticos evangélicos através do uso de técnicas de mineração de dados, o texto foi dividido em 7 capítulos. Considerando esta introdução como capítulo 1, teremos 4 capítulos de referencial teórico que embasarão a pesquisa empírica (capítulos 2, 3, 4 e 5), um capítulo específico que detalhará o tratamento de dados e os resultados obtidos e o capítulo final conterá as conclusões e propostas para trabalhos futuros.

No capítulo 2, foi realizada uma revisão da literatura sobre o que é MD e para tal, o conceito do termo “dado” será revisto em conjunto com a tríade dado-informação-conhecimento. Como na literatura da CI existem diversas interpretações e definições destes conceitos (CAPURRO e HJØRLAND, 2007), é importante delimitar qual o recorte teórico que será feito nesta tese, ou seja, qual definição será utilizada para que se possa realizar a conexão entre a MD, a ciência da informação e também ao conceito de dados abertos.

⁶ No original: “that cluster analysis is an important tool to classify units into groups. Its main advantage is to produce objective and replicable classification that can develop our knowledge regarding political phenomena”

O conceito de MD será revisto, não só sua história e atuação na ciência, mas também a metodologia e as estratégias utilizadas para sua implementação. Para cada estratégia de como utilizar as técnicas de mineração serão dados exemplos na área das ciências sociais para que seja possível se familiarizar com os termos empregados. Por fim, será feito um resumo mostrando com as técnicas de MD podem ser utilizadas na CI e vice-versa. Ou seja, a discussão dos usos das práticas da CI, com destaque para as teorias e métodos de classificação, pode ser de grande valia tanto para a MD como também para o campo mais amplo da Inteligência Artificial (IA)

Ainda no capítulo 2 serão mostrados exemplos de como a MD vem sendo utilizada no campo das ciências sociais aplicadas e em especial na Ciência da Informação e na Ciência Política. Por fim, como grande parte dos dados de pesquisa serão obtidos nos portais de dados abertos do Estado brasileiro, será feita uma breve análise do que são efetivamente dados abertos e a importância das políticas de dados abertos governamentais no Brasil e no mundo.

No capítulo 3, foi feito um levantamento do conceito de classificação e como os sistemas de organização do conhecimento podem ser utilizados para entender o processo social e político e como os marcos teóricos da classificação podem contribuir para o entendimento da classificação das religiões. Este capítulo foi construído baseado em artigo científico apresentado no congresso da ISKO 2020.

Em seguida, é feito um levantamento bibliográfico sobre a história da chegada dos evangélicos no Brasil e das principais denominações existentes atualmente. O protestantismo e o movimento evangélico foi analisado desde sua origem na Europa e como surgiu e cresceu no Brasil. O foco do levantamento será nas categorizações dos diversos grupos evangélicos na tentativa de construir uma relação entre as igrejas e grupos econômicos associados a elas.

O capítulo 4 trata especificamente da atuação dos evangélicos na política. Foi feito um levantamento de como se deu a entrada dos evangélicos na política, a criação e definição da chamada bancada evangélica nos últimos 20 anos e como a sua influência impacta o cenário político nacional.

O capítulo 5 busca descrever a importância do financiamento de campanhas política para as eleições no Brasil. Foi realizada uma revisão específica sobre as pesquisas de cunho empírico e quantitativo com o intuito de verificar quais as

principais variáveis utilizadas pelos pesquisadores para explicar o resultado eleitoral através do financiamento de campanha. O levantamento tomará como base o artigo do pesquisador Wagner Mancuso sob o título “Investimento eleitoral no Brasil: balanço da literatura (2001-2012)” mas também buscou pesquisas mais recentes sobre o tema.

Neste capítulo, também há o detalhamento de como está organizada a base de dados fornecida pelo TSE e quais os critérios utilizados para “enriquecer” os dados originais com os resultados eleitorais e outros dados que possam descrever o fenômeno eleitoral. Por exemplo, a base de financiamento fornecida pelo TSE traz o CNPJ da empresa que doou para um determinado partido ou candidato, entretanto quando a empresa doa diretamente para o partido, a contribuição fica diluída entre os candidatos e não é possível identificar quanto vai para cada candidato. É possível calcular quais os candidatos que mais se beneficiaram da verba recebida pelo partido?

Além disso, para aumentar a complexidade das análises que foram realizadas, a mini reforma eleitoral de 2016 trouxe um complicador a mais. A nova lei determina que somente pessoas físicas podem fazer doações e dessa forma, o método direto de encontrar os setores econômicos que efetivamente financiavam um determinado partido ou político não seria mais possível. Entretanto, era sabido que já nas eleições de 2016, empresários continuavam a fazer vultosas doações para as campanhas políticas.

Por fim, o capítulo 6 descreve como o processamento foi realizado e quais os resultados encontrados e o capítulo 7 traz as conclusões da pesquisa e sugestões para pesquisas futuras que podem advir desta tese.

2 DADOS, MINERAÇÃO DE DADOS E A CIÊNCIA DA INFORMAÇÃO

2.1 O CONCEITO DE DADO, INFORMAÇÃO E CONHECIMENTO

Para que se possa entender o conceito de Mineração de Dados é importante que se defina primeiramente o conceito do que é um Dado. Tradicionalmente, o termo “dado” é considerado como a base da hierarquia DIKW (*Data, Information, Knowledge and Wisdom*) onde informação é entendida como dado processado, conhecimento como informação processada ou interpretada e sabedoria como conhecimento interpretado (Hjørland, 2018). Entretanto, estas visões hierarquizadas trazem muitas questões conceituais e principalmente os termos conhecimento (Knowledge) e sabedoria (Wisdom) abrem um leque muito grande de interpretações sobre qual seria a diferença entre sabedoria e conhecimento e talvez esta hierarquia mais atrapalhe do que facilite o entendimento.

A tríade Dado, Informação e Conhecimento já foi amplamente discutida na CI apesar de não existir um consenso entre os principais autores da área. Buckland (1991) procurou trazer o aspecto da materialidade da Informação (*Information as Thing*), bem como da informação como processo e também como fenômeno cognitivo. Capurro e Hjørland (2007) discutiram sobre o conceito de Informação e seu caráter polissêmico e Furner (2016) realizou uma extensa pesquisa não só sobre a história e etimologia do termo dado (*data* em inglês), mas trouxe também os vários usos do termo em diferentes áreas do conhecimento.

Desta forma, um dos objetivos iniciais é propor uma estrutura teórica que traga uma definição o mais objetiva possível do que é um “dado” para, em seguida, se possa adentrar nos aspectos da Mineração de Dados e da sua interligação com a Ciência da Informação. Ou seja, não se busca aqui listar todas as definições existentes mas, ao contrário, delimitar o termo para que se possa construir um raciocínio sobre o conceito de mineração de dados.

Diferentemente dos dois textos de Furner (2016) e de Capurro e Hjørland (2007) onde os autores detalham a formação dos termos “dado” e “informação” através da história e de sua etimologia e, tal como em um livro de suspense, vão construindo uma narrativa onde somente no final da “trama”, no clímax dos respectivos artigos, trazem algumas possibilidades para a definição dos dois conceitos, aqui se fará o caminho inverso. A inversão tem uma intenção didática

para que se possa, de antemão, conhecer as definições propostas para então seguir o raciocínio utilizado nesta proposição teórica. Vale ressaltar que, obviamente, esta definição não tem a pretensão de ser definitiva. Pelo contrário, o objetivo é justamente suscitar o debate, ou seja, trazer novas reflexões que possam enriquecer o entendimento do assunto e delimitar o termo para o objetivo central da tese. Isso posto, a definição candidata dentro do modelo teórico aqui proposto para o termo dado é: **Dado é o registro do atributo de um ente, objeto ou fenômeno.**

Registro é o ato de registrar, ou seja, é a gravação ou a impressão de caracteres, símbolos que tenham um significado em algum documento ou suporte físico. A temperatura de uma lâmina de aço só se transforma em dado, quando ela é medida e registrada, a cor de uma planta só passa a ser um dado quando ela é vista e registrada seja através de uma fotografia, seja pela pintura feita à mão ou registrada digitalmente através de um espectrômetro. Assim, se a temperatura ou a cor forem apenas medidas mas não houver o registro, esses atributos não podem ser considerados um dado.

Propositalmente, o termo “registro” foi escolhido justamente para trazer a ideia de documento, ou seja, não há dado sem documento. Se não houver o registro, se não houver a medição, se não houver a observação, não existe dado. Desta forma, o termo registro está intimamente relacionado ao ato de documentar e organizar pois para que haja o registro, deve-se criar minimamente um conjunto de regras para que tal registro fique ordenado para que futuramente possa ser recuperado.

Já o termo atributo se refere a uma característica (uma propriedade) de um ente, objeto ou fenômeno. O termo “ente” foi uma opção preferencial pois englobaria tanto os seres vivos como os inanimados e também o que se encontra no universo da imaginação. Assim, recairia no conjunto dos entes, uma estrela, um elefante, um micróbio, um martelo, uma poesia ou um personagem de ficção. Ou a lista de atributos de um personagem de Machado de Assis não podem ser considerados como dados de tal personagem?

Cogitou-se não utilizar o termo “objeto” já que todo objeto é um ente. Entretanto, optou-se por utilizá-lo já que o termo é utilizado por vários autores de referência na CI (DAHLBERG, 1978; BRÄSCHER, 2014). Por fim, optou-se também por incluir o termo fenômeno para incluir o registro de um fato, de algo que ocorreu em um determinado intervalo de tempo, e que efetivamente não está ligado a um

objeto ou ente específico. Como exemplo, pode-se citar os fenômenos meteorológicos, a fala de um indivíduo ou a passagem de um cometa.

A definição de “dados” utilizada aqui tem semelhança com definições propostas por alguns autores. Christine Borgman (2016) reforça que não existe uma definição única para dados mas também procura propor uma definição que facilite o entendimento do termo. Para a autora, um dado é uma “entidade utilizada como evidência de um fenômeno para propósitos de pesquisa ou ensino” (Borgman, 2016, p.29, tradução nossa). Esta definição fica bem próxima da nossa pois reforça o conceito de registro através do uso do termo “evidência”. Entretanto, o uso do termo “entidade” não consegue retratar bem o que efetivamente está sendo registrado.

Já Semeler e Pinto definem dado como “todo objeto criado em formato digital ou convertido para o formato digital que possa ser usado para geração de *insights* de informação e conhecimento” (2019, p.115). Aqui a delimitação de que o dado só existe digitalmente diminui a universalidade do termo já que até hoje a sociedade continua criando dados não digitais. Fica também em aberto o que seria efetivamente um “insight”.

Dessa forma, tendo uma definição candidata para “dado”, passa-se a uma outra fase tão ou mais importante. Ora, se houve o registro do atributo de um objeto, como ele se deu? Em que suporte ele foi registrado? Quem fez o registro? O que significa esse atributo que está sendo registrado? Ou seja, por trás de todo o dado, existe sempre um conjunto de metadados. Dentro desta construção conceitual, os metadados são os dados que detalham como o dado foi registrado. **É o dado sobre o dado.**

O primeiro uso do termo “metadados” ocorreu em janeiro de 1967 por Stuart McIntosh e David Griffel que utilizaram o prefixo “meta” (do grego τὰ μετὰ) para tentar sintetizar o conceito “sobre os dados” e provavelmente a coleção de ensaios sobre metafísica de Aristóteles serviu de inspiração para a confecção do termo (STEINER, 2017). De fato, o conceito por trás da definição do termo pode ser rastreado desde a antiguidade: uma das primeiras formas de metadados é a obra de Pinakes, criada em 245 a.C. por Kallimachos de Cirene para catalogar a Biblioteca de Alexandria de forma sistematizada (POMERANTZ, 2015, p.6), ou seja, o termo metadado está intimamente associado à Biblioteconomia e por consequência à Ciência da Informação.

Nota-se que o uso do termo metadado associado ao documento bibliográfico difere da definição inicial já que, nesse contexto, a definição seria “os dados sobre o documento” e não a definição inicial “os dados sobre o dado”. Entretanto, se o dado (ou o conjunto de dados) for considerado um tipo de documento, as duas definições não se contradizem. E realmente, a partir da ideia inicial de que o dado pressupõe um registro, não seria estranho pensar que o dado pode ser considerado um tipo de documento. Furner corrobora essa relação entre documento e os dados:

De fato, não é verdade que os documentos sejam constituídos por dados, nem que o documento seja uma espécie de conjunto de dados: é o contrário, em ambos os aspectos. Um conjunto de dados é constituído por documentos; e o conjunto de dados é uma espécie de documento (p. 303, tradução nossa)⁷

Outras definições de metadado já incorporam o termo “informação” criando ainda mais ruído na tentativa de se encontrar uma definição mais precisa. Por exemplo, o *National Information Standards Organization* (organização norte-americana que organiza padrões informacionais naquele país) define metadado como “informação estruturada que descreve, explica, localiza ou facilita a recuperação, utilização ou gestão de um recurso de informação” (NISO, 2004).

Henning Spang-Hanssen (2001) corrobora a questão da necessidade de como o dado foi coletado, apesar de não utilizar o termo “metadado”:

A informação sobre as propriedades físicas de um material está realmente incompleta sem informação sobre a precisão dos dados e sobre as condições em que esses dados foram obtidos. Além disso, várias investigações de uma propriedade conduziram frequentemente a resultados diferentes que não podem ser comparados e avaliados para além das informações sobre os seus antecedentes.⁸ (p.3, tradução nossa)

A complexidade deste modelo teórico pode ainda aumentar um pouco com a seguinte reflexão: para todo dado, pode-se considerar que existem dois tipos de metadados que o acompanham: metadados explícitos e implícitos. Metadados explícitos são os metadados registrados em conjunto com o dado e os metadados implícitos são sabidos a partir de sua descrição ou de outros metadados explícitos.

⁷ No original “it is not in fact the case that documents are made up of data, nor that the document is a species of dataset: rather it is the other way around, in both respects. A dataset is made up of documents; and the dataset is a species of document”.

⁸ No original “Information about some physical property of a material is actually incomplete without information about the precision of the data and about the conditions under which these data were obtained. Moreover, various investigations of a property have often led to different results that cannot be compared and evaluated apart from information about their background”

Voltando ao exemplo da temperatura da chapa de aço, o dado é o número medido e registrado e os metadados explícitos são a indicação do instante em que a medição foi feita, do tipo de material (no caso o aço), do equipamento que realizou a medição e da unidade de medida utilizada. O próprio descritor “temperatura da chapa de aço” é um metadado. Entretanto, não se registra o conceito do que é uma temperatura, pois o produtor deste dado parte do pressuposto que o usuário ou leitor já conheça esse conceito. Ou seja, alguns metadados deste documento já são implícitos ao documento onde o dado foi registrado. São implícitos inclusive os metadados registrados, ou seja, está implícito para o usuário o que representa a unidade de medida de temperatura, as propriedades do instrumento de medição e uma infinidade de conceitos que não foram registrados pois se pressupõe que o leitor não necessite saber pois é um dado (ou melhor, um metadado) conhecido.

Desta forma, pode-se estender um pouco mais esse modelo teórico: quando se tem o dado e seus metadados, o usuário deste mesmo dado consegue ter uma compreensão maior do que aquela variável representa e do contexto em que ela foi registrada, ou seja, o conjunto (dado + metadados explícitos e implícitos) pode ser uma possível definição para informação. E esta definição não é meramente formal. Tal definição busca resgatar o aspecto cognitivo da informação que é contemplado pelos metadados implícitos que compõem a fórmula dado + metadado, isto é, da relevância do contexto e da percepção do usuário que estará recuperando esse dado. A análise ou percepção do conjunto de metadados implícitos é tão ou mais marcante que o aspecto físico ou documental.

Voltando ao exemplo da lâmina de aço, um leitor só conseguirá interpretar o dado da temperatura se ele souber o que significa uma temperatura. Ou seja, o dado foi registrado corretamente, o contexto em que o dado foi coletado pode ter sido fornecido mas se o leitor não tiver o conhecimento prévio dos termos empregados na definição dos metadados apresentados, o dado não se transforma em conhecimento para este mesmo leitor. Conforme atestam Marteleto, Nóbrega e Morado (2013, p. 100), o “conhecimento que resulta da ação e das relações dos sujeitos em seu fazer na história, em suas práticas culturais, é sempre um conhecimento contextual”.

O que está sendo chamado aqui de metadado implícito é efetivamente o que Dahlberg define como conceito em sua Teoria do Conceito (DAHLBERG, 1978). E

esta teoria procura tratar toda a problemática do que é um conceito, o que é um termo, o que é uma definição (a definição do que é uma definição), do que é a representação do conhecimento e dos aspectos semióticos que são intrínsecos a uma visão cognitivista e conceitualista.

Ainda segundo a autora, o conceito é um dos elementos centrais para a informação e para a geração do conhecimento tendo uma dimensão física e uma dimensão cognitiva. Capurro joga com as palavras quando afirma que, enquanto “information is a thing” (citando Buckland), “knowledge is no-thing”, ou seja, não existe materialidade no conhecimento (Capurro, 2017). E, por não ser material, a definição do que é conhecimento enfrenta grande resistência em áreas como a Ciência da Computação:

a visão “conceitualista”, apesar de servir a diversos propósitos na CI como, por exemplo, em instrumentos como os vocabulários controlados, não é adequada à realidade atual, caracterizada pelo uso de computadores. Sistemas automatizados e computadores se tornam elementos obrigatórios ante o grande volume de informação disponível. E computadores necessitam de representações formais (Souza, Almeida e Barracho, 2013 p.164)

E a construção destas representações formais se traduz justamente na transformação de conceitos (unidades de conhecimento) em metadados explícitos. Esta é a grande tarefa da elaboração de instrumentos de construção semântica e de ontologias para que as tecnologias de aprendizado de máquina e IA possam funcionar. Ou seja, quanto mais dados, quanto mais atributos, quanto mais metadados estiverem disponíveis para um objeto, melhor o algoritmo poderá “entender” e “aprender” de forma autônoma.

O grande esforço de construir uma Web Semântica (BERNERS-LEE; HENDLER; LASSILA, 2001), também denominada Web 3.0, é um exemplo claro da importância de construir insumos para que os algoritmos tenham a possibilidade de aprender. Viana (2018) reforça este argumento quando ressalta que o termo “semântica” visa “estabelecer uma interligação dos significados das palavras, o que as possibilita de serem compreendidas por computadores” e complementa que a Web 3.0 pretende ser um espaço digital “desambiguado” pois as máquinas poderão “entender” o significado do conteúdo de um site ou texto sem margem para dúvida na interpretação do que está sendo apresentado.

O termo Web 3.0 já vem sendo utilizado há quase duas décadas quando foi utilizado na obra de Tim Berners-Lee em 2001. Entretanto, percebe-se que o desenvolvimento da web semântica não atingiu seus objetivos com a mesma velocidade que as demais tecnologias da Internet (MICHALAKIDIS, 2016). Apesar dos inúmeros avanços tecnológicos como a criação de padrões e boas práticas com o uso de Linked Data, RDF e SPARQL, o ser humano continua sendo peça essencial para a criação da camada extra de dados (e metadados) que devem ser inseridas nas páginas web para que a máquina consiga interpretá-las. A inclusão automatizada de artefatos semânticos nas páginas web ainda não é possível e um dos motivos é a dificuldade em representar o conhecimento.

A Representação do Conhecimento (Knowledge Representation) é a disciplina que tem como objetivo a captura e modelagem de representações sobre o mundo que possam ser interpretadas por algoritmos para que estes possam resolver problemas complexos. E esta área tem uma extensa fronteira interdisciplinar com a CI, Linguística, Filosofia, Psicologia e Inteligência Artificial (BRÄSCHER, 2014).

A pesquisadora Marisa Bräscher (2014) ressalta ainda que o conhecimento acumulado na construção de sistemas de indexação e de organização do conhecimento como os tesouros, ontologias e taxonomias podem contribuir teoricamente para a superação das dificuldades enfrentadas na área de representação do conhecimento e, conseqüentemente para áreas como a IA e a MD. Na seção seguinte, a interdisciplinaridade das duas ciências será ainda mais ressaltada.

2.2 MINERAÇÃO DE DADOS

Se buscarmos na história da ciência da computação, as técnicas de mineração de dados vêm sendo utilizadas desde 1960, mas somente a partir da década de 80 que o termo passou a ser empregado para consolidar um conjunto de processos e algoritmos que tinham um mesmo objetivo:

[...] a extração automatizada de padrões que representem algum conhecimento implícito que esteja armazenado em grandes volumes de dados, sejam eles armazenados em banco de dados, data warehouses, na web ou qualquer outro tipo de repositório massivo ou fluxo de dados.⁹ (Han, Kamber e Pei, 2012, p. 23 do prefácio, tradução nossa).

Fayyad, Piatetsky-Shapiro e Smyth (1996) consideram que a mineração de dados é parte de um processo maior denominado *Knowledge Discovery on Databases* (KDD), que envolve a coleta, a limpeza dos dados, o processamento em si e a análise e visualização dos resultados encontrados. Entretanto considera-se aqui que a escolha do termo KDD é equivocada por diversas razões. Como primeiro ponto, se utiliza mineração de dados em repositórios que não estão necessariamente em banco de dados (*databases*), podendo estar em coleção de textos, em imagens e até mesmo em fluxos de dados enviados em tempo real por dispositivos como medidores de temperatura e de posicionamento georreferenciado.

Mesmo se o termo *databases* fosse substituído por *data* (dado), alguns autores ressaltam que a etapa de pré-processamento e limpeza de dados é intrínseca à mineração, ou seja, não existe mineração sem modelagem e pré-processamento dos dados (Hand, Mannila e Smyth, 2001). E, por fim, sob o ponto de vista epistemológico, o termo KDD é muito amplo e pode se confundir com a própria definição do que é Ciência. Afinal, o que é a ciência senão a busca de conhecimento através dos dados empíricos coletados?

Vale notar que a obra de Han e Kamber vai mudando a sua percepção entre os termos KDD e Data Mining ao longo de suas três edições. Em sua primeira edição (2001), os autores seguem a definição de Fayyad. Já na segunda edição (2006), eles já concordam que o Data Mining é considerado pela academia e pelo mercado como o processo integral e por fim, na terceira edição (2012), citam apenas a perspectiva de Fayyad como uma forma de ver o processo, mas que a grande maioria dos autores já colocam os dois termos como sinônimos, sendo que a preferência é pelo termo Data Mining.

O conceito KDD tampouco é compatível com o modelo teórico proposto no início deste ensaio já que conhecimento não pode ser “descoberto”, ele é transmitido e adquirido pelo ser humano através de processos de aprendizado que são alheios a

⁹ Do original “the automated or convenient extraction of patterns representing knowledge implicitly stored or captured in large databases, data warehouses, the Web, other massive information repositories, or data streams.”

técnicas de mineração de dados. Assim, o presente texto segue o entendimento de outros autores (COENEN, 2011; HAN, KAMBER e PEI, 2012; TRYBULA, 1999) e utiliza a expressão “Mineração de Dados” como mais apropriada que KDD, podendo ser utilizada para identificar todo processo que envolve a descoberta de padrões a partir de grandes volumes de dados.

O estudo dos algoritmos e usos de mineração de dados se originou na Ciência da Computação e recentemente passou a fazer parte de um novo campo se convencionou chamar de Ciência de Dados (*Data Science*). Segundo Carvalho (2016), a Ciência de Dados

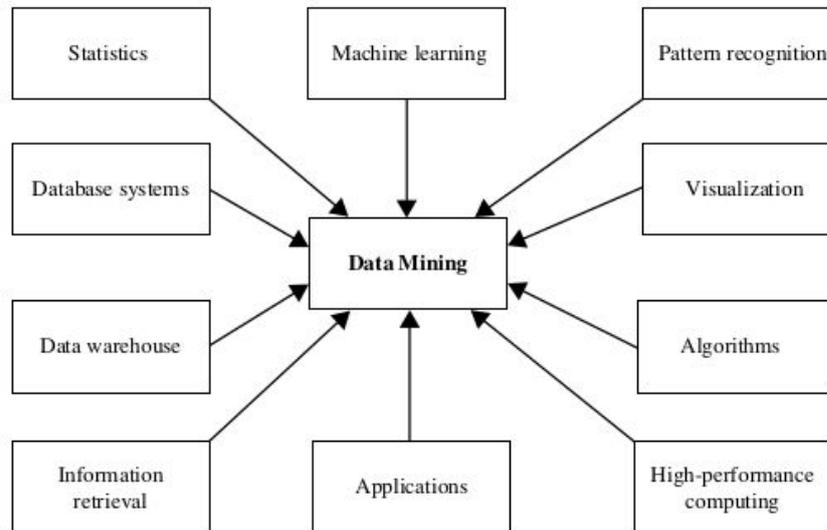
estuda princípios, métodos e sistemas computacionais capazes de extrair de forma eficiente conhecimento novo, útil e relevante presente em conjuntos de dados. Para isso, ela faz uso de técnicas de mineração de dados, particularmente de construção automática de modelos, capazes de extrair esse conhecimento. A construção automática de modelos permite que funções, hipóteses e regras sejam extraídas a partir de experiências passadas, representadas no conjunto de dados (Carvalho, 2016, p.63)

Na prática, a chamada Ciência de Dados vem sendo disputada também pelos cursos de Ciência da Informação, Ciências Estatísticas e Administração. No exterior, a maior parte dos cursos que tem o termo “*Data Science*” no título estão sendo organizados pelos departamentos de Ciência da Computação. Já no Brasil, o primeiro curso de graduação de Ciência de Dados, criado em 2020, está sob a égide do Departamento de Matemática Aplicada e Estatística da USP/São Carlos. Como o uso do termo Ciência de Dados ainda é muito recente, seria de bom tom que fosse utilizado com cautela já que existe muito ruído entre o que está no âmbito da ciência e o que é apenas marketing para que a indústria de TI possa vender novos softwares e novos cursos.

Outro termo que também merece desambiguação é o “Aprendizado por máquina” (*Machine Learning*) que algumas vezes é confundido com o processo de Mineração de Dados (LESKOVEC; RAJARAMAN; ULLMAN, 2014, p.20). De fato, a mineração de dados frequentemente utiliza algoritmos de aprendizado por máquina para resolver determinadas etapas do processo de mineração e nem todo algoritmo de aprendizado por máquina é utilizado no processo de mineração de dados. Resumidamente, o aprendizado por máquina é uma das tecnologias utilizadas na mineração de dados (HAN, KAMBER e PEI, 2012).

A figura abaixo mostra melhor essa relação:

Figura 2 – Tecnologias utilizadas pela Mineração de Dados



Fonte: Han, Kamber e Pei, 2011

O processo de Mineração de Dados pode ser dividido em três grandes atividades: pré-processamento dos dados, processamento de algoritmos para gerar modelos e padrões e por fim a avaliação dos modelos e padrões extraídos. A segunda etapa do processo pode utilizar de diversas estratégias para em função do tipo de dado e do tipo de análise que se quer obter. Estas diversas estratégias de mineração de dados podem ser categorizadas em dois grandes tipo em função dos processos que são empregados: descoberta supervisionada e não supervisionada.

A descoberta supervisionada requer inicialmente que seja realizada uma etapa de treinamento na qual parte dos objetos serão pré-classificados a priori. Seleciona-se um subconjunto de objetos que serão estudados e associa-se um termo único que identifique a classe à qual o item pertence. Dessa forma, o algoritmo terá uma base inicial de “conhecimento” e a partir daí o algoritmo “constrói” um modelo computacional se baseando no conjunto de objetos iniciais. Por fim, quando um novo elemento (do qual não se sabe a qual classe pertence) for inserido no sistema, o algoritmo indicará a qual grupo ele provavelmente pertencerá.

Para tornar o método mais claro, vale um exemplo na ciência política. Seleciona-se um grupo aleatório de parlamentares e cada um recebe um rótulo indicando se ele é de esquerda, de direita ou de centro. Ou seja, o ser humano informa ao algoritmo qual o grupo (segmento) do qual cada parlamentar faz parte e

indica também suas inúmeras características, como sexo, profissão, idade, partido político, votações anteriores etc. A partir destes dados, o algoritmo constrói um modelo e quando um novo candidato for inserido no sistema, o algoritmo indicará a qual grupo ele pertence.

Modelos supervisionados geralmente tem o objetivo de previsão. E no caso do uso de mineração de dados na pesquisa científica, estes algoritmos são muito úteis na fase de geração de hipóteses. Um estudo que tenha como objetivo prever em vez de explicar pode revelar associações entre variáveis que formam a base para o desenvolvimento de teorias e que podem ser posteriormente submetidas a outros tipos de testes empíricos mais rigorosos (SHMUELI, 2010).

Já as técnicas ditas “não supervisionadas” procuram categorizar os objetos que fazem parte de fenômeno que se quer estudar. Retomando o exemplo dos políticos, imagine que temos apenas as características dos políticos (dados de entrada) e nenhuma categorização aparente (dado de saída), ou seja, não temos um modelo ou categorização prévia para aplicar ao nosso conjunto de dados. Como podemos aprender mais sobre esse conjunto de dados? No nosso exemplo, qual seria a melhor forma de dividir os políticos em categorias? Será que existe apenas esquerda, direita e centro, ou há alguma outra categorização que permita distinguir os políticos, mas que não esteja aparente? O nome desse processo de classificação ou categorização é chamado de segmentação (clustering) e se denomina não supervisionado pois não há uma supervisão humana para determinar quais ou quantas categorias existirão para agrupar elementos de um grupo. Estes agrupamentos são denominados clusters, partições ou segmentos. O modelo teórico que fundamenta esse processamento é baseado no princípio de maximizar as similaridades intra-classes e minimizar a similaridade inter-classes (HAN, KAMBER, 2001).

Ainda dentro dos algoritmos não supervisionados, destaca-se a detecção de anomalias (outliers) ou, em jargão popular, os pontos “fora da curva”. A detecção de anomalias e os processos de segmentação são duas tarefas altamente relacionadas. O agrupamento encontra os padrões de maioria em um conjunto de dados e organiza os dados em função da similaridade de suas características, enquanto a detecção de outliers tenta capturar os casos excepcionais que se desviam substancialmente dos padrões da maioria (HAN, KAMBER e PEI, 2012).

É importante também ressaltar um outro recurso que pode ser utilizado tanto nas estratégias supervisionadas quanto nas não-supervisionadas: a visualização de dados. A construção de visualizações planas ou espaciais dos dados procura aproveitar as capacidades cognitivas dos seres humanos para favorecer percepções visuais sobre os dados que muitas vezes os algoritmos de mineração de dados não conseguem fornecer. Entretanto, o processo de montar visualizações em grandes bases não é trivial, pois é necessário encontrar uma visão ou projeção que reduza a complexidade enquanto captura informações importantes. O objetivo é reduzir a complexidade ao mesmo tempo em que se tenta perder a menor quantidade de informação possível (FAYYAD, 2001).

Também é possível categorizar a mineração de dados em função dos tipos de objetos e suportes onde o dado está representado. Além de minerações em *datasets* e tabelas relacionais, utiliza-se a mineração de dados em textos, na web, em imagens e vídeos e em grafos.

A Mineração de Textos (*Text Mining*) procura categorizar ou encontrar padrões em grandes coleções de documentos como notícias de jornal, páginas na Web, artigos de revistas científicas e obras de um determinado autor ou período (PINTO, 2007). Também é muito utilizado na mineração de opiniões ou análise de sentimentos (SOUZA e CAFÉ, 2018), quando se procura encontrar padrões em conversações realizadas nas redes sociais ou formulários online sobre algum tema específico.

A mineração na Web é uma categoria especial de mineração de texto pois necessita de um passo extra (denominado Web Scraping ou Crawling) para transformar os dados não estruturados de uma página web em um dado estruturado que possa ser minerado. Esse passo extra é quase sempre manual (realizado por humanos) se a página web não contém recursos semânticos como RDF ou outros tipos de metadados. Importante ressaltar que o simples processo de captura de dados na web não pode ser chamado de mineração de dados, ou seja, se a captura não vem acompanhada de um processo algorítmico que extraia alguma informação nova, ela é apenas uma captura.

Ainda na área de mineração de texto, uma técnica que vem sendo muito estudada dentro da Ciência da Informação é a Modelagem de Tópicos de

Conhecimento (*Topic Modeling*). Muitas vezes há a necessidade de categorizar coleções de documentos por assunto e para tal, existem métodos de classificação não-supervisionada, semelhante ao agrupamento em dados numéricos, que procura descobrir qual categoria (tópico de conhecimento) cada um dos elementos pertence.

Já a Mineração de Imagens analisa imagens ou vídeos basicamente para categorizá-las a partir de critérios pré-definidos ou para se encontrar padrões específicos em um grande volume de imagens, como, por exemplo, imagens de satélite ou recuperadas por telescópios. Aqui entram técnicas complexas de aprendizado de máquina, redes neurais (*neural networks*) e as chamadas *deep neural networks*.

Por fim, a Mineração de Grafos é o estudo de grafos gerados a partir do relacionamento entre entidades ou pessoas, como o relacionamento de pessoas nas redes sociais, operações financeiras ou citações bibliográficas. Apesar do uso de grafos na análise bibliométrica já ser uma prática comum na CI, novas técnicas vêm sendo empregadas na área da Economia Política, para encontrar, por exemplo, relações entre políticos e financiadores de campanhas eleitorais (JUNCKES et al, 2019). Outras relações mais complexas também são desenvolvidas, como doadores que tenham sido denunciados por prática de trabalho escravo¹⁰.

Ao analisar cada umas tipologias e estratégias de mineração de dados, percebe-se um elemento comum: a classificação. O objetivo final recai sempre no classificar, ou seja, buscar dentro do universo pesquisado um padrão que permita unir (ou segregar) os entes, objetos ou fenômenos que estão sendo estudados. Pode-se dizer que ao final de cada processo de mineração, emerge pelo menos um novo dado, um novo atributo do objeto que está sendo analisado. A metáfora que pode ser feita é comparar o processo de mineração tradicional (no mundo material) que tem como objetivo encontrar um metal raro (o ferro, o ouro ou uma pedra preciosa) em um grande volume de rocha bruta e, no caso da mineração de dados (no mundo informacional), o objetivo é encontrar um novo dado que não estava especificado dentro de um grande volume de dados.

¹⁰ O Ruralômetro (<https://ruralometro.reporterbrasil.org.br/>) é um exemplo de banco de dados que contém uma ferramenta interativa que mede como os deputados federais eleitos em 2014 agiram frente a projetos de lei e medidas provisórias que têm impactos sobre o meio ambiente, os povos indígenas e trabalhadores rurais.

E as classificações que emergem do processo de mineração nem sempre são triviais e quanto mais complexas forem, mais necessitarão de sistemas que organizem os resultados obtidos no processo. A partir deste raciocínio, pode-se supor que o uso de todo o conhecimento teórico acumulado na área de Organização do Conhecimento pode trazer benefícios também para essa etapa de análise dos resultados obtidos no processo de mineração de dados.

Em 1995, ao discorrer sobre a interdisciplinaridade entre Ciência da Informação e Ciência da Computação, Saracevic afirmou que a Ciência da Computação era o braço tecnológico da Ciência da Informação. Seria a Ciência da Informação o braço de construção semântica para a Ciência da Computação e da Ciência de Dados?

2.3 MINERAÇÃO DE DADOS NA CIÊNCIA DA INFORMAÇÃO

O uso de técnicas de mineração de dados (MD) tem sido crescente em pesquisas na área de ciências sociais e nas chamadas humanidades. Já em 2004, Chen e Liu detalham o uso de técnicas de MD na Ciência da Informação (CHEN e LIU, 2004). Só a revista científica especializada em Mineração de Dados para Humanidades Digitais “The Journal of Data Mining & Digital Humanities” já conta com mais de 58 artigos publicados e a partir de pesquisa bibliográfica preliminar na base Web of Science utilizando o termo “Data Mining”, foram encontrados 79 artigos na área de Sociologia e Ciência Política.

Para tentar mapear de forma ampla o que se tem estudado no Brasil sobre MD na CI, foram realizadas buscas com os termos “mineração”, “mining”, “machine” na BDTD (Biblioteca Digital de Teses e Dissertações), na BRAPCI (que agrega parte das revistas específicas de Ciência da Informação) e na BENANCIB (que agrega todos os trabalhos apresentados no ENANCIB) sendo que na BDTD foi aplicado um filtro extra para trazer registros apenas na área das Ciências Sociais. Por fim, nas 3 bases, aplicou-se um filtro extra para trazer documentos somente a partir de 2010 (inclusive).

A partir dos resultados iniciais, fez-se uma análise de cada um dos artigos para verificar qual tipo de técnica ou algoritmo de MD foi aplicado ou se o artigo relatava uma análise teórica ou descritiva sobre o tema. O objetivo desta análise era entender quais as técnicas que eram mais utilizadas na CI. A busca pelo banco de

teses e dissertações trouxe 69 teses na Grande Área de Ciências Sociais Aplicadas sendo 24 na área de Ciência da Informação (8 teses e 13 dissertações). Deste total, 10 trabalhos foram sobre Mineração de Textos, ou seja, quase a metade. A segunda maior frequência foi a de análises descritivas sobre o processo de mineração de dados e a sua interdisciplinaridade com a CI.

Já na base BRAPCI, foram encontrados 62 artigos entretanto 20 foram descartados pois continham os termos de busca no assunto, título ou palavra-chave mas efetivamente não faziam uso de nenhuma técnica de MD ou tampouco faziam alguma análise crítica sobre o uso de tais técnicas. O assunto que predominou foi também o de Mineração de Texto tal como podemos observar na tabela abaixo:

Tabela 1 - Total de Artigos na BRAPCI de 2010 a 2018 sobre Mineração de Dados

Categoria	Número de Artigos
Análise de Sentimentos	1
Árvores de decisão	1
Classificação	2
Descritivo	11
Regras de associação	4
Segmentação	3
Text Mining	14
Topic Modeling	3
Visualização	3
Total de Artigos	42

Fonte: Os autores

Por fim, na base do BENANCIB foram encontrados apenas 14 artigos sendo que em 2018 foram publicados 6 artigos, ou seja quase metade do total de artigos analisados. Da mesma forma que nas outras bases a preponderância se deu nas técnicas de Mineração de Texto, 6 dos 14 artigos utilizaram tais técnicas em suas pesquisas.

2.4 DADOS ABERTOS

Ao largo de todo o estudo epistemológico que foi levantado nesse capítulo, existe uma questão central que perpassa todos os fenômenos que foram enunciados: a circulação de informação, a criação de padrões computacionais onde cada bit que circula pela Internet e o próprio fazer da ciência são regidos por um

regime econômico e social capitalista que tem em seu cerne a propriedade privada e o controle dos meios de produção.

No livro “A Sociedade do Espetáculo”, Guy Debord procura sintetizar o momento histórico em que vivemos, quando se consolida a separação dos meios de produção do trabalhador e praticamente todos os aspectos da cultura e experiência são intermediados por uma relação social capitalista. O autor francês traz logo em sua primeira tese a concepção de que esta é global quando parafraseia Marx: “Toda a vida das sociedades nas quais reinam as condições modernas de produção se anuncia como uma imensa acumulação de espetáculos.” (Debord, 2003, Tese 1).

Jappe sintetiza o conceito de Espetáculo nos seguintes termos:

Existem dois fundamentos principais para o espetáculo: a renovação tecnológica incessante e a fusão econômico-estatal e, na sua fase mais recente, três grandes consequências: o segredo generalizado; o falso sem réplica e um presente perpétuo. O espetáculo não é, pois, uma pura e simples adjunção ao mundo [...] Na realidade a atividade social inteira é captada pelo espetáculo para servir os seus próprios fins. Do urbanismo aos partidos políticos de todas as tendências, da arte às ciências, da vida quotidiana às paixões e aos desejos humanos, em toda a parte a realidade foi substituída pela sua imagem. (JAPPE, 2008, p. 18)

E a circulação da informação também está inserida nesse processo global de acumulação de capital e de espetáculos. As plataformas sócio-digitais¹¹ (DANTAS, 2019) que controlam grande parte do tráfego da Internet são os grandes agentes e operadores deste fluxo. Sandra Braman trabalha com o conceito de Regime Global de Informação onde ressalta que a informação é uma das ferramentas de poder do regime global vigente e que tanto os governos como as empresas e a sociedade civil cooperam para a sua manutenção. E segundo a autora canadense, a transparência é um dos instrumentos políticos do regime global de informação:

¹¹ Dantas denomina de plataformas sócio-digitais todas as grandes corporações como Alfabet/Google, Facebook, Amazon, Apple, Ali-baba, etc. Fazem parte de uma mesma categoria de empresas que prestam serviços na Internet que tem um modelo de negócio comum: conectar compradores e vendedores de bens e serviços e que grande parte de sua receita vem da monetização dos dados que circulam por suas redes.

ela [a transparência] começou como um elemento do regime de segurança, foi retomada no comércio e agora é amplamente utilizada em todos os setores. A transparência substituiu as noções de livre circulação da informação em todas as suas variações como o ideal para os fluxos internacionais de informação e tornou-se agora um objetivo político por direito próprio. Esta transformação, por sua vez, muda as relações dos indivíduos com a sociedade e das sociedades entre si. (BRAMAN, 2004, p.32, tradução nossa)¹²

O que Braman quer ressaltar é que a transparência serve antes de tudo ao regime global de informação. Quanto mais informações públicas e governamentais estiverem disponíveis, menor será o custo para utilizá-las. É óbvio que a sociedade como um todo pode utilizar estes dados, mas como se verá a frente, os maiores beneficiários dos dados abertos são as empresas que agregam estes dados às suas próprias operações.

É nesse contexto de Transparência que o conceito de Dados Abertos se insere. Dados Abertos são definidos como dados que podem ser utilizados livremente por qualquer cidadão, podendo ser reutilizados e redistribuídos contanto que se credite a autoria e que sejam compartilhados sob a mesma licença.

O movimento de dados abertos surgiu nos EUA a partir das demandas da sociedade civil norte-americana que buscava uma maior transparência dos gastos e ações do governo e ganhou força com o ativismo hacker em todo o mundo. Esse movimento se mescla também com o movimento de software livre e com as próprias lutas de liberdade de expressão e de imprensa. Entretanto, ele ganhou força com as ações do governo Obama em 2009 que instituiu os primeiros atos para construir canais que disponibilizassem dados do governo americano na Internet.

Em 2011, durante a reunião anual das Nações Unidas, a administração Obama firmou uma aliança denominada “Parceria para Governo Aberto” (OGP, do inglês *Open Government Partnership*) com outras sete nações (Brasil, Indonésia, México, Noruega, Filipinas, África do Sul, Reino Unido) objetivando a publicação e disseminação de informações governamentais na web (BRAGA e TUZZO, 2016).

No Brasil, desde 2000 já havia uma movimentação para disponibilizar dados governamentais para a sociedade civil através do programa e-GOV. O e-GOV tinha

¹² Do original: “It began as an element of the security regime, was taken up in trade, and is now used widely across the board. Transparency has replaced notions of the free flow of information in all of its variations as the ideal for international information flows and has now become a policy objective in its own right. This transformation in turn changes relations of individuals to society, and of societies to each other.”

o objetivo de transformar “as relações do Governo com os cidadãos, empresas e também entre os órgãos do próprio governo de forma a aprimorar a qualidade dos serviços prestados” (PGE, 2019).

Em 2004, o Portal de Transparência do Governo Federal foi criado pela Controladoria Geral da União (CGU) com o objetivo de divulgar os dados sobre a execução orçamentária e financeira, sobre as previsões e as receitas, convênios firmados, dados sobre cargo, função, situação funcional e remuneração dos servidores civis e militares.

Finalmente, em 2011 foi promulgada a Lei de Acesso à Informação (LAI)¹³ durante o governo de Dilma Rouseff. A partir da criação da lei, uma profusão de dados começou a ser disponibilizada na Internet em diversos formatos por vários órgãos do governo em todos os níveis federativos. Junto com a lei, vieram também a criação de diversos padrões que normatizavam a forma como estes dados deveriam ser disponibilizados pelos órgãos de governo.

Em 2012, foram criadas a Infraestrutura Nacional de Dados Abertos (INDA) e o Portal de Dados Abertos¹⁴. A INDA é um conjunto de padrões, tecnologias, procedimentos que visam a disseminação e compartilhamento de dados e informações públicas no modelo de Dados Abertos e o Portal de Dados Aberto é um sítio na Internet que visa agregar todos os conjuntos de dados disponíveis.

A apropriação dos dados abertos tem sido cada vez mais frequente por vários setores da sociedade civil, como universidade, imprensa, ONGs e até mesmo por outros órgãos governamentais. Entretanto, apesar das iniciativas do governo com o portal de Dados Abertos e com o site do Governo Eletrônico, os formatos dos arquivos e a usabilidade dos sites ainda dificultam muito o acesso ao conteúdo ofertado.

Ou seja, apesar dos dados estarem disponíveis, não são necessariamente acessíveis, já que a compreensão e tratamento destes dados não são triviais para um cidadão comum pois os metadados são de difícil entendimento, muitas vezes não estão documentados e a quantidade e complexidade das variáveis envolvidas tampouco facilitam o processo de análise (BURLE, 2015; VAZ et al, 2013; BRISOLA et al, 2017).

¹³ Lei 12.527/2011

¹⁴ <http://dados.gov.br>

Se não trazem minimamente os metadados, são raros os casos em que o conjunto de dados (*datasets*) vem acompanhado de artefatos semânticos que permitam que estes possam ser lida por um algoritmo de forma automatizada.

Além da questão documental, os dados na maioria das vezes são disponibilizados em arquivos que requerem uma competência em TI que grande parte das pessoas não tem. Em boa parte das vezes, os arquivos vêm em formato tabular (em CSV, XLS ou ODS) mas por terem um tamanho muito grande não conseguem ser lidos por programas de planilhas eletrônicas e assim, necessitam de linguagens de programação para que o usuário consiga visualizar os dados.

Na prática, a simples disponibilização de dados brutos não é suficiente para que haja realmente transparência e para que o usuário, de fato, os utilize de maneira cidadã e efetiva. A forma como esses dados são apresentados e o grau de competência em informação do usuário são facilitadores ou barreiras para que as informações sejam realmente aproveitadas plenamente.

Dada esta realidade, diversos ativistas e organizações não governamentais procuram criar o que se convencionou chamar de aplicativos cívicos. Um aplicativo cívico são softwares ou sistemas que facilitam o uso dos dados públicos e procuram “capacitar os cidadãos a exercer uma maior participação na vida política da sua cidade, região ou país, mediante a cobrança e fiscalização das atividades dos mandatários de cargos eletivos, bem como do judiciário e do serviço público em geral” (GOMES; SCHNEIDER; BEZERRA, 2018, p.3).

Utiliza-se também o termo “tecnologia cívica” e “software cívico” entretanto o primeiro termo é mais abrangente pois envolve iniciativas de hardware ou de dispositivos físicos que vão além dos dados abertos e o uso de “software cívico” parece ser mais delimitador pois restringe as iniciativas específicas de uso de softwares ou sites.

Em resumo, o termo “aplicativo cívico” não se restringe a um software desenvolvido para aparelhos celulares inteligentes (*smartphones*) ou até mesmo computadores. O termo remete a qualquer software, sistema ou iniciativa que facilite o acesso aos dados públicos e/ou abertos. Por exemplo, quando um ativista adiciona metadados a um conjunto de dados do governo e disponibiliza para a sociedade, facilitando assim o acesso aos dados, tal iniciativa se enquadra na categoria de aplicativo cívico.

Uma das metas do uso de aplicativos cívicos é a construção de agentes computacionais (usuários não humanos) que consigam extrair significados desses recursos e das relações existentes entre eles (ISOTANI e BITTENCOURT, 2015). Apesar de tal intento ainda está longe de ser alcançado, o uso de aplicativos cívicos pode facilitar a implementação deste tipo de tecnologia no longo prazo.

Especificamente para os dados a serem utilizados nesta pesquisa, três fontes de dados são centrais: a base de Financiamento de Campanha, os Resultados das Eleições e mais recentemente a base de empresas e sócios (Cadastro Nacional de Pessoas Jurídicas – CNPJ). As bases referentes às eleições são mantidas pelo TSE, sendo que a primeira base contém dados desde de 2002 e a de candidatos e resultados eleitorais desde 1945. Já a base de CNPJ foi disponibilizada em 2018 pela Receita Federal em função da grande pressão feita pelos ativistas para que a base fosse disponibilizada. Antes desta data, havia um “mercado paralelo” que vendia a base completa ou serviços que utilizavam os dados que deveriam ser públicos.

Devido a complexidade do tratamento de dados, algumas organizações não governamentais e ativistas disponibilizam estes mesmos dados em um formato mais amigável para um público leigo e também para limpar os arquivos que muitas vezes são disponibilizados com inconsistências (CPFs inválidos, codificação de caracteres diferentes entre os arquivos, etc). É relevante citar dois casos pois os dados utilizados nesta pesquisa fizeram uso de tais iniciativas: o Brasil.IO¹⁵ coordenado pelo ativista Álvaro Justen e o BasedosDados¹⁶, uma organização sem fins lucrativos que busca universalizar o acesso a dados públicos.

Especificamente sobre a base de CNPJ, é importante ressaltar que seu uso não é nada trivial. São mais de 15 gigabytes de dados e, além disso, a base foi disponibilizada em um formato pouco usual que remonta o tempo dos mainframes da década de 70. Para dificultar ainda mais o processo, a base não é disponibilizada com uma frequência programada¹⁷ o que faz com que alguns dados podem estar defasados. Além do Brasil.IO e do BasedosDados, outras iniciativas

¹⁵ <https://brasil.io/>

¹⁶ <https://basedosdados.org/>

¹⁷ Até Agosto de 2021, havia um atraso de 6 meses por parte da Receita Federal na disponibilização de dados recentes. Para o efeito da presente pesquisa, não houve muito prejuízo mas se eles mantiverem essa inconstância, novas pesquisas podem se inviabilizar para as eleições de 2022.

como o [minhareceita.com](https://github.com/cuducos/minha-receita)¹⁸ e a base montada pelo pesquisador George Santiago¹⁹ também permitiram um acesso mais amplo a esse conjunto de dados.

Em resumo, existe um vasto campo a ser explorado no universo de dados abertos e a integração da comunidade acadêmica com os chamados aplicativos cívicos pode abrir caminho para novos tipos de pesquisas e metodologias. O fato concreto é que provavelmente esta pesquisa não teria conseguido trabalhar com estes dados em tempo hábil sem o suporte de tais iniciativas.

¹⁸ <https://github.com/cuducos/minha-receita>

¹⁹ <https://github.com/georgevsantiago/qsacnpj>

3 CLASSIFICAÇÃO DAS RELIGIÕES

Neste capítulo, é feita uma revisão da literatura sobre o que são os sistemas de organização do conhecimento, sua importância na Ciência e exemplos de uso nas ciências sociais. A partir do conceito mais amplo, buscar-se-á referências que tratam especificamente da classificação das religiões para que se possa, em seguida, detalhar a classificação específica dos evangélicos.²⁰

Classificar as religiões é uma tarefa árdua e não há consenso na literatura analisada sobre a melhor maneira de classificá-las (ADAMS, 2018). De fato, não há sequer consenso sobre os termos, ou seja, sobre o que é uma religião, denominação, seita ou culto (LIEBMAN, SUTTON, WUTHNOW, 1988).

Em uma análise sobre as diversas iniciativas que buscam classificar as religiões, Smith (1996) discorre sobre os graves problemas éticos em tentar hierarquizar ou ordenar as religiões. Em seu artigo denominado “Um assunto de classe: Taxonomias da Religião”, o historiador brinca com as palavras e ressalta que definir uma religião como hierarquicamente superior a outra, ou até mesmo considerar que práticas religiosas de um determinado povo não podem ser consideradas como religião, é algo que não cabe mais em nosso tempo e que os pesquisadores não estão mais dispostos a aceitar esse tipo de categorização utilizada no passado.

Assim, a presente revisão não tem como objetivo discutir questões teológicas das classificações em si, mas focar nos processos informacionais, éticos e políticos de como tais classificações foram construídas. Busca-se entender como os principais sistemas classificatórios categorizaram as religiões para que se possa ter uma visão mais ampla do assunto. Ou seja, a proposta aqui não é explicar ou detalhar os dogmas que determinam uma outra religião mas analisar criticamente como os sistemas construíram tais classificações e, em especial, como foi feita a classificação dos protestantes.

Buscou-se assim três fontes distintas para que fosse possível construir um panorama de como a classificação das religiões é realizada no Brasil e no mundo.

²⁰ Este capítulo é formado em grande parte pelo artigo aprovado e publicado nos Anais da ISKO 2020 sob o título “Ethical Perspective on Classifications of Religions. The Protestant Rise in Brazil”

A primeira fonte, mais geral e universal, são os sistemas de classificação bibliográfica que organizam as bibliotecas em todo o planeta. A partir da visão destes sistemas, busca-se entender como as religiões foram classificadas a partir do início do século XX.

A segunda fonte se baseou na categorização do Pew Research Center (PRC), uma organização não governamental de pesquisa estadunidense que realiza diversas pesquisas na área da Sociologia e Religião. Esta categorização tem como fonte os diversos censos nacionais e pesquisas amostrais realizadas nos cinco continentes e a escolha deste instituto se deu em função da sua abrangência global, do uso recorrente de suas pesquisas em estudos acadêmicos e do detalhamento fino que eles realizaram especificamente sobre o Protestantismo, foco da presente pesquisa.

A terceira fonte, já com um enfoque específico na realidade brasileira, foi o Censo do IBGE de 2010 que procurou classificar e enumerar quais são as religiões existentes no país e como estas estão distribuídas em seu território. Apesar de ter recebido críticas por diversos pesquisadores (ALTMANN, 2012; MAFRA, 2013; CAMURÇA, 2014), a classificação adotada pelo Censo serve como base para boa parte das pesquisas acadêmicas por conta de sua série histórica e por ser uma fonte de dados aberta.

3.1 SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO

Dahlberg (1978), define a organização do conhecimento como uma ciência com competência para ordenar a estruturação e sistematização dos conceitos a partir de suas características. Especificamente no âmbito da Ciência da Informação, segundo Hjørland (2003, p.1), a Organização do Conhecimento tem como objetivo a organização da informação em registros bibliográficos, incluindo a indexação de citações, o registro de textos integrais e da própria Internet.

Para atingir esse objetivo, busca-se construir Sistemas de Organização do Conhecimento (SOC). Por definição, SOC são sistemas semanticamente estruturados que através do uso de conceitos, termos, definições e relacionamentos entre estes procuram representar o conhecimento que se quer organizar. Os SOCs variam enormemente em formato e exibição, mas compartilham a característica

geral de auxiliar a produção e organização do conhecimento, visando promover e otimizar a recuperabilidade da informação (SOUZA et al, 2012).

A construção de um SOC depende essencialmente da necessidade do que se pretende organizar e de qual o objetivo que se quer atingir com tal sistematização. Alguns sistemas tem o objetivo de sistematizar uma área específica de um campo científico ou a base de conhecimento de uma instituição ou de um empreendimento. Já outros são generalistas e tem o intuito de serem universais, ou seja, abarcar todo o conhecimento humano tal como são os sistemas de classificação bibliográfica. São exemplos de SOC, os tesouros, as taxonomias, as ontologias e os sistemas bibliográficos presentes na grande maioria das bibliotecas do planeta.

Atualmente o sistema de classificação bibliográfica mais utilizado no mundo é a CDD (Classificação Decimal de Dewey) criado originalmente em 1876 pelo bibliotecário americano Melvil Dewey. Este sistema de classificação é dividido em dez classes principais, que tentam cobrir todas as áreas do conhecimento existente (NUNES et al, 2015). É um sistema mantido por uma organização privada²¹ e mesmo sendo atualizado de 10 em 10 anos, mantém suas bases em conceitos do século XIX.

Um outro sistema ocidental utilizado em mais de 130 países (inclusive no Brasil) é o UDC (Universal Decimal Classification). Criado pelos belgas Paul Otlet e Henri La Fontaine também no final do século XIX, o UDC tinha como objetivo ser um sistema mais aberto e mais amplo que o americano.

Grande parte dos países ocidentais utilizam estes dois sistemas em suas bibliotecas. Entretanto, países não alinhados à cultura ocidental como China e Índia buscaram construir seus próprios sistemas. Sob o ponto de vista histórico, a literatura ocidental comumente trata Aristóteles como um dos primeiros pensadores a estudar os sistemas de classificação bibliográfica. Entretanto, 200 anos antes do nascimento do filósofo grego, Confúcio já teria compilado os primeiros sistemas bibliográficos na China (TSIEN, 1958).

O caso da China é interessante pois o país contava com um rico acervo de sistemas bibliográficos até a ocupação militar e política por parte das potências europeias e do Japão no século XIX. E somente na década de 1970, 30 anos após o

²¹ A Online Computer Library Center é uma cooperativa sem fins lucrativos e adquiriu a marca e os direitos autorais do Sistema Decimal de Dewey quando incorporou a editora Forest Press em 1988.

início da Revolução Comunista, a China implementou seu próprio sistema de classificação bibliográfica. Na verdade, existem na China, até a presente data, quatro diferentes esquemas de classificação, cada um com propósitos diferentes. A partir do esforço para criar um único sistema classificatório por parte da Universidade de Beijin, foi criado o CCL (*Classification for Chinese Libraries*), mas este novo sistema ainda mantém as mesmas bases ideológicas dos quatro sistemas iniciais (STUDWELL, WU e WANG, 1994).

O CCL é norteado por uma arquitetura bem distinta da visão aristotélica ocidental. Enquanto a Filosofia é o primeiro item hierárquico dos esquemas classificatórios ocidentais, a classificação chinesa traz o Marxismo como sua primeira classe e a classe de religião, por exemplo, aparecia como um mero apêndice da classe Filosofia (ZHANG, 2003) até a penúltima revisão desta classificação. Na revisão atual, a classe “Filosofia” foi alterada para “Filosofia e Religiões”.

A Índia também construiu seu próprio sistema bibliográfico a partir do trabalho do filósofo e matemático indiano Ranganathan. Denominado Colon Classification, este sistema também foi construído para se adequar à realidade cultural do país e já em 1933 contava com uma inovação conceitual que ainda não estava presente nos sistemas americano e europeu: a classificação facetada. Neste modelo, um assunto teria sempre 5 tipos de facetas: personalidade, matéria, energia, espaço e tempo e a combinação destas facetas formariam a classificação de um dado item da biblioteca. Dessa forma, tal classificação pretendia abranger vários atributos ao invés de utilizar um único critério. Além disso, as classes mais gerais tratavam com maior detalhe os aspectos da cultura indiana que não apareciam com muita relevância nas classificações ocidentais.

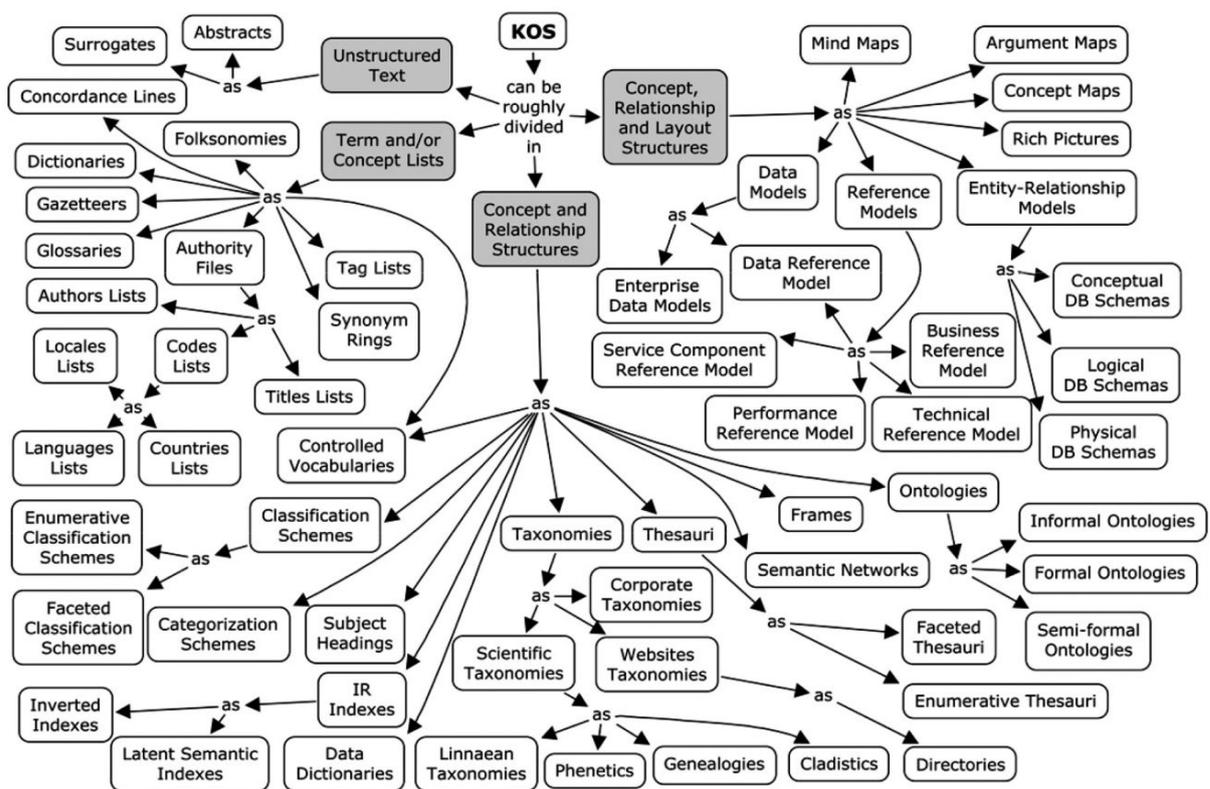
Existem muitos outros sistemas de classificação bibliográficos além da CDD, CDU e CC. Na Grã-Bretanha, por exemplo, boa parte das bibliotecas utiliza o Sistema Bliss desenvolvido no início do século XX, mas que também se estrutura a partir da CDD. Mesmo nos Estados Unidos da América, a CDD divide espaço com a CLC, que é a classificação do Congresso Americano e que foi criada a partir da classificação do filósofo Francis Bacon.

Todos esses exemplos mostram que muitos países procuraram criar seus próprios sistemas bibliográficos em função das suas necessidades locais pois

perceberam que os sistemas classificatórios são centrais para a organização do conhecimento de uma nação e são instrumentos importantes para reforçar valores culturais e políticos.

Além do universo dos sistemas bibliográficos, Souza et al (2012) ressaltam que existem dezenas de tipos de SOC, cada qual com a sua particularidade. Ainda segundo os autores, com o advento da Internet e dos avanços em torno da Web Semântica, o interesse por tais sistemas aumentou consideravelmente, já que os SOCs são essenciais para a implementação de sistemas de informação que necessitem de um suporte semântico, ou seja, um suporte que permita registrar em formato legível por máquina as definições e relações dos objetos a serem disponibilizados em tais sistemas. A lista dos tipos de SOC é explicitada pelos autores na Figura 3.

Figura 3 – Tipos de SOC (KOS)



Fonte: Souza et al (2012)

De fato, a pesquisa no campo é tão importante para a indústria da Internet que o *World Wide Web Consortium* (W3C) desenvolveu uma especificação denominada *Simple Knowledge Organization System* (SKOS) com o objetivo de

padronizar uma linguagem comum entre os SOC's para que houvesse interoperabilidade entre os sistemas na Internet.

Além disso, como já foi visto no capítulo anterior sobre Mineração de Dados, os algoritmos e sistemas de classificação gerados por máquina também buscaram na teoria sobre construção de SOC formas de otimizar e organizar os resultados obtidos nos processamentos automatizados.

O processo de construção de um SOC parte de 3 atividades centrais: definição de termos e conceitos, definição das relações entre os conceitos e, por fim, representação e classificação dos mesmos em estruturas lógicas que facilitem a recuperação da informação registrada. Ou seja, a classificação é uma das etapas do processo, mas necessita de outros passos para que possa ser bem sucedida.

No âmbito da CI, as duas primeiras atividades são trabalhadas, em grande medida, a partir da Teoria do Conceito de Dahlberg (1976), que procura mostrar a relevância de conceituar os termos e a linguagem a ser utilizada nos SOC. Carlan e Bräscher explicitam a importância desse aspecto:

Na organização e recuperação da informação, os SOC cumprem o objetivo de padronização terminológica para facilitar e orientar a indexação e os usuários. Quanto à estrutura variam de um esquema simples até o multidimensional, enquanto que suas funções incluem a eliminação da ambiguidade, controle de sinônimos ou equivalentes e estabelecimento de relacionamentos semânticos entre conceitos (Carlan e Bräscher, 2011, p.54).

O ato de classificar está no cerne do método científico. A criação de tipologias e de categorizações faz parte do ferramental que auxilia a ciência a organizar estruturas lógicas que auxiliem a representação do conhecimento. Bowker e Star (2000, p.131) vão mais além e apontam que “classificar é humano e todas as culturas, em todos os tempos produziram sistemas de classificação”. Entretanto os autores também alertam sobre como geralmente ocorre o processo de classificação:

[...] atribuir coisas, pessoas, ou suas ações a categorias é parte permanente do trabalho no burocrático estado moderno. Categorias nesse sentido surgem do trabalho e de outros tipos de atividade organizada, incluindo os conflitos sobre significados que ocorrem quanto múltiplos grupos lutam pela natureza de um sistema de classificação e suas categorias (Bowker; Star, 2000, p.285).

Mai (2013) corrobora essa visão ao indicar que qualquer sistema sempre conterá um viés de algum tipo. A questão é qual seria o viés aceitável e que tal viés

poderia ser em grande parte minimizado se tal ordenamento ou escolha fosse declarado e discutido com a comunidade usuária.

Tal como cita a pesquisadora Olga Pombo, “cada classificação tem por detrás um determinado mecanismo classificador que executa, melhor ou pior, as operações necessárias à classificação” (Apostel, 1963, p.195, apud POMBO, 2015, p.2).

A classificação mais simples possível é a binária, que parte da escolha de um determinado atributo dos objetos que se quer classificar e a separação é feita através da ausência ou presença daquele atributo. Retomando ao exemplo dos candidatos e candidatas a um cargo eletivo, uma classificação binária indica se ele foi eleito ou não. O processo fica mais complexo se ao invés de se ter ou não um atributo, o atributo assumir diversos valores, por exemplo, se ele é casado, solteiro, viúvo ou divorciado. Nesse exemplo, não existe alternativas intermediárias. Ele tem que necessariamente estar em um destes quatro estados.

E quando o atributo é um número discreto, por exemplo, a idade deste ou desta candidata? Neste caso, o ente classificador (quem realiza a classificação) pode optar por utilizar duas ou mais faixas de idade para agrupar os objetos a classificar. Como ele determina quais as faixas que irá utilizar? Nesse ponto, ele pode criar suas próprias faixas de idade, utilizar faixas construídas em pesquisas semelhantes à sua ou utilizar faixas universais que se tornaram padronizadas em seu país ou área de conhecimento.

O processo de classificação pode ser ainda mais complexo: e se o pesquisador quiser classificar, por exemplo, os candidatos de esquerda, centro e direita? Quais os critérios que determinam quem pertence a cada classe? Esse critério será unânime?

Obviamente nenhum objeto de estudo tem apenas um atributo. Cada objeto é formado por dezenas ou centenas de atributos que, se necessários na classificação, podem criar um sistema muito complexo que irá envolver hierarquias ou esquemas com múltiplas facetas. Novamente, a escolha do classificador fará toda a diferença no resultado final. Quais dos atributos serão utilizados?

Pode-se então pensar nas milhares de classificações que o ser humano vem criando ao longo do tempo. Desde as mais nobres e bem-intencionadas, como a classificação dos seres vivos desenvolvida por Carolus Linnaeus²² no século XVIII

²² Médico sueco que publicou a obra *Systema naturae* que buscou classificar todos os seres vivos. A

até as mais sórdidas, como as que classificavam os seres humanos em raças mais ou menos desenvolvidas. A própria linguagem está imbuída de séculos de esquemas classificatórios. Por exemplo, em relação à frase acima, que cita as mais “nobres” e bem-intencionadas classificações, pode-se questionar por que a nobreza tem uma conotação positiva? Por que não se diz, por exemplo, a mais plebeia das intenções?

Omnis determinatio est negatio (Spinoza apud Lenin 2011, p.111) ou traduzindo, “toda determinação exclui, em termos puramente lógicos”. Em termos éticos e políticos, o que significa cada conjunto de determinações consagradas, como os principais esquemas de classificação? Quais são as visões de mundo que operam em cada caso, por que razões, com que consequências?

Para Bourdieu (1989, p.142), o trabalho de categorização, de explicitação e de classificação “faz-se sem interrupção, a cada momento de existência corrente, a propósito das lutas que opõem os agentes acerca do sentido do mundo social e da sua posição nesse mundo”. De fato, as escolhas metodológicas em classificar um determinado domínio não são neutras e procuram fortalecer uma forma de enxergar o universo a ser classificado.

Ressalta-se assim a falsa neutralidade dos sistemas classificatórios, pois, como vimos acima, “essa não-neutralidade não se refere apenas ao fato de que eles são necessariamente tendenciosos, mas que preconceitos antiéticos específicos não são reconhecidos como tais”. (ICIE, 2021)

Dessa forma, todo processo de classificação que compõe o método científico deve ser feito com atenção redobrada para que as escolhas metodológicas sejam detalhadas de forma explícita, de modo a oferecer ao leitor a possibilidade de analisar criticamente os resultados obtidos. Partindo assim dessa visão crítica e atenta ao ato de classificar, adentremos no espinhoso terreno da classificação das religiões.

3.2 AS RELIGIÕES NOS SISTEMAS BIBLIOGRÁFICOS

Conforme visto na seção anterior, os dois principais esquemas de classificação bibliográfica se basearam na cultura ocidental e foram criados no auge da Era do Imperialismo, antes da Primeira Guerra Mundial. As religiões cristãs eram

obra é considerada o ponto de partida da moderna nomenclatura binominal utilizada até os dias de hoje.

hegemônicas na sociedade em que Dewey e Otlet viviam, enquanto a Ásia, África e América Latina (o agora chamado Sul Global) eram fornecedores exclusivos de matéria prima para a Europa e EUA.

Essa cultura hegemônica se refletiu em ambos os esquemas classificatórios. O CDD e o CDU definiram a classe Religião (representada pelo código 200 em ambos os esquemas). A CDD reservou a maioria das classes principais para o Cristianismo, mantendo a classe “Outros” (Código 290) para todas as outras religiões do planeta, agrupando em uma única classe, no mesmo nível hierárquico, a mitologia Grega, o Islamismo e o Budismo. Já a CDU foi bem mais inclusiva que a CDD, pois especificou bem mais religiões em seu esquema primário de classificação além de separar as religiões entre suas regiões de origem e a partir de uma ordem cronológica.

Seria justo, entretanto, culpar os profissionais que participaram da elaboração dos primeiros esquemas de classificação? Pragmaticamente, quantos livros relacionados ao hinduísmo, ao islamismo ou às religiões da China ou do Japão chegavam aos países ocidentais naquela época? Seria factível classificar as religiões de acordo com sua importância se o número de documentos que chegavam às bibliotecas (e aos classificadores) era muito pequeno?

Por exemplo, originalmente na UDC havia toda uma seção de classificação do judaísmo (com 49 subclasses), enquanto outras religiões com muito mais crentes naquela época não tinham tal relevância. Havia uma influência política ou socioeconômica para incluir tantas subclasses no esquema do Judaísmo? Ou simplesmente existiam mais documentos sobre o judaísmo nas bibliotecas ocidentais, já que a presença judaica no Ocidente é relevante desde os primeiros séculos do primeiro milênio da era comum, além do judaísmo estar na própria origem do cristianismo, apesar de ter sido frequentemente perseguido por este último.

Estes esquemas baseados principalmente em classes cristãs mudaram com a *Colon Classification* (CC). Pela primeira vez, a classificação construída pelo filósofo indiano Ranganathan tentava representar com mais detalhes as religiões orientais. Aparentemente, a classificação proposta pela CC listava as religiões em uma ordem histórica, ou seja, apresentando primeiro as religiões mais antigas nas primeiras

posições do esquema classificatório. Entretanto, não foi possível encontrar uma referência bibliográfica que explicitasse essa escolha metodológica.

Vale notar também que na CC, existe uma distinção entre Religião (Classe Q) e Experiências Espirituais e Misticismo (Classe Δ), ou seja, dentro desta classificação (baseada na cultura indiana) as experiências espirituais não são fenômenos necessariamente ligados à religião e estão hierarquicamente no mesmo nível que a Religião ou a Filosofia.

As três classificações podem ser vistas nas tabelas 2, 3 e 4, agrupadas em uma única página ao final da seção para facilitar a visualização e sem a tradução para o português, para que os termos em inglês pudessem ser comparados nas três classificações.

Em contrapartida, como podemos ver na tabela 5, a CC também foi muito concisa ao classificar as religiões cristãs da mesma forma que a CDD e a CDU tampouco se aprofundaram nas religiões orientais. Vale notar também que nas três classificações, os termos seita (*Sect*), igreja (*Church*), movimentos (*Movement*) e denominação (*Denomination*) se confundem corroborando a dificuldade em se utilizar tais conceitos. Na tabela comparativa, os termos foram deliberadamente mantidos em inglês e enfatizados com o intuito de demonstrar como estes aparecem sem um critério aparente.

De fato, diversos autores abordaram como os esquemas de classificação são tendenciosos quando o assunto é religião. Idrees e Mahmood (2009) investigaram o viés negativo dado ao Islamismo. Miranda e Silva (2019) realizaram extensa revisão bibliográfica e também um levantamento terminológico de assuntos ligados ao Islamismo. Já no Brasil, Costa e Miranda (2019) e Miranda (2018) mostram como a Umbanda, uma das diversas religiões originárias do povo africano existentes no Brasil, também é sub-representada ou em diversas vezes classificada erroneamente como “magia” ou “feitiçaria” e não como uma religião.

Ainda sobre a questão da visibilidade (ou invisibilidade) do Sul Global, nas três classificações não há uma classe primária ou secundária para as religiões africanas, para as religiões originárias das Américas e também as religiões de toda a região do Pacífico. Elas foram categorizadas como sendo todas do mesmo tipo em função da sua localização geográfica independente de terem histórias e ritos totalmente distintos. Tal como Adams (2018) ressalta

no classification should ignore the concrete historical individuality of religious manifestations in favour of that which is common to them all, nor should it neglect to demonstrate the common factors that are the bases for the very distinction of types of religious experience, manifestations, and forms (Adams, 2018).²³

Vanda Broughton (2000) identificou três áreas principais onde o viés pode ocorrer:

an illogical order, or distribution of notation, that causes one system to appear as dominant, use of vocabulary that has a strong flavour of one system or is special to that system and inadequate provision of detail other than for the 'favoured' religion”²⁴ (Broughton 2000, p.56).

A autora observa, entretanto, que é possível minimizar o viés através do uso de técnicas analíticas de facetas e que a mais recente revisão da Classe 2 da CDU pode ser composta e pode atingir um melhor grau de especificidade. Idrees e Khalid (2009) também propuseram emendas e expansão a fim de dar uma melhor orientação aos bibliotecários sobre como classificar o Islã.

Os editores da CDD também admitem o problema: “Nós somos os primeiros a admitir que a visão mais geral da classe 200 Religião é problemática” e a partir desta constatação, eles iniciaram um estudo “sobre o futuro da estrutura 200” (Green, 2006).

A preocupação em tentar minimizar o viés sobre Religião é exemplificada também pelo artigo de McIlwaine e Mitchell (2006), primeiro na Conferência ISKO e depois replicado em um dos próprios relatórios da UDC, onde os autores sugeriram formas de minimizar o impacto do viés de classificação através de uma tabela auxiliar que lista as religiões em ordem cronológica, ou seja, das religiões mais antigas para as mais novas. No mesmo artigo, eles concordaram com Broughton que a revisão da classe 2 da UDC traria benefícios a tal sistema de classificação.

Ao longo desta seção, foi demonstrada a dificuldade em classificar as religiões em sistemas tradicionais e de grande alcance como os sistemas

²³ nenhuma classificação deve ignorar a individualidade histórica concreta das manifestações religiosas em favor daquilo que é comum a todas elas, nem deve negligenciar a demonstração dos fatores comuns que são as bases para a própria distinção dos tipos de experiência religiosa, manifestações e formas.

²⁴ uma ordem ilógica, ou distribuição de notação, que faz com que um sistema apareça como dominante, uso de vocabulário que tem um forte sabor de um sistema ou é especial para esse sistema e provisão inadequada de detalhes que não seja para a religião "favorecida". inadequado de detalhes que não sirvam para a religião "favorecida".

bibliográficos. Nas próximas seções, veremos especificamente como os protestantes são classificados nos censos internacionais e na realidade brasileira.

Tabela 2 - Classificação das Religiões na CDD

Philosophy & theory of religion	210
The Bible	220
Christianity & Christian theology	230
Christian practices & observances	240
Christian pastoral practice & religious orders	250
Christian organization, social work & worship	260
History of Christianity	270
Christian denominations	280
Other religions	290

Fonte: Elaborado pelo autor a partir de Dewey, 2003

Tabela 3 - Classificação das Religiões na CDU

Prehistoric religions. Religions of early societies	21
Religions originating in the Far East	22
Religions originating in Indian sub-continent	23
Buddhism	24
Religions of antiquity. Minor cults and religions	25
Judaism	26
Christianity	27
Islam	28
Modern spiritual movements	29

Fonte: Elaborado pelo autor a partir da CDU

Tabela 4 - Classificação das Religiões na Colon Classification

Hinduism, Vedic	1
Hinduism, Post Vedic	2
Jainism	3
Buddhism	4
Judaism	5
Christianity	6
Muhammadanism	7
Other religions	8

Fonte: Elaborado pelo autor a partir da CC

Tabela 5 - Classificação das Religiões Cristãs na CDD, CDU e CC

CDD	CDU	Colon (Q Class)
281 Early church and Eastern churches	281 Primitive Churches. Eastern Churches .	61. Early Churches
	281.93 Russian Orthodox Church	618. Russian
	281.94 Greek Orthodox Church	611. Greek
		613. Armenian
282 Roman Catholic Church	282 Roman Catholic Church	62. Roman Catholic
283 Anglican churches	283 Episcopal Churches – not Roman Catholic (Protestants)	
	283(410.1) Anglican Church	
	283(73) Episcopal Church on USA	
	283.5 Old Catholics	
284 Protestant denominations of Continental original	284 Continental Protestant Sects	63. Protestant
	284.1 Lutherans	
	284.2 Calvinists	
	284.3 Utraquists. Tabortas	
	284.4 Coterons. Gazarists	
	284.5 Huguenotes. French Movements	
	284.6 Moravian Brothers. Hernutos	
	284.98 Old Lutherans. Free Lutheran church	
285 Presbyterian churches , Reformed centered in America	285 Puritanism	65. Puritanism

CDD	CDU	Colon (Q Class)
	285.1 Presbyterian Churches	64. Presbyterian
	285.8 Brownists. Barrowists.	
286 Baptist, Disciples of Christ, Puritanism	286 Movements that accentuate the baptism of Adults by immersion	
	286.12 Anabaptists. Menonists.	
	286.15 Baptists.	
	286.3 Adventists.	
	286.4 Other movements. Campbellists. Christ Disciples (USA)	
287 Methodist churches	287 Methodist	68L6. Methodist
289 Other denominations and sects	289 Other Movements	
	289.3 Mormons	
	289.4 New Jerusalem (Swedenborg)	
	289.6 Quakers	66. Quakers
	289.954 Jeovah Witnesses	
	289.956 Liberal Catholic Church	

Fonte: Elaborada pelo autor a partir da CDD, CDU e Colon Classification

3.3 O PROTESTANTISMO NO MUNDO

São denominadas protestantes, as pessoas que professam a fé cristã e fazem parte de grupos religiosos ou igrejas que emergiram da reforma protestante iniciada por Lutero no século XVI. Segundo o relatório "Global Christianity" elaborado pelo PRC (2011), havia 801 milhões de protestantes no mundo em 2010 e eles representavam 36,7% da população cristã global.

Em relação às demais religiões, segundo outro relatório da mesma instituição denominado "The Future of World Religions: Population Growth Projections, 2010-2050", o Cristianismo ainda é a religião com o maior número de adeptos. Segundo o relatório, as religiões foram categorizados em 6 grupos: Cristianismo, Islamismo, Budismo, Hinduismo, Religiões Tradicionais ou "Populares" (*Folk Religions*) e o Judaísmo.

Tabela 6 - Religiões no Mundo em 2010

Religião	População
Cristianismo	2.168.330.000
Islamismo	1.599.700.000
Não afiliados / Sem religião	1.131.150.000
Hinduismo	1.032.210.000
Budismo	487.760.000
Religiões Populares (<i>Folk Religion</i>)	404.690.000
Outras	58.150.000
Judaísmo	13.860.000

Fonte: Pew Research Center. Traduzido pelo autor.

É possível notar que nesta classificação se repetem os mesmos problemas já discutidos na seção anterior quando a classificação cria dois grandes grupos denominados “Religiões Populares” e “Outras” que englobam religiões que são muito distintas entre si. Segundo o relatório, as Religiões Populares são “as ligadas a um determinado povo, etnia ou tribo e essas religiões muitas vezes não têm credos formais ou textos sagrados”. Exemplos incluem religiões tradicionais africanas, religiões populares chinesas, religiões nativas americanas e religiões aborígenes australianas e as outras são as religiões que não se encaixam nos outros seis grupos.

Sanchis procura problematizar o que se chama de Religião Popular:

O que seria, de fato, uma religião popular? Rapidamente, impõe-se a constatação de que religião popular não é simplesmente religião "outra", mas é inteiramente religião "relativa". Ela integra um "campo", no sentido de Bourdieu, opondo-se dinamicamente. A que? Na totalidade do campo, pode ser a outras religiões (ou outra religião), dominantes(s): existem religiões institucionalmente populares neste sentido. Mas também no interior de uma denominação, correntes "populares" existem, que se opõe (ou se articulam com) à religião oficial da instituição. (Sanchis, 2007, p.7)

Tafjord também ressalta esse viés sobre as religiões denominadas “Indígenas” mas que se aplica também na classificação do PRC:

Eles não são colocados na mesma categoria porque tiveram entre si algo de especial em comum, independente do olhar europeu. O que eles compartilham, histórica e estruturalmente, que ao mesmo tempo não compartilham também com outras "religiões", é quase nada. (Tafjord, 2013, p.6, tradução nossa)²⁵

²⁵ Do original: They are not placed in the same category because they amongst themselves have had something special in common, independent of the European eye. What they share, historically and

Este agrupamento das religiões ditas populares e até mesmo o grupo de “Outros” novamente invisibiliza um enorme contingente populacional que professa outras religiões. Por exemplo, somente no Censo Indiano de 2001²⁶ o número de Sikhs chegou a 19 milhões de pessoas. Por que o PRC especifica em sua categorização o Judaísmo com 13 milhões de praticantes e não especifica o Sikhismo que também tem adeptos em todo o planeta?

Uma segunda análise interessante pode ser feita a partir do número de respondentes da categoria “Não afiliados / Sem religião”. A princípio, este número elevado reforça o discurso de secularização da sociedade moderna que indica que existe uma tendência de que a humanidade está deixando de ser religiosa. Entretanto, ao analisar a metodologia da pesquisa da PRC, nota-se que a invisibilidade de outras religiões também pode estar ocorrendo já que, segundo o relatório, estão inseridos nos “sem religião” não apenas os ateus e agnósticos, mas também pessoas que disseram que acreditam em Deus, só que não professam uma religião específica. Ora, que Deus é esse que a pessoa entrevistada acredita? E qual o percentual de ateus em relação aos que disseram que acreditam neste ou nestes Deuses? O relatório do PRC não detalhou estas questões em sua metodologia.

Já o relatório específico sobre o Cristianismo veio bem detalhado. Categorizou-se os cristãos por sua história e origem. Esta forma de classificação se assemelha à classificação bibliográfica vista na seção anterior. O relatório categorizou o cristianismo em quatro grandes grupos: Católicos, Ortodoxos, Protestantes e "Outros Cristãos", tal como pode ser visto na tabela abaixo:

Tabela 7 - Cristianismo no Mundo em 2010

Tradição Cristã	%
Católicos	50,1
Protestantes	36,7
Ortodoxos	11,9
Outros cristãos	1,3

Fonte: Pew Research Center. Traduzido pelo autor.

structurally, that they at the same time do not also share with other "religions," is close to nothing.
²⁶ https://censusindia.gov.in/Census_and_You/religion.aspx

O último grupo refere-se a todos os sub-grupos que não se enquadram nos três primeiros grupos cristãos. A classificação de católicos e ortodoxos parece ser fácil de realizar, pois são igrejas que têm atributos teológicos, históricos e geográficos bem definidos ao longo de quase 18 séculos. No entanto, como fazer a diferença entre protestantes e outros tipos de cristãos, que são bastante emblemáticos como os mórmons, mas que fazem questão de não serem classificados como protestantes?

Tal dificuldade em categorizar os protestantes foi observada desde o início do estudo do protestantismo. Uma das bibliografias mais respeitadas sobre o assunto foi a *Ética Protestante e o Espírito do Capitalismo*, escrita pelo sociólogo alemão Max Weber. Neste trabalho, Weber tentou categorizar os vários grupos de protestantes que existiam no seu tempo e já indicava a dificuldade em realizar tal categorização:

E só podemos fazê-lo apresentando as ideias religiosas com a consistência lógica de um “tipo ideal”, que só raramente se deixa encontrar na realidade histórica. Precisamente por causa da impossibilidade de traçar limites nítidos na realidade histórica, nossa única esperança ao pesquisar as mais coerentes de suas formas é atinar com os seus efeitos mais específicos (Weber 1999, p. 90).

O autor categorizou os protestantes em cinco grandes grupos (Luteranos, Calvinistas, Pietistas, Metodistas e Anabatistas), baseando-se nos dogmas morais que cada grupo seguia. O interessante da categorização adotada por Weber veio da necessidade de perceber quais grupos mais se ajustavam no chamado espírito do capitalismo, a nova ética que tinha surgido no mesmo período histórico que a reforma protestante e que se tornou mais forte justamente nos países que abraçaram o capitalismo com mais intensidade. Ou seja, a opção metodológica de categorização tinha um objetivo de entender como o novo ethos capitalista se fortalecia em cada um dos grupos.

O PRC também procurou realizar um outro tipo de categorização, agrupando os cristãos em três grandes “movimentos”: Pentecostais, Carismáticos e Evangélicos. Nesta segunda categorização, o critério utilizado se baseou nos dogmas centrais de cada grupo religioso, independente do contexto histórico e geográfico. Os pentecostais são os que acreditam em uma experiência pós-conversão baseada no batismo pelo Espírito Santo. Os evangélicos são os que procuram seguir estritamente o que está escrito no Evangelho e os carismáticos são

os que não estão nos primeiros grupos. Destaque-se que aqui os evangélicos são o termo utilizado nos EUA e Europa para um movimento conservador entre protestantes, como assinalado na introdução do trabalho, ou seja, o termo não tem o mesmo significado utilizado no Brasil e América Latina.

Vale notar também que nesta categorização, o PRC não utilizou o termo protestante, pois os grupos denominados “protestantes” estavam presentes nas três categorias definidas pela pesquisa, o que demonstra como a forma de categorizar as religiões difere do uso comum do termo.

Sobre a distribuição geográfica, o relatório demonstrou que a maior nação protestante em números absolutos continua sendo os EUA. A tabela abaixo mostra os países com o maior número de protestantes e o percentual em relação ao total da população de cada país. Ressalta-se a presença de inúmeros países africanos, o que demonstra como a presença dos protestantes cresceu ao redor do planeta no século XX.

Tabela 8 - Países com maior número de protestantes em 2010

Países	População de Protestantes (em milhões)	%
EUA	159,85	51,5
Nigéria	59,68	37,7
China	58,04	4,3
Brasil	40,50	20,8
África do Sul	36,55	72,9
Reino Unido	33,82	54,5
Congo	31,70	48,1
Alemanhã	28,64	34,8
Quênia	24,16	59,6
Índia	18,86	1,5
Total do resto	308,82	1,8
Total Mundial	800,64	11,6

Fonte: Pew Research Center. Traduzido pelo autor.

A metodologia da coleta de ambos os relatórios (o Global Christianity e o The Future of World Religions) foi baseada em 3 fontes distintas: os dados dos censos populacionais dos países quando estes coletaram dados sobre religião, de institutos de pesquisas locais e das próprias instituições religiosas quando não foi possível obter os dados das duas primeiras formas. Entretanto, a pesquisa da PRC não

especificou qual a fonte utilizada para cada país, ou seja, não foi possível identificar o percentual de fontes em relação ao total de países analisados.

A partir de pesquisa documental realizada neste trabalho em cada um dos países da América do Sul, Central e do Norte, pode-se constatar que apenas seis países conduziram uma pergunta sobre a religiosidade em seus censos nacionais após 2010: Brasil, Canadá, México, Peru, Jamaica e Haiti. Todos os outros países do continente americano dependem de institutos privados para estimar quais as religiões professadas por sua população e segundo os dados coletados pela ONU, dos 285 países existentes atualmente no planeta, apenas 121 conduziram pesquisas sobre religião em seus censos nacionais²⁷.

Ainda sobre a análise dos censos nacionais, um caso interessante ocorre nos Estados Unidos, onde existe uma lei específica que proíbe o Censo Nacional de conduzir qualquer investigação sobre religião (PRC, 2010). A lei promulgada em 1950 veio na mesma época em que os horrores do nazismo foram descobertos após o fim da segunda guerra mundial, e a preocupação de que o Estado pudesse distinguir seus cidadãos com base em sua religião pode ter sido a causa da implementação da lei. Entretanto, durante a pesquisa, nenhuma evidência foi encontrada de que a motivação para tal lei tenha ocorrido por conta do Holocausto ocorrido em solo europeu.

Ainda sobre os EUA, é importante se aprofundar um pouco mais nas estatísticas sobre o protestantismo norte-americano, pois tal investigação será útil para a análise da realidade brasileira como veremos a frente.

Boa parte das pesquisas sobre o ambiente religioso nos EUA são baseadas no modelo “Denominalista” (STEENSLAND, 2000). Neste modelo, não há uma estrutura maior “Religião” que defina os grupos principais. Eles trabalham com o conceito de Congregação (menor grupo), Denominação e Tradição (ou Família). Assim, para tal modelo, a Igreja Católica, o Islamismo e o pequeno grupo de fiéis composto por duas ou três igrejas no meio oeste americano são considerados igualmente como denominações.

O Hartford Institute estima que existam 314 mil congregações cristãs nos EUA. Segundo a Enciclopédia Britânica, uma Congregação é “uma assembleia de

²⁷ Os dados brutos sobre os censos podem ser obtidos em <https://data.un.org/Data.aspx?d=POP&f=tableCode%3A28>

peçoas, especialmente um corpo reunido para o culto religioso ou para frequentar habitualmente uma determinada igreja”²⁸. Pode-se dizer que uma congregação é equivalente a uma paróquia no jargão católico.

Uma congregação pode estar ou não associada a uma denominação e uma denominação pode ser categorizada uma determinada “família” ou tradição cristã. Entretanto não existe um consenso sobre o que é uma denominação. Segundo Richey,

A denominação é um corpo eclesial voluntarista. A denominação é voluntária e, portanto, pressupõe uma condição de tolerância legal ou de facto e uma liberdade religiosa. É eclesial, um movimento ou corpo que se entende em ser uma igreja (ou movimento religioso) legítima e auto-suficiente. É uma igreja espontânea, um corpo que concede autenticidade de outras igrejas, mesmo que reivindique sua própria.²⁹ (RICHEY, 2013, p.2, tradução nossa).

Ainda na estrutura do modelo Denominalista, algumas denominações são formais e têm uma figura central, como a Igreja Católica. Já outras denominações se organizam sob a forma de “Convenções”, como as convenções batistas e outras denominações, que não respondem administrativamente ou teologicamente a nenhum órgão superior, mas se identificam dentro de um mesmo nome.

Efetivamente, as denominações (cristãs ou não) são fruto de divisões internas, os chamados cismas, e estas divisões não ocorrem apenas em função de disputas doutrinárias e teológicas, mas principalmente em função de fatores sociais, tais como raça, classe sociais e conflitos étnicos ou políticos (LIEBMAN; SUTTON; WUTHNOW, 1988). As questões sociais e políticas soltam aos olhos quando se analisa a categorização utilizada por pesquisas baseadas no denominalismo.

Steensland (2000) agrupa as denominações da seguinte forma: “Conservative Protestant”, “Mainline Protestants”, “African American Protestant”, “Catholic and Orthodox”, “Sectarian groups”, “Jews” e “Other religion”. Apesar das congregações afro-americanas terem uma história de luta e terem realmente um formato de culto próprio, separá-las, ou melhor, segregá-las das denominações “euro-americanas”

²⁸ No original: Congregation is an assembly of persons, especially a body assembled for religious worship or habitually attending a particular church.

²⁹ No original: Denomination is a voluntaristic ecclesial body. The denomination is voluntary and therefore presupposes a condition of legal or de facto toleration and religion freedom. It is eclesial, a movement or body understanding itself to be a legitimate and self-sufficient, proper church (or religious movement). It is a voluntary church, a body that concedes the authenticity of other churches even as it claims its own

apenas reforça que o histórico racista dos EUA também se mostra neste tipo de classificação.

De fato, o protestantismo é um componente central na sociedade e na cultura norte-americana tanto quanto o capitalismo. Diferentemente do Brasil, que nasceu católico, o protestantismo faz parte da história dos EUA desde sua fundação. Apesar da constituição americana ter a laicidade do Estado como um dos seus pilares, o protestantismo anda de braços dados com as grandes corporações e com os poderes legislativos, executivos e judiciário; foi peça central na propaganda ideológica durante a guerra fria e até os dias de hoje influencia a política e as eleições americanas (KRUSE, 2015).

Wuthnow (1994) ressalta especificamente a dualidade entre a religiosidade americana e o capitalismo:

A cada sábado e domingo, cerca de setenta e cinco milhões de americanos se reúnem em sinagogas, igrejas, mesquitas, templos, casas de reunião e salões de confraternização. A cada semana, pelo menos cinquenta milhões participam de pequenos grupos de convivência que estudam a Bíblia, oram e discutem questões espirituais. [...] Como uma nação de crentes, devemos esperar que nossos compromissos religiosos tenham um impacto decisivo em nosso comportamento econômico. Mas também estamos apaixonadamente comprometidos com o todo-poderoso dólar. Dedicamos a maior parte das horas em que estamos acordados para ganhá-lo e muito do resto do nosso tempo para encontrar maneiras de gastá-lo. (WUTHNOW, 2014, p.3, tradução nossa)³⁰

Desde o final do século XX, há dois fenômenos que tem chamado a atenção dos estudiosos em religião: o não-denominalismo e as *mega-churches*. O não-denominalismo tem ocorrido em um crescente número de congregações que não querem se identificar com nenhuma denominação específica. Procuram com isso ter uma maior liberdade teológica e também política (CHAVES, 2011).

Já as *mega-churches* são grandes congregações que contém um número elevado de fiéis e contam com uma estrutura organizacional de grande porte, com centenas de funcionários para receber milhares de pessoas em um culto dominical. Muitas vezes, contam com a figura central de um pastor com grande expressão

³⁰ Do original: Each Saturday and Sunday, some seventy-five million Americans flock to synagogues, churches, mosques, temples, meeting houses and fellowship halls. Each week at least fifty million participate in small fellowship groups that study the Bible, pray, and discuss spiritual issues. [...] As a nation of believers, we should expect our religious commitments to have a decisive impact on our economic behavior. But we are also passionately committed to the almighty dollar. We devote the bulk of our waking hours to earning it and much of the rest of our time to finding ways to spend it.

mediática e exercem grande influência na comunidade onde está instalada. (CHAVES, 2006; WILSON, 2010)

Tal como foi visto nos sistemas de classificações bibliográficas e nos relatórios do PRC, existe uma grande dificuldade em classificar as religiões de forma mais abrangente. E no caso do protestantismo, essa dificuldade é ainda maior. A partir desta constatação, busca-se na próxima seção articular tais classificações à realidade brasileira.

3.4 O PROTESTANTISMO NO BRASIL

Historicamente, os primeiros protestantes que chegaram ao Brasil a partir das invasões franceses e holandeses no século XVI e XVII. O primeiro culto foi realizado por franceses no Rio de Janeiro ainda em 1557 e durante a presença holandesa mais de 22 igrejas foram construídas na região do Recife (MATOS, 2011).

A permissão de que outras expressões religiosas que não fossem a católica pudessem se estabelecer na colônia portuguesa coincide com a abertura dos portos com a vinda de D. João VI. Entretanto, essa primeira leva de protestantes tinha como objetivo primário atender os estrangeiros que chegaram junto com a corte portuguesa e atos de evangelização eram muito tímidos. Com a imigração europeia durante o Império, novos missionários anglicanos, luteranos e metodistas chegaram ao Brasil, mas, novamente, o foco era atender as comunidades que imigravam.

A partir da década de 1910 chegam ao Brasil os primeiros missionários com o intuito de evangelizar a população. Foram fundadas à época a Assembleia de Deus por missionários suecos e a Congregação Cristã do Brasil por missionários italianos, ambas de influência pentecostal. Vale ressaltar que, apesar de sua origem europeia, estes missionários vieram dos Estados Unidos com a missão de evangelizar a população dentro dos preceitos pentecostais (MARIANO, 2004; BAPTISTA, 2007).

Um segundo grupo mais expressivo de protestantes pentecostais começou a se instalar nas décadas de 1950 e 60 com a fundação da Igreja do Evangelho Quadrangular e em seguida das igrejas Brasil Para Cristo, Deus é Amor e Casa da Bênção. Especificamente a Quadrangular passa a agir com maior força e conforme atesta Mariano, “além da ênfase na cura, essa vertente pentecostal notabilizou-se

pelo intenso uso do rádio e pela pregação itinerante com o emprego de tendas de lona” (2004, p.123).

Ao longo da segunda metade do século XX, o termo “evangélico” passou a ser utilizado pela população para identificar os protestantes. Esse nome surgiu a partir da própria autodenominação dos missionários americanos que procuravam se diferenciar dos demais “pois eram adeptos da corrente protestante, com raízes no movimento fundamentalista que desejava afirmar a sua fidelidade ao Evangelho e não à Ciência ou à razão humana” (CUNHA, 2019, p.15).

A partir do final da década de 70, o número de evangélicos começou a crescer expressivamente com o surgimento das igrejas denominadas pelo pesquisador Ricardo Mariano como neopentecostais. Essa nova vertente do protestantismo trazia em seu bojo uma doutrina originária dos EUA, na década de 40: a chamada Teologia da Prosperidade (TP) que se afastava do ascetismo originário dos metodistas e trazia uma nova crença que prezava os bens materiais. Em linhas gerais, define que, quanto maior a fé da pessoa, maior seria a sua prosperidade material (MARIANO, 1996, p.26).

O termo neopentecostal, apesar de não ser uma unanimidade entre os pesquisadores da área (SIEPIRSKI, 1997), passou a ser uma referência para este grupo de igrejas que surgiram a partir da década de 70 e que enfatizavam em suas práticas a TP, a cura divina e a guerra espiritual entre o bem e o mal.

Diferentemente das denominações tradicionais que eram fundadas por missionários estrangeiros, as igrejas neopentecostais foram criadas por pastores brasileiros que importaram suas bases teóricas em grande parte do pastor norte-americano Kenneth Erwin Hagin, que pregava desde a década de 40 no meio oeste norte-americano. Nos EUA, essa teologia conhecida como “*the health-and-wealth gospel*” se fortaleceu justamente nos momentos de crises econômicas tal como atesta Sousa:

Naturalmente, as crises sazonais da economia conduzem a um clima de insegurança, sobretudo entre aqueles que são mais diretamente afetados por elas. Portanto, não é possível entender a TP sem uma compreensão adequada do contexto em que ela se solidificou como doutrina; a TP representou a adequação de determinados segmentos do protestantismo norte-americano tanto ao rápido crescimento da economia como também às flutuações desta e ao desejo de participar das riquezas deste mundo sem desvincular-se da religião ou ser censurado por ela. Por isso, o enriquecimento e a prosperidade, agora ao alcance de muito mais pessoas, representam tanto uma acomodação a este mundo, como também o resultado de concessões divinas (SOUSA, 2011, p.224).

Em menos de 10 anos, estas igrejas que pregavam a TP se transformaram em corporações com uma enorme capacidade financeira, midiática e política. Dentre elas, se notabilizam a Igreja Universal do Reino de Deus (IURD), a Internacional da Graça de Deus (IIGD), a Comunidade Evangélica Sara Nossa Terra e a Renascer em Cristo. Turbinadas por tal ideologia que tornava o fiel “sócio” de Deus, numa sociedade na qual o fiel entrava com o dízimo e Deus entrava com as graças (principalmente de ordem material), estas Igrejas prosperaram. Somente a IURD cresceu 2.600% e no final da década de 90 se estimava que tinha mais de 2 milhões de adeptos (MARIANO, 2004, p.125).

A disputa pelo chamado mercado da fé originou inúmeras divisões entre as novas igrejas, como por exemplo a criação da Igreja Mundial do Poder de Deus, criada por antigos pastores da IURD. Sottani relata bem essas disputas:

Num mercado self service de opções religiosas (Mariano, 2013), a expansão acelerada do modelo da Universal nos anos 90 parece ter suscitado a proliferação de igrejas clones da IURD, com a mesma ênfase na Teologia da Prosperidade e a mesma forma de proselitismo, inclusive via mídia. Uma das principais opções em franco crescimento nas últimas décadas foi justamente a Igreja Mundial do Poder de Deus, do apóstolo Valdemiro Santiago, ex-bispo da IURD, que atraiu aos seus cultos parte significativa do rebanho iurdiano e de outros grupos do protestantismo histórico. Enquanto no Censo do ano 2000 ela ainda não tinha expressividade suficiente para ser contabilizada, em 2010 já aparecia com 315 mil fiéis declarados no censo, mais de quatro mil templos no país, e outros 250 no exterior. Talvez não por acaso as críticas à Mundial tenham se proliferado nos espaços da IURD, inclusive em reportagens da Rede Record, emissora de Edir Macedo, que a acusa de charlatanismo e enriquecimento ilícito (SOTTANI, 2016, p.30)

Também surgiram igrejas que buscavam “nichos de mercado”, tal como a Igreja Bola de Neve, que visava a juventude da classe média paulistana. Nessa

Igreja, os pastores são jovens, seus altares são pranchas de surf e a linguagem no púlpito totalmente diferente da empregada por pastores que pregavam na periferia das grandes cidades.

As igrejas pentecostais tradicionais também cresceram a reboque desta “onda evangélica” e a maior delas é a Assembleia de Deus (AD). Diferentemente das grandes organizações com um comando central como a IURD, a AD é organizada em formato de federação, onde as milhares de igrejas presentes em grande parte dos municípios brasileiros não respondem administrativamente ou financeiramente a nenhum órgão central, mas são filiadas à Convenção Geral das Assembleias de Deus no Brasil (MARIN, 2014). Tal como em outras denominações, vários cismas e divisões também ocorreram na Assembleia de Deus, tal como o Ministério Madureira³¹ e a Convenção Geral das ADs. Entretanto os pastores de ambas as convenções se identificam como pertencentes à denominação genérica “Assembleia de Deus”.

Ao longo dos anos 80 e até os dias de hoje, a presença dos evangélicos no Brasil vem crescendo inclusive na mídia e na cena cultural. Os evangélicos passaram a compartilhar um modo de viver diferenciado que se convencionou chamar de cultura gospel. Conforme atesta a pesquisadora Magali Cunha,

Este jeito moderno na forma de ser religioso trouxe ainda mais visibilidade para os evangélicos por meio de adesões de celebridades (artistas, esportistas, políticos e afins), que passam a se apresentar como fiéis deste segmento religioso. A cultura gospel é um fenômeno em curso que torna a presença evangélica na esfera pública mais intensa, ainda mais recentemente, por meio da ocupação cada vez maior de espaço nas mídias não-religiosas. (CUNHA, 2019, p. 26)

Entretanto esse crescimento não foi pacífico. Além das querelas e divisões internas e dos escândalos políticos, o discurso da guerra espiritual e do “nós contra eles” utilizado por parte das igrejas neopentecostais gerou diversos conflitos religiosos no Brasil. O número de casos de intolerância religiosa aumentou consideravelmente, em particular contra religiões de matrizes africanas mas também contra a Igreja Católica. O caso do pastor que chutou a imagem de Nossa Senhora Aparecida em um programa de televisão repercutiu muito no Brasil e no exterior³² e

³¹ Fundada em 1958, a Convenção Nacional de Madureira se originou no bairro de Madureira no Rio de Janeiro e na década de 80 criou sua própria convenção nacional denominada CONAMAD que, segundo a organização, conta atualmente com mais de 2 milhões de membros.

³² <https://www.uol.com.br/splash/noticias/ooops/2020/10/12/ha-25-anos-bispo-chutava-imagem-de->

os recentes casos de traficantes de drogas evangélicos que expulsaram religiosos de matriz africana de suas áreas de influência³³ são exemplos de como a competição pelo mercado da fé não se dá de forma pacífica.

Mesmo com os cismas e conflitos internos entre pastores das diversas denominações, além de escândalos financeiros e políticos, estas organizações continuam exercendo grande influência na política e no meio evangélico. Atualmente a IURD tem o controle da segunda maior emissora de TV do país (TV Record) e possui filiais em diversos países do mundo.

Da mesma forma, as demais igrejas que ainda não conseguiram seus próprios canais de TV, pagam vultuosas quantias por espaços nas programações de canais da televisão leigos para apresentarem seus cultos e oferecer os produtos de suas “marcas”. Se tornaram corporações com ramificações em vários setores econômicos e com uma grande gama de “produtos” que são oferecidos aos seus consumidores. Além de seu produto principal, que é a “salvação”, estas corporações oferecem produtos culturais como músicas, livros, shows, centros de apoio médico e uma infinidade de serviços que levam o selo (a marca) desses empreendimentos.

Para tais organizações, ao longo do restante da pesquisa, o termo **corporações pentecostais** será utilizado para referenciar especificamente este tipo de organização e para diferenciá-las das demais organizações religiosas, sejam elas protestantes ou não. O termo “corporação” aqui utilizado tenta sintetizar uma forma de ser particular de determinadas organizações religiosas, chamando atenção para o fato de que tais organizações contém os mesmos atributos das corporações relatados na obra “The Corporations”, de Joel Bakan. Em seu documentário (e posterior livro), o jurista define que “a corporação é uma instituição patológica, uma perigosa detentora do grande poder que exerce sobre as pessoas e as sociedades” (BAKAN, p.2)³⁴ e complementa:

nossa-senhora-na-tv.htm

³³ <https://www.agazeta.com.br/brasil/policia-prende-trafficantes-evangelicos-que-atacavam-terreiros-no-rio-0819>

³⁴ No original: the corporation is a pathological institution, a dangerous possessor of the great power it wields over people and societies.

Uma premissa chave é que a corporação é uma estrutura única e um conjunto de imperativos que direcionam as ações das pessoas dentro dela. É também uma instituição legal, cuja existência e capacidade de operar dependem da lei. O mandato legalmente definido da corporação é perseguir, incessantemente e sem exceção, seu próprio interesse, independentemente das conseqüências muitas vezes prejudiciais que possa causar a outros.(BAKAN, p.1, tradução nossa)³⁵

O uso de termos ligados ao mundo empresarial para se referenciar às organizações protestantes não é uma inovação desta pesquisa. Além do clássico de Weber já citado aqui, inúmeras obras se utilizam de analogias entre religião e capitalismo, principalmente nos EUA, berço das igrejas pentecostais. No livro *Pillars of faith: American congregations and their partners*, Ammerman (2005) trata da associação das congregações protestantes como fornecedores e sócios (*partners*) que fazem parte do complexo ecossistema protestante existente naquele país. O escritor americano metodista Russel Richey ressalta a semelhança entre as corporações e as denominações religiosas:

A denominação é uma criatura eclesial da modernidade, uma forma social emergente e próxima ao partido político, à imprensa livre e à livre iniciativa [...] De fato, não deve ser surpreendente que a denominação tem se assemelhado à corporação do século passado. Ela tem se assemelhado à forma comercial atual da época, e também à forma atual do partido político e da imprensa. (RICHEY, p.3)³⁶

Baptista (2015), ao analisar as formas de organização interna de cada um dos tipos de igrejas evangélicas, ressalta também o caráter societário das igrejas neopentecostais onde a figura dos sócios proprietários centraliza o poder e as decisões organizacionais:

³⁵ No original: A key premise is that the corporation is an institution- a unique structure and set of imperatives that direct the actions of people within it. It is also a legal institution, one whose existence and capacity to operate depend upon the law. The corporation's legally defined mandate is to pursue, relentlessly and without exception, its own self-interest, regardless of the often harmful consequences it might cause to others

³⁶ Do original: The denomination is a ecclesial creature of modernity, a social form emerging with and closely akin to the political party, the free press and free enterprise [...] Indeed, that the denomination has resembled the corporation for the last century should not be surprising. It has resembled the current business form of the day, and also the current form of the political party and of the press

O governo neopentecostal é assumidamente episcopal, mas seu paradigma é a grande empresa. No topo da pirâmide eclesiástica está o executivo, com características de proprietário, embora não assuma essa condição, de forma explícita. Esse executivo pode ostentar os títulos de apóstolo, bispo ou “missionário”, mas o conteúdo da função que exerce é sempre a de maior executivo da empresa-igreja, que, geralmente, fundou ou ajudou a construir. Esse tipo de governo não guarda semelhança com o episcopado das igrejas tradicionais protestantes nem com o episcopado medieval da Igreja Católica. Como afirmou Carreiro (2011, p. 29), o governo neopentecostal “é autoritário não porque é antigo, mas porque é moderno”. (BAPTISTA, 2015 apud BAPTISTA, 2021, p.8)

Assim, o termo “corporação pentecostal” tem o objetivo de sintetizar o interesse econômico, a busca por “consumidores”, a consolidação de sua marca no segmento religioso com intuito de aumentar a presença de tais marcas na sociedade. Como será relatado no capítulo seguinte, esse poder econômico se converteu em poder político, se consolidou e hoje tem grande influência na sociedade brasileira, tentando impor uma agenda de costumes, pautas morais e mantendo assim as benesses concedidas pelo Estado Brasileiro, tal como isenção de impostos e cargos em ministérios e secretarias.

Em resumo, para esse conceito que a presente pesquisa procura trabalhar, uma corporação pentecostal se diferencia das demais organizações religiosas a partir de três critérios bem claros: o domínio societário de um pastor ou pequeno grupo de pastores que controlam a organização; um braço midiático gerenciado por empresas externas à pessoa jurídica da Igreja, de tal modo que possam auferir lucro; e, por fim, a presença na política partidária com o objetivo de manter seguras suas atividades do ponto de vista jurídico, além de se beneficiar de ações do Estado.

A partir destes critérios, pode-se elencar as maiores das corporações pentecostais existentes no Brasil a partir da literatura existente. A maior delas é a IURD, seguida da IEQ, IIGD e Maranata. Entretanto, o processo de categorização das organizações que compõem o conjunto de igrejas que fazem parte das congregações das Assembleias de Deus e também das igrejas Batistas já não é tão trivial. O caráter federativo de tais organizações, no formato de “convenções” e “ministérios”, dificulta a identificação de concentração ou aglutinação de diversas pessoas jurídicas na mão de um único controlador.

Indiretamente, a partir da presença de determinados empresários na mídia, é possível listar algumas delas, tal como o AD-VEC, controlada por Silas Malafaia, e a

AD-MM, liderada pelo pastor Manoel Ferreira. Uma linha de pesquisa para tentar encontrar outras corporações pentecostais poderia trabalhar a partir destas figuras midiáticas e, através dos padrões de financiamento de campanha e da rede de empresas associadas a elas, encontrar outros empreendimentos que seguem esse “modelo de negócio”.

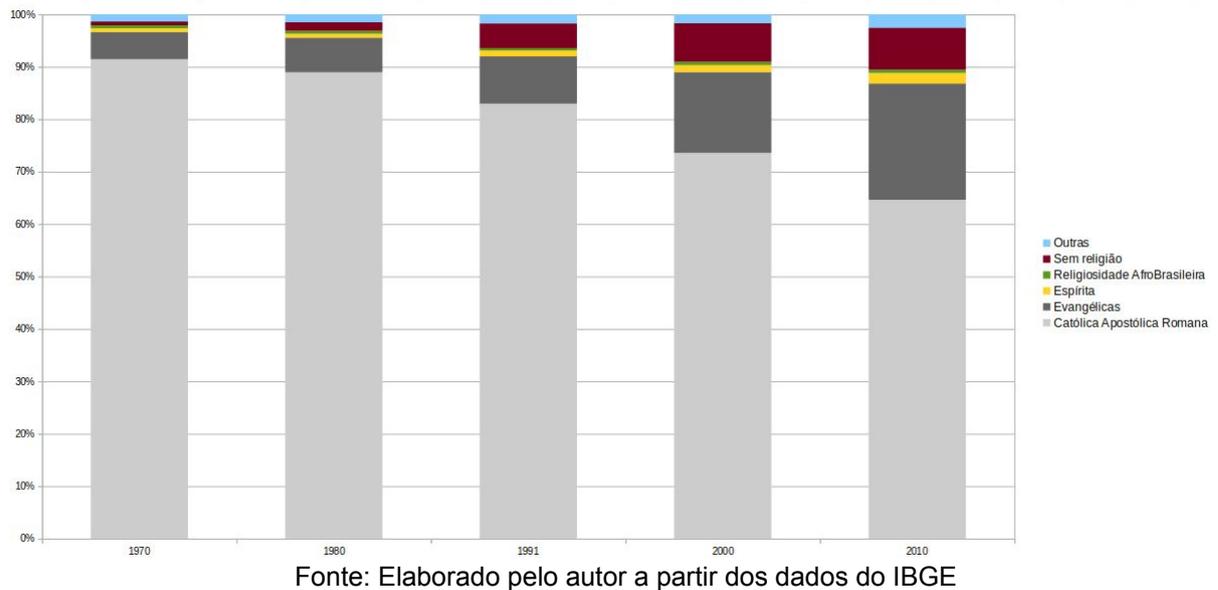
3.5 A CLASSIFICAÇÃO DOS EVANGÉLICOS NO CENSO DO IBGE

O Instituto Brasileiro de Geografia e Estatística (IBGE) é um instituto público da administração federal brasileira e o principal fornecedor de dados demográficos do país. Como declarado em seu site, "tais informações atendem às demandas de vários tipos de segmentos da sociedade civil, bem como dos órgãos em nível federal, estadual e municipal". E uma de suas principais missões é realizar o Censo Nacional que ocorre a cada 10 anos.

Desde 1872, o censo brasileiro inclui uma pergunta sobre Religião e até 1970 o catolicismo romano reinou absoluto com 91,8% da população declarando-se católica. Entretanto, após o Censo de 1991, houveram mudanças significativas na composição religiosa pela primeira vez, quando houve um crescimento no número de respondentes que se declararam evangélicos. Na década de 1980, o percentual da população que se declarou evangélica era de apenas 6,6% e subiu para 9,0% na década seguinte.

Nos anos 1990, houve crescimento mais expressivo e de acordo com o último censo de 2010, a população total brasileira era composta por 22,2% de evangélicos. Houve também um ligeiro crescimento daqueles que se declararam espíritas e de todas as outras religiões, mas o fato mais notável é que a população católica vem diminuindo em ritmo constante, tal como podemos ver no gráfico 1.

Gráfico 1 – Percentual da população residente, por grupos de religião



Após a divulgação dos resultados do Censo de 1990, houve um descontentamento por parte das igrejas evangélicas e também de pesquisadores da área em função da não categorização de determinados grupos ou denominação. Conforme atesta Mariano,

Cumprir observar que o Censo de 1991 não obteve êxito na identificação da denominação de considerável parcela dos evangélicos (...). Quer dizer, o IBGE não identificou a igreja de 8,5% dos evangélicos tradicionais, colocando-os na categoria classificatória 'evangélica tradicional não determinada'. Essa cifra mais que triplicou para o caso pentecostal. Nada menos que 31,9% deles (2.609.526 fiéis), quase um terço, não tiveram a denominação identificada, sendo arrolados na categoria 'evangélica pentecostal não determinada' (MARIANO, 2001: p.48).

Devido a este novo cenário, o IBGE viu a necessidade de uma categorização mais apropriada que contemplasse as novas religiões que surgiam no Brasil, e procurou a consultoria de uma ONG brasileira especializada em Religião chamada ISER (Instituto de Estudos Religiosos). Fundado nos anos 70, o ISER tem uma tradição não apenas na área da religião, mas também na defesa dos direitos humanos, da segurança pública e das questões ambientais. Publica regularmente duas revistas de renome na área e livros sobre religião, além de promover encontros ecumênicos entre os diversos representantes religiosos. Essa parceria já havia sido firmada no Censo de 1990, mas para o Censo de 2000, o trabalho do ISER foi estendido sob a supervisão da pesquisadora Clara Mafra.

O resultado desta consultoria não foi utilizado integralmente pelo IBGE mas trouxe uma categorização bem mais detalhada que a utilizada no Censo de 1990. Em função deste aprimoramento, pesquisadores de Ciências Sociais passaram a utilizar com mais frequência os dados do Censo do IBGE (CAMURÇA, 2014). A aplicação deste formato de categorização foi muito criticada na época, pois o IBGE não deu continuidade à consultoria firmada com o ISER quando não permitiu que os pesquisadores tivessem acesso à categorização que seria utilizada no relatório final do Censo.

Já no Censo de 2010, especificamente para o grupo evangélico, o IBGE trabalhou com três categorias principais: Protestantes de Missão, Pentecostais e Outros. O grupo categorizado como “de missão” se refere às igrejas tradicionais, principalmente as de origem europeia e de missão dos EUA. As denominadas Pentecostais são as fundadas no século XX no Brasil e que seguem as tradições pentecostais e as “Não determinadas” são aqueles que não se encaixam nas duas primeiras categorias ou que se declara evangélica mas não tem vinculação com nenhuma em específico. Os resultados agrupados pelos três grandes grupos podem ser vistos na tabela 9. Já na tabela 10, é possível visualizar o detalhamento por denominação.

Tabela 9 - População de Evangélicos – CENSO 2010

	População	%
Evangélica de Missão	7.686.827	18,18
Evangélica de origem pentecostal	25.370.484	60,01
Evangélica não determinada	9.218.129	21,80
Total	42.275.440	100

Fonte: O autor. Baseado no CENSO do IBGE 2010

Tabela 10 - Denominações cristãs não católicas em 2010

Evangélicas de Missão		
Igreja Evangélica Luterana	999.498	2,29
Igreja Evangélica Presbiteriana	921.209	2,11
Igreja Evangélica Metodista	340.938	0,78
Igreja Evangélica Batista	3.723.853	8,51
Igreja Evangélica Congregacional	109.591	0,25
Igreja Evangélica Adventista	1.561.071	3,57
Evangélicas de Missão - outras	30.666	0,07
Evangélicas de origem pentecostal		
Igreja Assembléia de Deus	12.314.410	28,16
Igreja Congregação Cristã do Brasil	2.289.634	5,24
Igreja o Brasil para Cristo	196.665	0,45
Igreja Evangelho Quadrangular	1.808.389	4,13
Igreja Universal do Reino de Deus	1.873.243	4,28
Igreja Casa da Benção	125.550	0,29
Igreja Deus é Amor	845.383	1,93
Igreja Maranata	356.021	0,81
Igreja Nova Vida	90.568	0,21
Evangélica renovada não determinada	23.461	0,05
Comunidade Evangélica	180.130	0,41
Evangélicas de origem pentecostal - outras	5.267.029	12,04
Evangélica não determinada	9.218.129	21,08
Não identificadas como evangélicas		
Outras religiosidades cristãs	1.461.495	3,34
Igreja de Jesus Cristo dos Santos dos Últimos Dias	226.509	0,52
Testemunhas de Jeová	1.393.208	3,19

Fonte: O autor. Baseado no CENSO do IBGE 2010

A partir destes números, é possível verificar um resultado expressivo para a categoria “Evangélica não determinada” que representa mais de 20% do total de cristãos não católicos. Se somarmos o número de “Evangélicas de origem pentecostal – outras” que aparece na Tabela 10, o percentual sobe para 33% do total de evangélicos. Este número confirma assim um grande pluralismo dentre os evangélicos.

Uma possibilidade sobre o elevado número de “Outros” pode estar relacionada ao mesmo fenômeno das igrejas independentes (que não estão ligadas a nenhuma denominação) tal como ocorre nos EUA. Devido aos escândalos políticos, aos casos de violência e intolerância religiosa praticada por determinados atores evangélicos, talvez algumas congregações (utilizando o termo norte-americano) possam ter escolhido um caminho independente do trilhado pelas

corporações pentecostais e os praticantes de tais igrejas não se identificaram com nenhuma das grandes categorias definidas pelo IBGE.

A categorização do IBGE também apresentou problemas na transparência de como os dados foram tabulados, tanto que enquanto o resultado do Censo Geral foi divulgado em 2011, o relatório específico sobre os dados de Religião só foi apresentado em 2012. A dúvida era que a tabulação dos resultados não conseguisse classificar as respostas de maneira adequada. Como o formulário do censo vinha com uma pergunta aberta: "Qual é sua religião?", havia dúvidas de como o entrevistado teria respondido e quão precisa era a resposta que o técnico reproduziu no questionário.

O ISER solicitou ao IBGE o banco de dados com as respostas individuais dos entrevistados (mesmo que anonimizados) para que fosse possível avaliar a exatidão da coleção. No entanto, o IBGE negou o pedido. Todo esse processo foi detalhado em artigos científicos pelos pesquisadores Marcelo Camurça e Clara Mafra, pesquisadores estes que participaram da consultoria fornecida pelo ISER para o IBGE (CAMURÇA, 2014).

Segundo Altmann (2012), o IBGE acertou no "atacado" e errou no "varejo", ou seja, conseguiu demonstrar as tendências gerais que mostravam um aumento da população evangélica, um aumento no número de espíritas e de pessoas que se consideram sem religião, mas errou ao tentar especificar a população evangélica e também não procurou determinar os casos de duplas (ou múltiplas) respostas, ou seja, os casos de sincretismo religioso que são tão comuns no Brasil.

É importante notar que o total de adeptos das igrejas pentecostais que foram discriminadas durante o censo e que não operam no formato federado (como a Assembleia de Deus) é muito pequeno em relação ao total de evangélicos, representando apenas 18% do total deste segmento religioso. Entretanto, como veremos no capítulo seguinte, a representação política e midiática destas corporações é bem maior que o indicado no Censo. Ou seja, proporcionalmente, o número de deputados de determinadas corporações pentecostais é maior que a população indicada no censo para o grupo que eles supostamente representam.

Efetivamente, ao categorizar os pentecostais a partir das principais organizações religiosas, o Censo de 2010 construiu gratuitamente um diretório que listava as principais corporações pentecostais e onde se localizavam seus

“consumidores”. Serviu, em grande medida, como uma fonte para informes publicitários utilizados por tais corporações para mostrar quão influentes eram e em que regiões seus adeptos estavam mais concentrados no território nacional.

Como pode ser visto ao longo do capítulo, há muitas possibilidades de examinar a classificação dos grupos evangélicos e pesquisas neste assunto podem abrir novas possibilidades de esclarecer melhor como esta população se move entre as várias denominações religiosas no campo evangélico. A análise de diferentes SOCs sobre um tema controverso como a religião, em muitos países e culturas, pode abrir espaço para novos e menos etnocêntricos esquemas de classificação, interculturais e multifacetados, reforçando assim as metas que Otlet e Ranganathan visavam ao criarem os seus sistemas de classificação bibliográfica.

4 O ESPÍRITO EVANGÉLICO NA POLÍTICA BRASILEIRA

Há coisas demais a que se quer dar o nome de 'religiosas'

Geertz, 2001, p.151

Política e Religião são elementos centrais da sociedade e estão integradas antes mesmo das dinastias dos faraós do Egito e dos jardins suspensos da Babilônia. Somente no século XVII, as revoluções francesa e americana trouxeram a novidade da laicidade do Estado. A laicidade e a separação entre política e religião constituem, como sustenta Bauer, “características do mundo ocidental. Elas possuem pouco ou nenhum sentido na maioria das outras sociedades” (1999, p. 8). O autor ressalta ainda que as próprias definições de fé, religião e igreja são fortemente influenciadas pela experiência cristã.

Na América Latina, a integração da religião na política se deu através da onipresença da Igreja Católica apoiando inicialmente as monarquias coloniais e em seguida as elites oligárquicas que estiveram no poder após a consolidação da independência dos países que surgiam no continente. Até os dias de hoje, a Igreja Católica continua exercendo grande influência seja através de uma presença simbólica com cruxifixos espalhados na maioria dos tribunais do Brasil seja através de leis em países como Argentina, Bolívia e Costa Rica que adotam a Igreja Católica como uma religião de Estado (ORO, 2011).

Muito em função da mão forte da Igreja Católica, a presença dos evangélicos na vida política no Brasil sempre se deu de forma tímida até 1940. Somente com a chegada de um grande contingente de missionários americanos e de uma constituição que garantia uma maior liberdade religiosa, foi possível o crescimento da presença evangélica nos meios de comunicação e, em parte, graças a este fator, na política nacional.

Nos últimos 40 anos, pesquisadores e analistas políticos vem acompanhando o crescimento da população evangélica e o aumento da influência de seus líderes no campo político. Não apenas os deputados ou senadores eleitos, mas pastores com grande poder midiático passaram a ter grande presença na cena política, não só na

defesa de uma pauta conservadora nos costumes, mas também influenciando a escolha de candidatos à presidência da república e na convocação de grandes manifestações populares, como a Marcha para Jesus e outras semelhantes.

Entretanto, conforme foi visto no capítulo anterior, os evangélicos não representam uma população homogênea e fazem parte de centenas, senão milhares de denominações, igrejas e movimentos. Será que todos os evangélicos comungam desta pauta conservadora ou tal pauta fica concentrada em determinadas vozes com maior poder midiático e político?

Este capítulo procura resumir como a literatura acadêmica analisou a atuação dos evangélicos na política brasileira, realizando inicialmente um recorte histórico de como esta presença se manifestou nas eleições e quais os principais atores políticos que surgiram desde a redemocratização. Em seguida, procura mostrar como os movimentos evangélicos e corporações pentecostais investiram em uma atuação mais ativa dentro dos partidos políticos e como é a composição ideológica desta atuação.

Para o restante da pesquisa, cabe explicitar uma questão metodológica importante: a pesquisa assume que um candidato é evangélico quando este se autodeclara como tal. Desta forma, evita-se associar apenas os candidatos apoiados formalmente por uma determinada organização ou grupo e, por outro lado, evita-se falsos positivos, ou seja, candidatos ligados a grupos evangélicos, mas que não se declaram como tal.

Além disso, a pesquisa procura seguir a tipologia sugerida por Lacerda (2017) que agrupou os candidatos evangélicos em três categorias: os de origem evangélica mas que não são apoiados formalmente por suas igrejas; os candidatos evangélicos oficiais que saem dos quadros das corporações pentecostais para se candidatarem e os políticos tradicionais que se tornam evangélicos para se beneficiar das “redes” que apóiam as candidaturas evangélicas. Sempre que houver necessidade, cada um dos três tipos será apontado durante o capítulo.

Por fim, cabe ressaltar que a pesquisa segue o entendimento de Baptista (2007) de que não existe uma “bancada evangélica” que vota em bloco e que tem uma coordenação unificada. Existe sim um conjunto de políticos evangélicos que, para determinadas pautas, se unem para reforçar sua identidade no Congresso.

Desta forma, quando utilizamos o termo “bancada evangélica”, estamos nos referindo ao conjunto de congressistas evangélicos.

4.1 HISTÓRIA DOS EVANGÉLICOS NA POLÍTICA

A eleição do primeiro representante evangélico na Câmara Federal ocorreu em 1933 com a eleição do pastor metodista Guaracy Silveira pelo PSB e que foi reeleito na constituinte de 1946 já pelo PTB. Participou ativamente do processo de autonomia da Igreja Metodista no Brasil e lutava por pautas liberais e progressistas, como a tese de que o ensino religioso não fosse obrigatório nas escolas (OLIVEIRA, 2008).

Até 1987, a presença dos evangélicos foi tímida tanto que ao longo das 13 legislaturas desde 1933 até 1987, apenas 50 deputados evangélicos foram eleitos. Um dos fatores que impediam o acesso dos evangélicos na esfera político-partidária era uma feroz vigilância da Igreja Católica (não só sobre os evangélicos, mas também sobre os espíritas kardecistas e adeptos do Candomblé) através da restrição do acesso aos meios de comunicação, já que constantemente a criação de rádios e jornais evangélicos era impedida por membros da Igreja Católica e seu clero através de pressões sobre os órgãos que concediam tais concessões (FAJARDO, 2011).

A tímida presença no Legislativo não significava, entretanto, que as lideranças evangélicas já não participassem da vida política. Apesar do embate entre alas progressistas e lideranças conservadoras evangélicas, uma parcela considerável destas lideranças apoiou a ditadura civil-militar implantada a partir de 1964 e se beneficiou deste apoio, recebendo suas primeiras concessões de rádio (ALMEIDA, 2016).

O fortalecimento de candidatos e grupos políticos patrocinados e ligados a algum segmento evangélico se deu concomitantemente ao surgimento de corporações midiáticas que serviam de suporte para a transmissão da mensagem das igrejas que surgiram na década de 70 em diante. Um dos primeiros exemplos mais marcantes foi do pastor Nilson Fanini, que tinha um programa de TV na Rede Bandeirantes e que servia explicitamente como canal de propaganda para o PDS

(Partido Democrata Social³⁷) nas eleições de 1982 (ASSMANN, 1986, p.84). Foi dono da TV Rio e tinha trânsito livre no Palácio do Planalto durante a ditadura militar.

Na eleição federal de 1982, apenas quatorze evangélicos foram eleitos para a Câmara dos Deputados, com seis representantes da Igreja Batista, dois metodistas e um da Assembleia de Deus. Já em 1986, com a primeira eleição geral após a redemocratização, diversas igrejas pentecostais entraram na política parlamentar trocando a postura tradicional de "os crentes não se metem com a política" por uma nova política de "irmão vota por irmão", buscando assim o envolvimento ativo nas eleições. Como resultado, os evangélicos elegeram trinta e quatro deputados para a Assembleia Constituinte e houve uma mudança na composição desta representação pois os Pentecostais compuseram dois terços deste total (FREESTON, 1996).

Ainda na eleição de 1986, surgiu a figura do candidato oficial de uma determinada corporação pentecostal que, apesar de não se identificar como candidato de uma dada igreja, era financiado e apoiado por tal. Um dos exemplos marcantes foi a do candidato Jefferson Campos (PSD), pastor da Igreja do Evangelho Quadrangular (LACERDA, 2017).

Ressalta-se, entretanto, que nem todos os candidatos evangélicos seguiam esse modelo corporativo. Várias candidaturas eram independentes de suas igrejas de origem mas se autodeclaravam evangélicos. Exemplo deste tipo de candidatura é a de Benedita da Silva (PT), que era da Assembleia de Deus mas não era apoiada explicitamente por este grupo religioso.

Já nas eleições presidenciais de 1989 quando Collor foi eleito, houve intensa movimentação dos evangélicos na disputa presidencial. Inicialmente houve uma movimentação de empresários e pastores pela candidatura de Íris Resende, ex-ministro do governo Sarney, evangélico e representante do PMDB de Goiás. Entretanto, o PMDB desprezou tal articulação ao escolher o Senador Ulysses Guimarães como candidato à presidência. Na falta de um candidato "puro sangue", a IURD e a Assembleia de Deus passaram a declarar explicitamente o apoio a Collor. O próprio Edir Macedo declarou: "Após orar e pedir a Deus que indicasse uma pessoa, o Espírito Santo nos convenceu de que Fernando Collor de Mello era o escolhido". Já outras corporações como a IEQ se limitaram a pedir que seus fiéis

³⁷ O PDS era formado pelas mesmas lideranças políticas que formavam a ARENA, o partido que representava a ditadura militar no Congresso Nacional e que no futuro se tornou o PFL e atualmente está representado majoritariamente pelo DEM e PP.

não votassem em candidatos de esquerda mas não declararam o voto a qualquer candidato (MARIANO; PIERUCCI, 1992).

Apenas uma pequena parcela de igrejas evangélicas apoiou a candidatura de Lula (PT). Segundo Mariano e Pierucci (1992), um dos apoios mais contundentes foi o manifesto "Uma opção ético-política, Manifesto de Cristãos do Comitê Inter Religioso" em apoio à Frente Brasil Popular, assinado por várias lideranças protestantes e católicas como as igrejas Metodista, Episcopal, Presbiteriana Unida, Presbiteriana do Brasil, Congregacional, Evangélica de Confissão Luterana, Evangélica Luterana do Brasil e Batista. Um fato irônico é que Silas Malafaia, um dos mais ferrenhos críticos da esquerda nos dias de hoje apoiou Lula no segundo turno (CAMPOS, 2002).

Apesar do apoio das corporações pentecostais e da vitória de Collor nas eleições presidenciais, o número de candidatos evangélicos eleitos no congresso foi menor que o das eleições de 1987. Os deputados estiveram na base de apoio do governo Collor, mas a maioria votou pelo impeachment (DA SILVA, 2017).

Em 1990, foi criado o Movimento Evangélico Progressista (MEP) com o intuito de fazer frente a bancada evangélica conservadora eleita na Assembleia Constituinte e às corporações pentecostais que já tinham se envolvido em “práticas não-republicanas” tal como relata a pesquisadora Magali Cunha (2019):

A atuação da primeira bancada no Congresso Constituinte foi marcada pelo fisiologismo e pela histórica farta distribuição de estações de rádio e canais de TV aos deputados evangélicos (com estreita relação com a ampliação de presença de evangélicos nas mídias no período), entre outras práticas de corrupção (p.39)

Por conta deste desgaste e de escândalos financeiros das corporações pentecostais, foram criados outros movimentos além do MEP que buscavam se diferenciar de tais práticas. Foram criados o Movimento Cristão Democrático de Centro (MCDC) e a Associação Evangélica Brasileira (AEVB) que representavam o centro e a esquerda política (CUNHA, 2019).

Para se opor à AEVB, as corporações pentecostais criaram o Conselho Nacional de Pastores do Brasil (CNPB) em 1993 e já nas eleições de 1994 apoiaram a candidatura de Fernando Henrique Cardoso (PSDB) à presidência apesar da controvérsia de que ele teria dito que era ateu em uma entrevista de grande repercussão. Ainda nas eleições de 1994, o número de deputados evangélicos

eleitos retornou ao patamar de 34 deputados e o predomínio dos candidatos ligados às corporações pentecostais se manteve.

Nas eleições de 1998, com as benesses do governo FHC aos evangélicos sob forma de concessões de rádio e TV, o apoio ao presidente já se deu logo no primeiro turno. Nessas eleições, a bancada evangélica deu um salto para 55 deputados e a IURD passou pela primeira vez a ter quase o mesmo número de congressistas que a AD. Entretanto o crescimento de todos os segmentos evangélicos cresceu naquele ano.

Ressalta-se também que os deputados evangélicos passaram a se concentrar em menos partidos e também passaram a ter o controle sobre o PL, o PSC e PSL. Em 2001, o Partido Liberal (PL) era o que mais abrigava os deputados evangélicos na Câmara Federal, com o predomínio de representantes da IURD que tinha escolhido o partido em função da falta de uma liderança forte que pudesse se opor ao seu projeto de poder e também pela presença de diretórios em boa parte do país. Em 2002, o PL se aliou ao PSL (Partido Social Liberal) que também era dominado pela IURD e passou a controlar 29 votos no Congresso (FONSECA, 2008).

Nas eleições de 2002, Lula amplia sua base de apoio quando escolhe o seu vice José Alencar que se filia ao PL, partido ligado a setores evangélicos e principalmente à IURD. Perdeu, entretanto, o apoio do PSB, que lança Anthony Garotinho como candidato que “utiliza sua identidade religiosa evangélica em campanha e, como resultado, conquista o voto de 51,3% dos evangélicos e de somente 6% dos católicos no 1o turno” (TREVISAN, 2013, p.5).

No segundo turno de 2002, apenas a IEQ oficializa o apoio a José Serra (PSDB). A grande maioria das lideranças evangélicas declara apoio a Lula, inclusive o candidato evangélico derrotado Anthony Garotinho:

Em um concorrido evento numa churrascaria na zona norte do Rio de Janeiro, que reuniu cerca de 900 pessoas, Lula obteve o apoio de representantes de várias igrejas, incluindo a Metodista, Batista, Sara Nossa Terra, Igreja Universal e pastores da Assembléia de Deus, que não seguiram a decisão das duas principais convenções da igreja de apoiar Serra. Ao discursar para a platéia de evangélicos, Lula recorreu a imagens bíblicas para rechaçar a campanha do adversário, que insiste na tese do medo para desconstruir sua candidatura às vésperas do segundo turno das eleições (UOL Notícias, 17/10/2002)³⁸.

³⁸ <https://noticias.uol.com.br/inter/reuters/2002/10/17/ult27u27501.jhtm>. Acesso em 01/05/2021

Nas eleições de 2002, a bancada evangélica do Congresso teve um aumento considerável passando a ter sessenta e oito deputados e quatro senadores. Em 2003, é criada a primeira Frente Parlamentar Evangélica, que se tornou a primeira organização formal dos parlamentares evangélicos dentro do Congresso, contendo não apenas os representantes da bancada evangélica, mas também deputados que apoiavam as reivindicações trazidas por tal frente.

Já em 2006, em função de dois grandes escândalos políticos que envolveram vários deputados evangélicos (Máfia das Sanguessugas e o Mensalão), houve uma redução considerável no número de evangélicos no Congresso (REICH e DOS SANTOS, 2013). Dos sessenta e oito deputados eleitos em 2002, a bancada passou a ter quarenta e nove deputados em 2006.

No âmbito da disputa presidencial, a grande maioria das lideranças evangélicas manteve o apoio a Lula. Somente parte da Assembleia de Deus (CGAD) declarou voto em Alckmin (PSDB) enquanto que o Ministério Madureira (CONAMAD), que nas eleições de 2002 tinha optado pelo PSDB, declarou apoio a Lula. A IURD não só manteve o apoio à reeleição de Lula como o seu vice José de Alencar migrou para o PRB, partido esse controlado naquele momento pela IURD (MARIANO et al, 2007).

Da mesma forma, na primeira eleição de Dilma Rouseff (PT) em 2010, as lideranças evangélicas se dividiram no apoio aos candidatos à presidência. Com Marina Silva (PV) na disputa, uma fatia considerável do eleitorado evangélico optou pela candidatura de uma evangélica, mesmo tendo ela um histórico progressista e de esquerda. Já a IURD optou por apoiar a candidatura de José Serra (PSDB) e, para complicar o quadro, o Pastor Everaldo (líder do PSC) acabou declarando o voto à coligação petista apesar de estar inclinado inicialmente a apoiar José Serra,. A mudança de “lado” de Everaldo foi detalhada na imprensa que apurou que haveria uma promessa de cargos após as eleições e o PT consolidou a aliança mesmo antes dos resultados das eleições depositando 4,7 milhões de reais para a campanha do PSC (CUNHA et al, 2017, p.26).

A bancada evangélica em 2010 deu um saldo de 49 para 71 deputados. Contabilizando somente os candidatos a deputado federal eleitos, os evangélicos obtiveram mais de 7 milhões de votos em todo o país. A Assembleia de Deus se manteve na liderança com vinte deputados eleitos seguida de deputados batistas e

presbiterianos. Nessa eleição, a IURD teve um desempenho bem abaixo dos demais “concorrentes” evangélicos, elegendo apenas 6 deputados.

Logo no primeiro mandato da presidenta Dilma, grande parte das lideranças evangélicas passou a ser de oposição ao governo Dilma. Esse movimento foi capitaneado pelo presidente da Câmara Eduardo Cunha (PMDB) (que vinha tendo embates constantes com o governo) e por várias figuras de grande expressão midiática no meio evangélico, como os pastores Silas Malafaia, o próprio Pastor Everaldo e Marco Feliciano. As entrevistas de Everaldo retomam o discurso anticomunista mas também explicitam o real motivo para a quebra da aliança que foi a negativa do governo Dilma em conceder os cargos desejados pela base de apoio do Pastor Everaldo. Cunha et al. relatam uma das falas de Everaldo:

“Nós decidimos ter candidato próprio em Janeiro de 2011, porque o governo do PT aparelhou o Estado para atender seus interesses partidários. Nós não indicamos nem um garçom. (...) O governo deixou de ser dos brasileiros para ser de um partido só, para a hegemonia de um sistema que está vencido no mundo. Nós não queremos que o Brasil se torne uma Cuba nem uma Venezuela.” (CUNHA et al, 2017, p.26)

Nas eleições de 2014, os setores evangélicos contaram com 2 candidatos para a corrida presidencial: Marina Silva (PV) e o Pastor Everaldo (PSC), sendo que a primeira por um motivo circunstancial pois era vice na chapa de Eduardo Campos e assumiu a disputa em função da morte do candidato pernambucano. Entretanto a candidatura de Everaldo foi um fiasco. Apesar da candidatura do Pastor ser mais conservadora que a de Marina já que esta tinha o histórico de defender causas progressistas e apesar das lideranças evangélicas optaram pela candidatura de Everaldo, Marina obteve 20 milhões de votos a mais que Everaldo.

Nessas eleições a bancada evangélica no Congresso Nacional também cresceu, mas não tão acentuadamente quanto nas eleições de 2010. Saiu de setenta e um para noventa e três deputados, mantendo os mesmos quatro senadores. Entretanto, o número de votos de alguns deputados evangélicos foi muito expressiva, principalmente no Rio e São Paulo. Em São Paulo, o Pastor Marco Feliciano (PSC) foi o terceiro deputado mais votado, ficando atrás apenas do apresentador de TV Celso Russomano (PRB) e do comediante Tiririca (PR) e bem à frente de nomes de peso como Bruno Covas (PSDB), Carlos Sampaio (PSDB) e Paulinho da Força (SD).

Já no Rio de Janeiro, as segundas e terceiras posições foram ocupadas por Clarissa Garotinho (PR) e Eduardo Cunha (PMDB) e a soma dos votos a candidatos evangélicos eleitos no Estado do Rio de Janeiro chegou a 26% do total de votos válidos, metade deles apenas para os candidatos ligados às corporações pentecostais, confirmando a força de tais organizações no Estado.

Em 2018, após o golpe que removeu Dilma Rouseff da presidência (BASTOS, 2016; OLIVEIRA, 2016), a onda conservadora (ALMEIDA, 2019) se intensifica com uma grande campanha midiática contra o campo da esquerda que culmina com a condenação e prisão do ex-presidente Lula, impedindo assim que ele pudesse concorrer ao pleito. Com a estratégia de não comparecer aos debates e fortalecido com o atentado contra ele, Bolsonaro supera os outros candidatos da Direita e é eleito no segundo turno das eleições.

Contando com o apoio de grande parte das lideranças evangélicas, sua campanha fez questão de fazer crer que ele tinha se tornado evangélico e passa a mostrá-lo com frequência em cultos e eventos religiosos. Tais demonstrações culminam escatologicamente com o seu “batismo” no Rio Jordão em Israel pelo mesmo Pastor Everaldo que em 2014 tinha disputado as eleições presidenciais (ALMEIDA, 2019) e que viria a ser preso por corrupção em 2021.

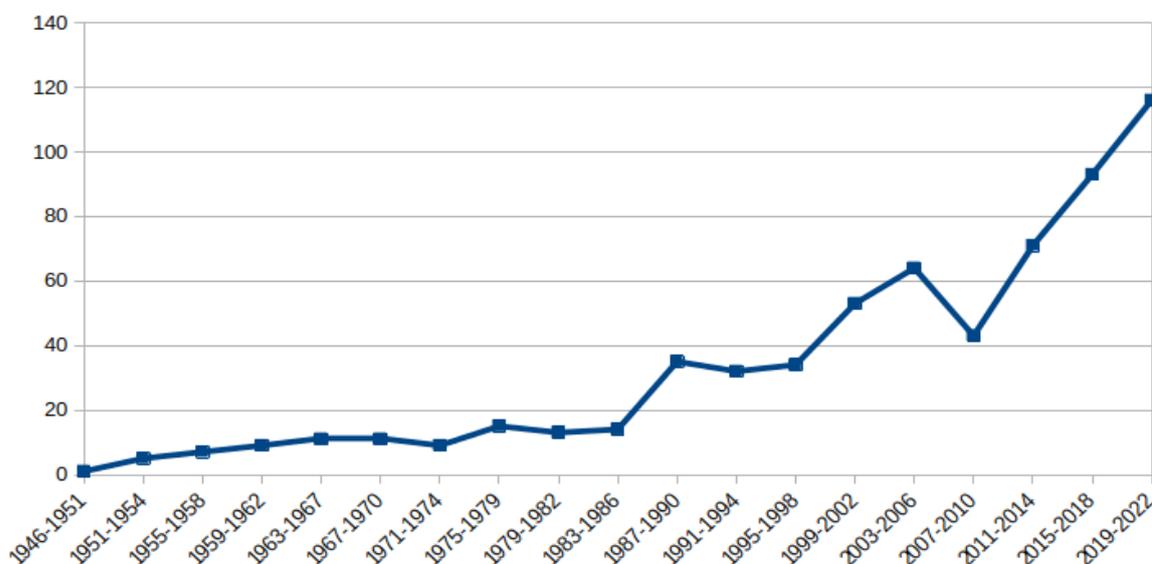
Freston (2020) propõe que Bolsonaro seria “o primeiro presidente pan-cristão” pois fizeram dele “um candidato híbrido ideal reunindo as vantagens eleitorais da identidade evangélica, mas evitando as desvantagens”:

A eleição de Bolsonaro foi um movimento amplo, o qual boa parte dos líderes evangélicos abraçou tardiamente e sem a centralidade que eles mesmos se atribuíam. Podemos prever que não terão nenhum pudor em repudiá-lo, se necessário, com o mesmo vigor com que o apoiaram. Afinal, muitos deles, antes, haviam apoiado os governos do PT. E se a identificação ideológica era fraca naquele caso, não devemos presumir necessariamente que seja forte agora. (FRESTON, 2020, p.373)

Já no Congresso, a onda conservadora auxiliou sobremaneira a eleição dos candidatos evangélicos. A bancada evangélica passou de noventa e três para cento e dezesseis deputados e elegeu o número recorde de quatorze senadores. Os três filhos de Bolsonaro tiveram as maiores votações em seus estados, sendo um para o Senado, um para a Câmara Federal e um como Deputado Estadual. Os deputados eleitos se dividiram em vinte e quatro dos trinta e cinco partidos que disputaram as

eleições e o estado que mais elegeu deputados evangélicos continuou sendo o Rio de Janeiro. O crescimento da bancada evangélica desde 1946 até 2018 pode ser visto no gráfico abaixo:

Gráfico 2 – Total de Deputados Federais Evangélicos de 1946 a 2019



Fonte: Elaborado pelo autor³⁹

Ao longo de todo o levantamento bibliográfico sobre a bancada evangélica, foram encontradas várias discrepâncias entre os números apresentados pelas pesquisas analisadas. Como a presente pesquisa tinha como base metodológica nominar e categorizar cada um dos candidatos evangélicos e a organização religiosa à qual ele ou ela estavam associados, a correta quantificação dos eleitos era um passo fundamental.

Um dos exemplos de discrepância de dados foi em relação às eleições de 1998. Fonseca (2004) indicou que foram eleitos quarenta e nove deputados, Rodrigues-Silveira e Cervi (2019), cinquenta e um. Como saber qual a soma correta? Na verdade, todos esses números podem estar corretos, pois cada pesquisador pode ter escolhido uma metodologia distinta para determinar quem pode ser considerado evangélico ou não. Mas, na prática, como a metodologia e a

³⁹ Só foram considerados os deputados evangélicos que tomaram posse no início da legislatura, ou seja, suplentes que tomaram posse ao longo da legislatura e deputados que desistiram ou perderam o mandato não foram computados nesta totalização.

listagem dos deputados nem sempre foi explicitada nas referidas publicações, não houve como fazer tal comparação.

De todas as publicações encontradas durante a pesquisa, somente seis explicitavam os nomes de cada um dos deputados e senadores que faziam parte de suas análises: Freston (1993), Gonçalves (2011), Batista (2013), Da Silva (2017), Lacerda (2018) e Cunha (2019). Fábio Lacerda (2018) foi o único de todos os pesquisadores a disponibilizar a fonte de dados em formato legível por máquina (em formato de planilha eletrônica) e em um repositório de dados público.

Inicialmente, a presente pesquisa estava assumindo que os resultados de Lacerda seriam suficientes para identificar os candidatos evangélicos, mas dada as discrepâncias acima citadas, optou-se por validar os dados da planilha disponibilizada. E após a verificação, verificou-se que a base do pesquisador não continha vários políticos evangélicos de renome e desta forma, foi necessário rever toda a planilha, incluindo novos nomes e aprimorando a classificação das organizações associadas aos políticos em questão.

A coleta se deu da seguinte forma: uma planilha eletrônica foi criada e para cada publicação científica que continha a lista dos candidatos evangélicos, os nomes dos candidatos eram adicionados. A planilha continha 6 colunas:

1. ano em que o candidato foi eleito;
2. nome de urna ou seu nome completo;
3. partido a que pertencia no momento da eleição;
4. unidade da federação;
5. denominação/igreja de origem ou que frequentava;
6. Fonte ou Evidência: em qual artigo científico ou fonte jornalística houve a evidência de que o nome da(o) parlamentar era identificado como pertencente a determinada denominação/igreja.

A partir da listagem resultante, foi realizada uma busca primeiro no site da Câmara e do Senado para confirmar se os dados de cada um dos deputados e senadores estavam corretos. Em seguida, buscou-se a igreja ou denominação do congressista. A busca inicial foi feita no Google e na Wikipedia pelo nome do candidato e, a partir de resultados que remetessem especificamente a artigos científicos, matérias jornalísticas em veículos de imprensa e em sites de análise

política como o DIAP e o Congresso em Foco, buscava-se a igreja a qual o candidato pertencia.

Também se optou por estender o período de coleta para um período anterior a 1998 pois dessa forma, seria possível encontrar o histórico de determinados políticos que não apareciam nas listagens existentes. Conforme já foi dito no início do capítulo, foi considerado deputado evangélico todo o candidato que se declare evangélico e não apenas os candidatos oficiais ou ligados às igrejas ou corporações pentecostais.

Durante a pesquisa, um outro revés impactou a busca dos candidatos. O DIAP, que era uma fonte muito utilizada nos artigos científicos e conferências sobre o tema, reformulou todo o seu site e todos os links (URLs) referenciados em artigos científicos deixaram de existir fazendo com que os dados não pudessem mais ser conferidos. Houve uma tentativa de fazer a busca do histórico no site *Wayback Machine*⁴⁰, mas as páginas não foram encontradas. Foram enviados e-mails para o DIAP mas não houve nenhum retorno.

Foram feitas também consultas por e-mail para alguns dos pesquisadores mas apenas dois responderam aos questionamentos. Os pesquisadores Fábio Lacerda e Dirceu Gerardi foram muito solícitos e enviaram os dados que tinham à disposição para que uma segunda conferência pudesse ser realizada. É interessante citar que ambos os pesquisadores procuraram trabalhar não só com os eleitos mas também com os não-eleitos pois suas pesquisas tinham o interesse em saber qual a taxa de sucesso de candidatos evangélicos, ou seja, do total de candidatos evangélicos, que concorriam às eleições, quanto eram eleitos e se estes eram eleitos com mais facilidade que os demais candidatos.

Para tal, utilizaram uma técnica engenhosa para obter os candidatos não-eleitos: buscaram os nomes de urna que continham os termos “Pastor”, “Bispo”, “Presbitero”, “Irmão”, etc. Obviamente a técnica não conseguia obter todos os evangélicos mas através dela, era possível filtrar um número considerável de candidatos que deliberadamente queriam ser reconhecidos como religiosos.

Efetivamente, o processo de coleta atrasou a pesquisa em mais de três semanas. Foram encontrados vários falsos positivo, ou seja, deputados que

⁴⁰ O site Wayback Machine procura copiar os dados de todos os sites da Internet como forma de guardar o histórico de várias épocas do site para que em caso de destruição do site, o conteúdo original possa ser consultado. O link para o serviço é <https://archive.org/web/>

estavam indicados em alguns artigos como evangélicos mas que não eram evangélicos (ou por serem notadamente católicos ou por não ter sido encontrada nenhuma evidência de qual igreja pertenciam). Notou-se também que vários deputados evangélicos do campo progressista não apareceram em algumas referências. O caso mais contundente foi o do deputado mineiro Paulo Delgado que é declaradamente batista e que só foi aparecer em apenas uma das referências como evangélico. Em função da metodologia utilizada, é possível também que outros deputados evangélicos não estejam presentes pois se eles não constaram em nenhuma das bibliografias estudadas, eles também não estarão presentes aqui.

Foram encontrados também vários falsos-positivos, ou seja, deputados que estavam indicados como evangélicos, mas que definitivamente eram católicos ou não eram ainda evangélicos quando foram eleitos pela primeira vez. São exemplos destes casos, a deputada Chris Tonietto que é católica mas aparece em várias listas como evangélica e a deputada Fátima Pelaes que se tornou evangélica apenas em seu terceiro mandato.

Ao longo da pesquisa, confirmou-se a tipologia de Lacerda (2017) que já indicava que haviam de três tipos de candidatos evangélicos: os que já eram evangélicos e entraram para a política sem o apoio de suas igrejas, os que faziam parte da estrutura organizacional das corporações pentecostais e foram escolhidos para representá-las no Congresso e os candidatos que se tornaram evangélicos após o início da carreira política com o intuito de se aproveitar dos benefícios que tal posicionamento religioso pudesse proporcionar para o seu sucesso nas urnas. Exemplo maior deste último tipo de político foram os filhos e o próprio presidente Bolsonaro que se aproximou do campo evangélico durante a campanha eleitoral em 2018.

O resultado da coleta dos deputados evangélicos foi incorporado à base de dados relacional onde a mineração de dados seria realizada e também publicada no repositório de dados do LARHUD para que outros pesquisadores tivessem acesso ao conteúdo e possam contribuir antes mesmo do fechamento da presente pesquisa. Houve intenção também de que o processo de montagem desta base de candidatos evangélicos fosse relatado em um artigo científico específico para que a planilha resultante pudesse ser divulgada em periódico especializado.

4.2 AS CORPORAÇÕES PENTECOSTAIS E OS PARTIDOS POLÍTICOS

Desde a aparição dos políticos ligados às corporações pentecostais, as pautas apresentadas por este grupo estão, em geral, ligadas ao conservadorismo e ao espectro política da direita. Antes da redemocratização, havia ainda alguma representação do campo progressista entre os deputados evangélicos no Congresso contabilizada pela presença de presbiterianos (IPB) e Luteranos (IECLB), mas já no primeiro parlamento eleito sob a nova constituição de 1988, houve uma maior participação de parlamentares de direita que se uniram com o intuito de lutar contra questões que atentassem contra a moral cristã e também para defender interesses comerciais, principalmente em relação a concessão de canais de TV e rádio.

Posto que esta predominância do campo de direita é um fato que já está consolidado na literatura, cabe apenas questionar em que grau ela ocorre, ou seja, todos os deputados evangélicos são de direita ou qual o percentual de deputados progressistas estão presente entre os evangélicos eleitos. E esta representação condiz com o que pensa a população evangélica?

Segundo Simone Bohn (2004), a população evangélica no Brasil não é tão homogênea quanto aparenta, e sua associação com condições indicativas de pobreza e suas afinidades com a chamada “direita” no espectro ideológico não se comprovam. Ao mesmo tempo, a ideia de que a filiação religiosa tende a se traduzir em lealdades religiosas quase que automáticas aparece em sua pesquisa, mesmo que essa lealdade não tenha uma matriz ideológica clara:

[...] posições tradicionalistas dos evangélicos em relação ao aborto e ao homossexualismo não são necessariamente acompanhadas por posturas políticas conservadoras (o exemplo que utilizamos foi o posicionamento em relação à proibição de greves contra o governo). Essa dissociação torna difícil caracterizar os evangélicos como membros da base do que vem sendo identificado como “nova direita”. As opiniões desse grupo diante de temas econômicos, além disso, acentuaram ainda mais a dificuldade de posicioná-los claramente à direita ou à esquerda do espectro ideológico. (BOHN, 2004, p. 335)

Tal análise pode ser confirmada nas duas eleições onde Lula foi vitorioso, já que houve uma aproximação com setores evangélicos por conta do vice-presidente José Alencar, que tinha o apoio da Assembleia de Deus e da IURD. Já em 2010, houve uma divisão da chamada bancada evangélica entre a candidatura do PT e do PSDB. Entretanto, em 2014, a pauta conservadora e contra a chamada “ameaça

comunista” se acirrou e grande parte dos políticos evangélicos que já não tinha apoiado a campanha do PT, deixou de apoiar o governo eleito.

Como ocorreu essa mudança de posicionamento? Houve uma mudança no discurso da esquerda tão grande que justificasse a perda de apoio por parte dos evangélicos? A maior visibilidade das causas LGBTI+ e de direitos humanos em geral, defendidas pelo campo progressista, pode ter sido a principal causa de reação da bancada evangélica? Mas por que somente a partir de 2014 já que desde a década de 1990, a esquerda já lutava contra a homofobia e pelos direitos civis em geral?

Uma outra possibilidade é que tal mudança possa ser explicada pelas forças econômicas e pelas corporações pentecostais que apoiam e financiam boa parte da bancada evangélica. Pode-se aventar a hipótese de que a motivação de tal bancada seja apenas um simulacro para executar uma agenda determinada pelos grupos econômicos que conduzem as ações de tal frente parlamentar. Ou seja, por trás de todo o discurso baseado na moral cristã, permanece a mesma lógica que Engels escreveu no prefácio da obra “18 de Brumário de Luís Bonaparte” de Karl Marx:

[...] a grande lei do movimento da história, a lei segundo a qual todas as lutas históricas travadas no âmbito político, religioso, filosófico ou em qualquer outro campo ideológico são de fato apenas a expressão mais ou menos nítida de lutas entre classes sociais, a lei segundo a qual a existência e, portanto, também as colisões entre essas classes são condicionadas, por sua vez, pelo grau de desenvolvimento da sua condição econômica, pelo modo da sua produção e pelo modo do seu intercâmbio condicionado pelo modo de produção (ENGELS in MARX, 2011, p. 22).

Ou seja, o direcionamento ideológico de parte da bancada evangélica se adequaria aos interesses dos grupos econômicos que são efetivamente os comandantes da atuação de seus deputados. Conforme será visto no capítulo de Financiamento de Campanha, tal hipótese é de difícil comprovação empírica e um estudo mais aprofundado sobre os setores que financiam os deputados evangélicos estaria bem além da proposta da presente pesquisa. Desta forma, serão utilizados aqui os instrumentos que a literatura nos apresenta que é a tipologia dos partidos políticos em função de sua ideologia.

Resumidamente, a opção metodológica utilizada para classificar ideologicamente os deputados evangélicos se baseia em quais partidos eles escolheram para se candidatar, assumindo que estes seguem (ou deveriam seguir)

a mesma ideologia dos partidos escolhidos. Para tal intento, tomaremos como base a classificação indicada por KRAUSE et al (2017) já que é fruto de uma extensa compilação feita por especialistas na área em conjunto com o índice de Power e Rodrigues-Silveira (2018). O resultado de ambas classificações fica muito próximo da classificação realizada por Tarouco e Madeira (2015).

Segundo Machado (2017), não existe unanimidade sobre a classificação ideológica de partidos políticos e especificamente no Brasil, a tarefa de classificação é ainda mais árdua, não só em função da quantidade de partidos que são criados e desfeitos a cada legislatura, como também em função do posicionamento distinto dos partidos nas diversas regiões do país. Um partido que em determinado estado seria considerado de centro, pode ser de extrema direita em outro estado da federação em função das coligações e dos grupos econômicos que o apoiam.

Entretanto, essa complexidade não inviabiliza a categorização dos partidos políticos em alguma medida. Machado (2017) reforça a necessidade deste tipo de classificação:

(...) negar que existam distinções entre os partidos políticos seria reduzir as eleições a um teatro para recondução de uma elite política única e indivisível que dominaria a esfera pública independente dos interesses sociais predominantes na sociedade. Mesmo em sistemas políticos onde ideologias partidárias não representem projetos rígidos quanto ao direcionamento da ação, subsiste outra face da ideologia: a capacidade de permitir aos atores políticos – eleitores e membros de partidos – distinguir opções políticas através de referências quanto à posição de aproximação ou distanciamento das demais (p.46)

Bolognesi et al (2019) realizou recentemente extensa revisão de literatura sobre o tema e uma pesquisa do tipo *web survey* comparando os resultados obtidos especificamente para partidos pequenos com os resultados de outras tipologias utilizadas em diversas pesquisas científicas. Outras abordagens interessantes que valem a pena ser citadas são a de Codato et al (2018) e Power e Rodrigues-Silveira (2018). Codato et al. buscaram adicionar o viés religioso na classificação de determinados partidos políticos, distinguindo os partidos de direita como seculares ou cristãos.

Já Power e Silveira-Rodrigues (2012) trabalharam com um indicador numérico que visa não apenas classificar os partidos nos três blocos ideológicos, mas criar uma gradação que varia de -1 (mais à esquerda) a 1 (mais à direita). Este indicador

foi gerado a partir de questionários realizados com os deputados e senadores entre os anos de 1990 e 2017 e tal como em Krause et al (2017), tem a vantagem de categorizar os partidos ao longo do tempo, isto é, consideram que existe uma instabilidade ideológica temporal. Demonstram que um partido que era considerado de Centro em 2001 pode passar a se encaixar no espectro da direita em 2015 ou vice-versa.

Para que se possa entender melhor esta chamada instabilidade ideológica, é importante rever a genealogia dos partidos políticos no Brasil. No início do processo de redemocratização, antes da nova constituição, foi realizada uma reforma partidária e os primeiros partidos políticos que surgiram a partir da reestruturação do MDB e da ARENA foram PMDB, PDS, PP, PTB, PDT e PT. Os partidos de direita eram o PDS e o PTB, ambos formados principalmente por políticos da ARENA, partido este que apoiava o regime militar.

Já o PMDB era formado por políticos de diversos matizes, de antigos apoiadores do regime militar a representantes do extinto Partido Comunista. O PP (Partido Popular) teve vida breve e foi incorporado ao PMDB logo após a sua criação⁴¹. No campo da esquerda, surgiram o PDT sob a liderança de Leonel Brizola e o PT tendo Lula como figura central, sendo este eleito o primeiro presidente do partido e nas eleições de 1987 o deputado federal mais votado no Estado de São Paulo.

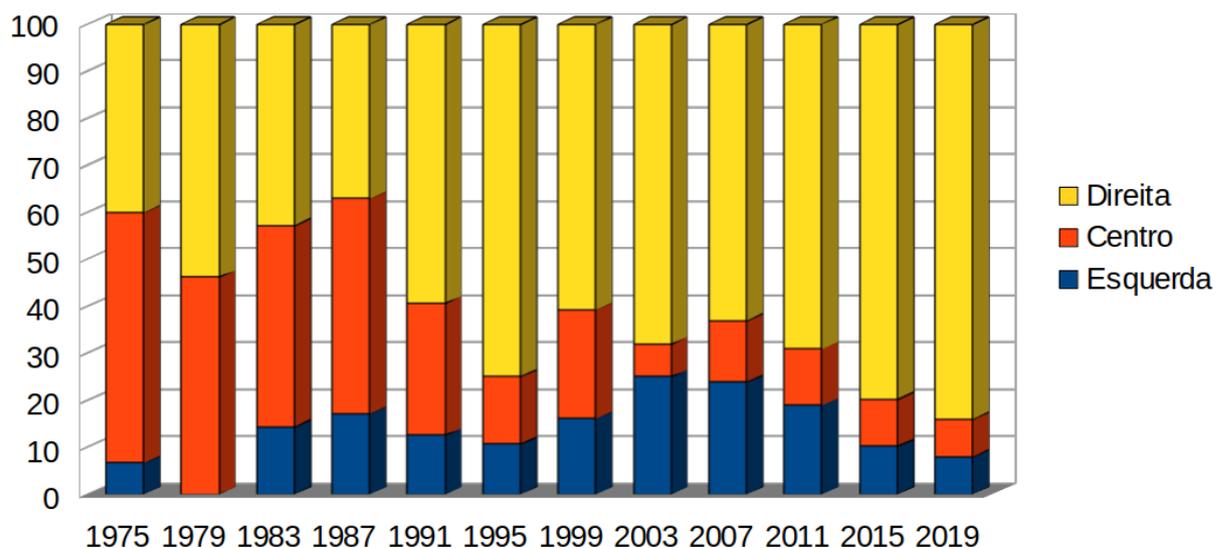
Após as eleições de 1982, houve uma ruptura no PMDB e as lideranças dissidentes criaram o PCdoB, PCB, PSB e o PSDB, sendo este último proveniente da união dos “originais do MDB”. Os três primeiros foram classificados como de esquerda enquanto que o PSDB foi classificado pelos especialistas como um partido de centro, tal como o PMDB. Ao longo da década de 80, o PDS sofreu um fatiamento ainda maior que o PMDB e gerou diversos partidos menores. Entretanto, as lideranças mais influentes no Congresso e nos Estados passaram a fazer parte do PFL que já no século XXI teria o seu nome alterado para Democratas (DEM).

Durante toda a ditadura militar os candidatos evangélicos estiveram divididos entre o MDB e a ARENA, ou seja, não havia uma tendência por um partido em detrimento ao outro na Câmara Federal. Somente a partir das eleições de 1991, com

⁴¹ A sigla viria a ser reutilizada 4 anos depois pelo Partido Progressista mas com outra configuração e organização partidária

o aumento de candidatos pentecostais eleitos, o equilíbrio se desfaz pois os candidatos evangélicos passam a optar mais por partidos de direita, tal como podemos ver no gráfico abaixo, que detalha o percentual de deputados federais no início de cada legislatura baseado na tipologia de Krause et al.⁴²

Gráfico 3 – Deputados Evangélicos por Bloco Ideológico (%)



Fonte: Elaborado pelo autor

O gráfico apresenta alguns detalhes interessantes como o aumento de deputados evangélicos no campo da esquerda durante o governo Lula, principalmente no PSB e PT que elegeram 13 deputados evangélicos em 2002, quase o mesmo contingente que partidos que abrigavam tradicionalmente os candidatos evangélicos como PFL e PL.

Ainda sobre a tipologia apresentada por Krause et al (2017), vale uma nota metodológica. Para a geração do gráfico 3 e também para as subseqüentes análises, foram realizadas três alterações na classificação dos partidos. Na tipologia original de Krause et al, os partidos PPS, PV e PMN eram considerados de esquerda, mas, para a presente análise, iremos considerá-los como partidos de Centro a partir de 2006. Optou-se por este ajuste pois notou-se que houve uma unanimidade tanto em Codato et al (2018) como em Bolognesi et al (2019) e Power e Rodrigues-Silveira (2018) que classificaram tais partidos como de Centro.

⁴² A lista completa de partidos nesta análise encontra-se no repositório do LARHUD no endereço: <https://doi.org/10.5281/zenodo.5515647>

É importante frisar que essa guinada à direita antecedeu e foi bem mais acentuada que a ocorrida para o total de congressistas, ou seja, se compararmos o crescimento da direita no Congresso utilizando o mesmo critério partidário, pode-se notar que o crescimento do número de deputados evangélicos foi mais acentuado e ocorreu 10 anos antes.

Foi possível notar também que a migração maior ocorreu dos partidos de centro (PMDB e PSDB) para os partidos de direita, sobretudo para os partidos que foram criados especificamente para abrigar candidaturas evangélicas, tal como PSC, PR e PRB. A partir da criação destes partidos, foi possível mapear de forma mais clara a influência das corporações pentecostais na política partidária.

O PRB foi criado em 2005 como o braço político da IURD. Apesar de ter eleito apenas um deputado nas eleições de 2006, reelegeram José de Alencar como vice de Lula e em 2010 elegeram nove deputados (sendo sete deles evangélicos) e o senador Marcelo Crivella. Já o PR, criado em 2006 a partir da fusão do Partido Liberal (PL) com o PRONA, também tinha inicialmente uma forte presença da IURD, mas em 2010 elegeram apenas deputados de outras corporações pentecostais, sendo 4 deputados da Assembleia de Deus, um da IIGD e um da IMPD. Além disso, o PR elegeram três senadores em 2006 sendo um deles evangélico.

O terceiro e o único partido partido que procurou um nome ligado ao cristianismo para identificá-lo foi o Partido Social Cristão (PSC) que, a partir das eleições de 2002, foi ocupado pela AD tal como relata Cunha (2019)

O objetivo da AD era competir como o poder que a IURD adquiria no apoio ao governo, o que gerou um novo fôlego ao PSC. A AD alocou [no partido] parte de seus parlamentares e aumentou a importância da sigla com vistas a composições futuras, que a colocassem como parceira no poder, em alianças regionais e nacionais (p.54)

Vários outros partidos também se mobilizaram para receber candidatos de outras corporações pentecostais e o “troca-troca” de deputados entre os diversos partidos durante seus mandatos também foi uma constante desde 2006. Só em 2016, trinta e dois dos deputados evangélicos mudaram de partido (CUNHA, 2019). Estas mudanças constantes foram fruto de criação de novos partidos, mas também de arranjos para o uso do fundo partidário para as eleições municipais e outros que talvez não tenham sido tão republicanos.

A presença de deputados que representam as corporações pentecostais é preponderante em relação às demais denominações. A maior “bancada” é de assembleianos seguida da IURD e de Batistas. Entretanto, como já foi visto na seção anterior, tanto os assembleianos quanto os batistas se organizam sob forma de ministérios ou convenções, ou seja, não existe uma figura central, uma organização que comanda de forma unificada seus representantes no parlamento.

Dessa forma, a corporação pentecostal com maior presença no Congresso é a IURD, seguida de forma mais modesta pela IEQ e pela IIGD. Ao longo da presente pesquisa, procurou-se diferenciar duas organizações assembleianas de maior poder no congresso e na mídia: a AD-VEC (Vitória em Cristo) controlada por Silas Malafaia e a AD-MM (Ministério Madureira) controlada pela família Ferreira e tendo sob sua égide personagens importantes como o ex-presidente da Câmara Eduardo Cunha e o pastor Everaldo. Entretanto, não foi possível atestar se o número de assembleianos ligados a estas duas corporações pentecostais é maior dentre o total de deputados que se autodenominam apenas como pertencentes ao nome genérico “Assembleia de Deus”. A representação das organizações religiosas no Congresso ao longo do tempo pode ser vista na tabela abaixo:

Tabela 11 - Deputados Federais evangélicos por Denominação Religiosa

Denominação	1983	1987	1991	1995	1999	2003	2007	2011	2015	2019
AD	1	13	12	10	11	18	9	21	22	34
IURD		1	3	6	12	18	6	7	13	21
Outras	1	5	2	2	2	1	2	6	14	18
Batista	6	10	8	5	14	10	9	13	14	15
Presbiteriana	2	4	1	5	3	2	3	8	7	7
IEQ	1	2	1	1	2	5	2	6	6	5
IIGD			1	1	1	1	2	3	3	3
Luterana	1		3	4	6	3	3	1	2	3
AD-MM					1	1			1	2
AD-VEC						1			3	2
Maranata						1	3	2	2	2
Metodista	2		1		1	1	1	2	3	2
Anglicana						1	1	1	1	1
Sara Nossa Terra						1	2	1	2	1
Total	14	35	32	34	53	64	43	71	93	116

Fonte: Dados da Pesquisa

Dentre as denominações tradicionais, ditas “de missão”, os Batistas são os que mais tem representantes no Congresso seguidos dos Presbiterianos, Luteranos, Metodistas e Anglicanos. Vale notar que para as três últimas denominações, existe uma constância de deputados e percebe-se que um mesmo deputado é reeleito várias vezes. Essa característica destoa bastante dos candidatos ligados às denominações pentecostais que, a cada legislatura, apresentam quase sempre candidatos diferentes. Além disso, não existe um apoio formal destas igrejas e em todos os casos analisados, os candidatos já eram evangélicos antes de entrar para a vida pública.

Um outro detalhe é a quantidade de deputados identificados dentro da categoria “Outras” ocupando a 3ª colocação a partir de 2015. Esta categoria foi criada para abarcar todas as denominações que tiveram menos de 5 deputados ao longo de todas as legislaturas ou quando não foi possível identificar a qual igreja um determinado deputado pertencia. Nesta categoria estão desde adventistas até igrejas que empregam o uso de ervas medicinais em seus cultos.

Ao longo deste capítulo, foi possível perceber quão complexa é a representação dos evangélicos no Congresso e também quão forte é a presença das corporações pentecostais na atuação dos parlamentares apoiados por elas. Não foi possível até esse ponto da pesquisa, identificar quais dos deputados foram apoiados formalmente por tais organizações, apesar da literatura indicar um ou outro candidato que tiveram o apoio formal de uma determinada denominação através de análise documental em matérias jornalísticas. Espera-se que, ao investigar os padrões de financiamento dos deputados evangélicos no próximo capítulo, seja possível encontrar mais subsídios para tal identificação.

5 FINANCIAMENTO DE CAMPANHA POLÍTICA

Este capítulo tem o objetivo de realizar um levantamento bibliográfico que procura resumir o tema de financiamento de campanha no Brasil com o foco no levantamento e tratamento dos dados disponíveis e dos principais eixos temáticos que são utilizados pelos pesquisadores da área.

5.1 HISTÓRICO

O Código Eleitoral de 1950⁴³ foi a primeira legislação que trouxe um ordenamento jurídico sobre a condução do sistema partidário e eleitoral no Brasil e tinha uma preocupação especial com as questões financeiras do processo, além de prever regras de funcionamento e de organização partidária (CAMPOS, 2009). Conforme relata Campos

também definia que os estatutos partidários deveriam conter regras específicas que obrigassem os partidos a “fixar e apurar as quantias máximas [de recursos] que os candidatos poderiam despende pessoalmente com a própria eleição”. Além da questão do limite de gastos (que prevaleceu nas legislações seguintes), o Código considerou que os partidos contemplassem em seus estatutos os limites para as contribuições e os auxílios dos seus filiados. (CAMPOS, 2009, p.26)

A partir de 1965, o financiamento de campanhas no Brasil passou a ser misto: parte do financiamento é feito pelo Estado e parte é arrecadado de pessoas (físicas ou jurídicas) dispostas a patrocinar um candidato ou partido. O financiamento feito pelo Estado se dá por dois canais: a partir do fundo partidário e pela concessão do horário gratuito eleitoral, além de spots no rádio e na televisão.

A Constituição de 1988 trouxe mais liberdade e autonomia organizativa para os partidos políticos além do acesso ao Fundo Partidário (FP). Entretanto trouxe novas regras como a obrigatoriedade na prestação de contas com o TSE e a necessidade de criação de uma personalidade jurídica para cada partido. E mesmo com estas exigências, a primeira eleição para presidente foi cercada de irregularidades e foi uma prévia de como seriam os anos futuros:

⁴³ Lei no 1.164, de 1950

A primeira eleição direta para presidência do país, em 1989, trouxe uma experiência amarga, levando inclusive ao impeachment em 1992 do presidente eleito. No centro do problema estava justamente a corrupção devido às contribuições ilegais e tráfico de influência entre doadores, partido e governo. A reação foi a criação de uma lei temporária (Lei no 8.713/1993) para regulamentar o pleito de 1994. Dessa forma, foi permitido às empresas financiar partidos, mas com limites estabelecidos. (KRAUSE et al, 2015, p.9)

Segundo Mancuso (2015), que realizou uma extensa revisão bibliográfica sobre o tema, as primeiras pesquisas com um corte empírico sobre o financiamento de campanha no Brasil datam de 1997, tendo o trabalho de Samuels na década de 2000 se convertido na base metodológica para as pesquisas subsequentes. Ainda segundo Mancuso, até 2014 pode-se categorizar as linhas de pesquisa em três grandes questões: “(i) a relação entre investimentos e resultados eleitorais; (ii) a relação entre investimentos eleitorais e benefícios para os financiadores e (iii) os determinantes do investimento eleitoral”.

Em toda a literatura, é ponto pacífico que o financiamento empresarial tem sido uma das variáveis de grande peso no resultado eleitoral no Brasil e, talvez por isso, o total de investimento veio crescendo acima da inflação desde 1998, eleição após eleição. Samuels (2001) já indicava que, desde 1994, a maior parte do financiamento vinha de fontes empresariais e que a participação da sociedade civil tenderia a diminuir. Só no ano de 2010, o investimento para todos os cargos em disputa chegou a quase 3 bilhões de reais e 74% deste valor veio de pessoas jurídicas (MANCUSO, 2015, p.156).

5.2 SISTEMA ATUAL

A partir de 2015, ocorreram mudanças significativas que afetaram a realidade do formato como o financiamento foi conduzido. Até as eleições de 2014, tanto cidadãos quanto empresas privadas estavam aptos a contribuir com as campanhas eleitorais. Mas a partir da minirreforma eleitoral (Lei No 13.165, de 29 de setembro de 2015), foi vedada a contribuição de pessoas jurídicas. Para compensar essa “queda da arrecadação”, os partidos políticos aumentaram consideravelmente o valor que passaram a receber do Fundo Especial de Financiamento de Campanha (FEFC) através da Lei 13.487/17, de 6 de outubro de 2017) que previa um aumento significativo do financiamento público para as campanhas políticas.

De acordo com a nova lei, 2% do total do fundo é dividido entre os partidos existentes (tendo ou não representação no Congresso Nacional) e 98% do total restante é dividido apenas entre os partidos que têm representação na Câmara dos Deputados e no Senado. Já nas eleições de 2018, foram destinados 1,7 bilhões de reais⁴⁴ e nas eleições de 2020, o valor passou para 2,07 bilhões⁴⁵.

Por conta desta minirreforma e do agravamento das denúncias e condenações pelo uso de recursos não contabilizados, o chamado “caixa dois” (dinheiro fornecido por empresas para políticos e partidos, não contabilizado no Tribunal Eleitoral), o volume de dinheiro envolvido nas eleições em 2016 e 2018 foi bem menor que o apurado nos anos anteriores, como podemos ver na tabela abaixo:

Tabela 12 - Financiamento Eleitoral Privado entre 2010 e 2018

Valores em R\$ corrigidos pelo IPCA de 01/01/2019

	Pessoas Físicas	Próprio Candidato	Empresarial	Total
1994	219.457.903	113.304.475	1.518.497.440	1.851.259.818
1998	160.497.053	176.052.172	869.503.506	1.206.052.731
2002	367.979.805	332.972.490	1.437.358.941	2.138.311.236
2006	565.459.879	377.712.570	1.947.756.075	2.890.928.524
2010	740.088.991	535.193.837	3.855.829.609	5.131.112.437
2014	735.084.304	491.267.922	3.976.788.149	5.203.140.375
2018	581.003.526	448.601.963	0	1.029.605.489

Fonte: Elaborado pelo autor baseado nos dados do TSE e Santos (2016)⁴⁶

Como podemos ver na tabela 12, até 2014 a doações de pessoas físicas e contribuições de parlamentares são muito menores quando comparadas com a doação de pessoas jurídicas. E em 2018, o total de doações caiu a 1/5 do total doado em 2014.

⁴⁴ Dados disponíveis em <http://www.tse.jus.br/eleicoes/eleicoes-2018/prestacao-de-contas-1/fundo-especial-de-financiamento-de-campanha-fefc>

⁴⁵ Dados disponíveis em <https://www.tse.jus.br/eleicoes/eleicoes-2020/prestacao-de-contas/fundo-especial-de-financiamento-de-campanha-fefc>

⁴⁶ Os dados de 1994 e 1998 foram obtidas a partir de planilha do pesquisador Bruno Carazza Santos (2016) que disponibilizou os dados compilados pelo professor David Samuels já que o TSE não disponibilizou estes dados no portal de Dados Abertos. Todos os valores foram atualizados pelo IPCA de 01/01/2019.

Vale notar, entretanto, que mesmo com a nova legislação, as empresas continuaram patrocinando legalmente os candidatos indiretamente, mediante doações feitas pelos sócios das mesmas empresas que antes financiavam as campanhas. Segundo matéria do Jornal O Globo, somente para o candidato Pedro Paulo (candidato a prefeito pela cidade do Rio de Janeiro), 256 sócios de empresas de Construção, Engenharia e Infra-estrutura doaram acima de 30 mil reais durante a campanha de 2016⁴⁷.

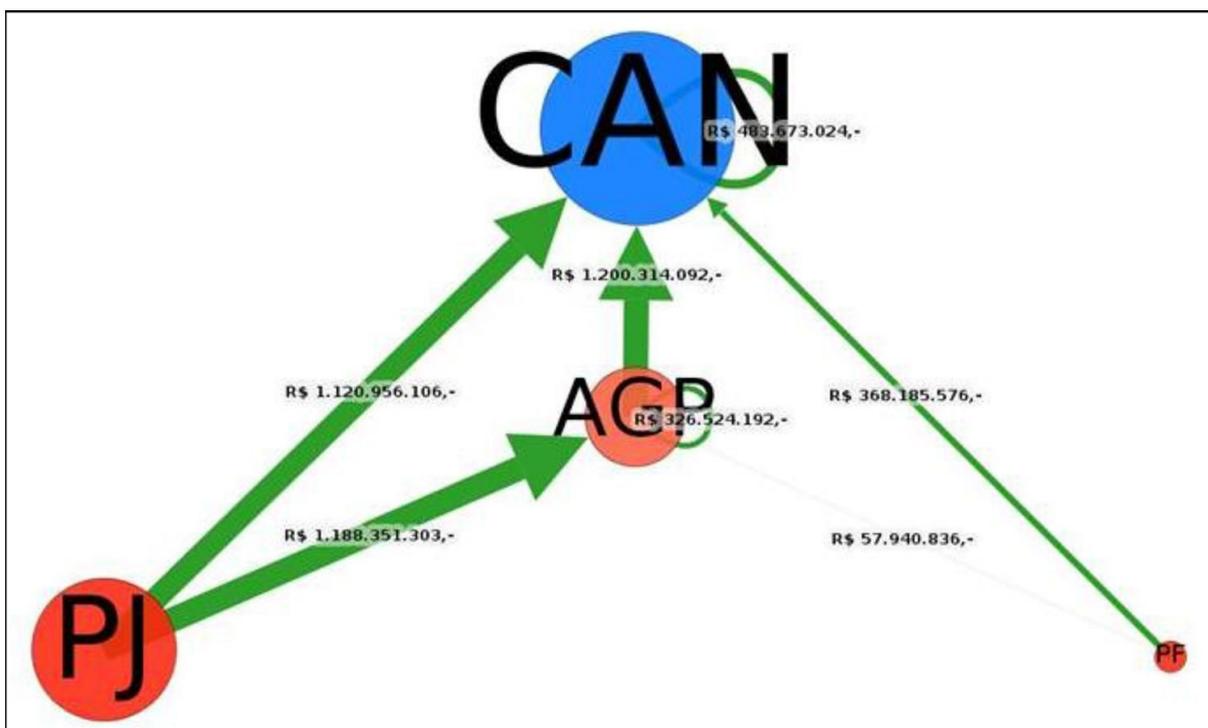
Além disso, novos casos de corrupção e crimes eleitorais continuaram acontecendo. A mais grave é a denúncia sobre vultuosos investimentos patrocinados por empresários nas redes sociais, em favor da candidatura do presidente da república eleito em 2018 e que ainda não foram apuradas. Também houve operações bem sucedidas da Polícia Federal, que ocasionaram na perda de mandato de uma deputada estadual e um deputado federal no Acre em função do uso de laranjas (Operação Santinhos), e na queda do ministro Gustavo Bebião já no governo Bolsonaro, em função da mesma prática no Estado de Pernambuco (BELO, 2021).

Até 2014, havia uma preferência dos financiadores de campanha, especialmente empresas, por candidatos mais competitivos e que apresentassem maiores chances de se eleger (HOROCHOVSKI et al. 2016; MANCUSO HOROCHOVSKI e CAMARGO 2016). As empresas financiavam individualmente determinados candidatos e também faziam doações diretamente às agremiações partidárias e estas decidiam a melhor estratégia de dividir os valores entre os candidatos. A divisão desses recursos pode ser visualizada na figura abaixo onde os nós são os tipos de atores que compõe o sistema: candidatos (CAN), Pessoas Jurídicas (PJ), Pessoas Físicas (PF) e Agentes Partidários (AGP)⁴⁸:

⁴⁷ <https://oglobo.globo.com/brasil/empresas-driblam-lei-para-doar-campanhas-eleitorais-20132632>

⁴⁸ Os autores denominaram de Agentes Partidários os Diretórios Nacionais, Estaduais, Municipais, Zonais e Comitês de Financiamento, estes últimos criados durante as eleições para movimentar as doações e despesas de uma região ou grupo de candidatos. O presente trabalho segue com a mesma terminologia.

Figura 4 – Fluxo de Doações nas Eleições Presidenciais em 2010



Fonte: HOROCHOVSKI et al. 2016

Ressalta-se que a espessura das arestas procura ser proporcional ao montante de recursos nos fluxos e que o diâmetro do nó procura representar o montante de recursos recebidos ou doados por cada ator. É importante notar que o peso das arestas que ligam os agentes partidários (AGP) a candidatos (CAN) é próximo ao peso das arestas que ligam pessoas jurídicas (PJ) aos partidos. Além disso, existe uma grande movimentação financeira intra-partidária e entre candidatos, ou seja, uma instância partidária transferindo recursos para outra instância, estando elas no mesmo partido ou não. Tal como atestam HOROCHOVSKI et al (2016):

as transferências entre agremiações partidárias entre si (loops) e de candidatos entre si (loops) revelam a dimensão estratégica de busca de aliados no pleito eleitoral, o que acontece também nas transferências das agremiações partidárias diretamente para os candidatos (p. 6)

Possivelmente, as doações feitas diretamente às AGPs são uma estratégia para que a empresa não “marque” sua doação a um determinado candidato ou então para que o candidato não fique “marcado” como sendo financiado por uma determinada organização que não é de seu interesse. Um caso notório que exemplifica essa movimentação foi a doação da JBS ao PP que prontamente

transferiu o mesmo valor ao então candidato a deputado federal Jair Bolsonaro. Ao ser questionado de sua ligação com a empresa que estava sendo investigada pela Lava-Jato, negou que o dinheiro teria sido originário da JBS⁴⁹ já que o dinheiro não tinha vindo diretamente para a sua campanha.

Com o fim do financiamento empresarial, já nas eleições de 2016, essa configuração foi alterada sensivelmente. A nova legislação proporcionou uma maior influência do partido político sobre o financiamento dos candidatos pois, como os repasses dos partidos às campanhas dos candidatos é controlado pelos diretórios nacionais e estaduais e o aporte de recursos privados diminuiu drasticamente, as candidaturas passaram a depender ainda mais da decisão dos dirigentes partidários que detém o poder de definir quem receberá as verbas do FEFC e do FP.

Inclusive, tal poder e conseqüente concentração de recursos se dá independente da estrutura partidária do partido, tal como atesta Bolognesi et al (2020) que realizaram extenso estudo empírico sobre o tema:

Os partidos apresentam concentração de recursos de campanha em poucos candidatos, invariavelmente nos mais competitivos e que compuseram a lista dos eleitos nas eleições de 2014. Frisamos que tal dinâmica já fora diagnosticada para o financiamento empresarial e, como vimos, vale para o financiamento partidário, independentemente do grau de estruturação dos partidos. À exceção do PCO, todas as agremiações agem da mesma forma, o que sinaliza para a pouca utilidade prática de se reduzir o financiamento privado como medida de redução das assimetrias, especialmente as internas, sem que isso venha acompanhado de dispositivos institucionais que obriguem os partidos a distribuir seus recursos de maneira mais equitativa (p. 25)

Para este estudo, Bolognesi et al (2020) construíram uma série de indicadores quantitativos que procuram endereçar a questão da concentração de recursos para poucos candidatos e buscam avaliar também a capacidade organizacional de cada um dos partidos. Como indicador de concentração de recursos, os autores utilizaram o coeficiente GINI⁵⁰ mensurando assim a distribuição interna das doações feitas por cada partido a seus candidatos.

Já sobre a capacidade organizacional de um partido, os autores construíram um índice de estrutura organizacional agregando diversos indicadores sobre a

⁴⁹ <https://www1.folha.uol.com.br/poder/2017/05/1886798-qual-partido-nao-recebe-diz-bolsonaro-sobre-propina-a-radio.shtml>

⁵⁰ O coeficiente GINI é comumente utilizado para calcular a distribuição de recursos entre uma determinada população e varia de 0 (nenhuma concentração) e 1 (total concentração).

burocracia e a capilaridade territorial de cada um dos partidos mas como os dados brutos não foram disponibilizados, não foi possível reproduzir o resultado da pesquisa, o que inviabilizou o uso de tal índice neste trabalho⁵¹.

Uma outra forma de analisar os dados de financiamento é através das despesas realizadas, pois na prática, nem toda doação é utilizada pela campanha e, além disso, uma campanha pode gastar mais do que arrecada, fazendo com que o partido ou o candidato termine as eleições com uma dívida a pagar. Analisar o tipo de despesas realizada por cada candidato pode trazer algum conhecimento sobre as estratégias utilizadas por cada campanha. Entretanto, ao longo do levantamento bibliográfico não foram encontrados estudos que fizessem uso específico de tais análises.

5.3 BASES DE DADOS ABERTOS DO TSE

Conforme já levantado no capítulo 2, os entes governamentais têm a possibilidade de disponibilizar dados públicos de diversas formas. A primeira é através de aplicativos na Internet onde o cidadão pode consultar dados individuais sobre o objeto a ser disponibilizado. No caso dos financiamentos de campanha, o TSE disponibiliza diversos serviços de “Consulta aos Doadores e Fornecedores de Campanha de Candidatos”⁵², onde é possível encontrar as doações individuais de uma pessoa física ou jurídica ou de um determinado candidato ou comitê de campanha.

Este tipo de serviço é útil para encontrar um dado específico para uma matéria jornalística ou até mesmo para os candidatos conferirem suas prestações de contas. Mas para uma análise agregada ou estatística, onde é necessário analisar o total de doações de um partido ou de uma cidade, este tipo de serviço não atende às necessidades de consulta. Dessa forma, o TSE disponibiliza desde 2002 as bases de dados de financiamento que podem ser utilizadas em softwares de análise como planilhas eletrônicas ou programas estatísticos.

⁵¹ Foi enviado e-mail para o primeiro autor do artigo solicitando os dados brutos mas este negou o envio alegando que os dados “foram base de publicação conjunta e ainda serão usados pelos pesquisadores”.

⁵² <https://divulgaspca.tse.jus.br/> e <https://inter01.tse.jus.br/spceweb.consulta.receitasdespesas2014/abrirTelaReceitasCandidato.action>

É interessante notar que as bases de dados de financiamento de campanha passaram a ser disponibilizadas antes mesmo de resultado final das eleições, assim, a medida que as prestações de contas dos partidos são registradas no sistema do TSE, o tribunal já disponibiliza os resultados no site da instituição através de arquivos no formato CSV agrupados por estado⁵³.

Até 2010, o TSE disponibiliza o conjunto de dados a partir de quatro tipos de arquivos para cada unidade da federação: doações feitas diretamente aos candidatos, doações recebidas pelos comitês partidários, doações recebidas diretamente pelo partido e um quarto arquivo, com as despesas realizadas. O tratamento destes arquivos não é trivial, pois a cada ano, a disposição dos atributos é modificada o que dificulta a importação dos dados de forma automatizada. Além disso, a discriminação dos metadados não era bem documentada e só em 2016, o TSE forneceu uma documentação apropriada no mesmo repositório que continha os dados.

Apesar da abertura dos dados de financiamento para a sociedade ter sido um grande avanço em relação ao aspecto financeiro das eleições, a base de dados apresenta ainda hoje diversas inconsistências e problemas de interpretação nos atributos das tabelas disponibilizadas. Além da pouca documentação disponível, alguns atributos não são claros e consistentes. Por exemplo, algumas doações são marcadas como recurso do “Fundo Partidário” e ao mesmo tempo como sendo feita por uma empresa privada (Recibo Eleitoral 401000700000AC000004 de 2014). Para exemplificar alguns dos problemas, foram construídos scripts em Python com alguns dos casos de erro para que outros pesquisadores possam reproduzir o que está aqui relatado e todos os códigos fontes foram divulgados no site do projeto FARMÍ⁵⁴.

Efetivamente, a análise dos dados requer atenção redobrada em função do tipo de doação, se foi destinada a um dos agentes partidários (AGP) ou se diretamente ao candidato. Por exemplo, uma doação feita ao partido por uma pessoa física ou jurídica pode não ser encaminhada a um determinado candidato, ou seja, uma vez que a doação entra no caixa de um AGP, este não necessariamente será utilizado para um determinado candidato. Esta realidade se encaixa no ditado popular que diz “dinheiro não tem rastro”, ou seja, uma vez que o dinheiro entra em

⁵³ <https://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais-1>

⁵⁴ As análises podem ser reproduzidas no link: <http://www.farmi.pro.br/explicando-base-de-financiamento-do-tse/>

uma conta-corrente única (no caso a conta corrente de um agente partidário), perde-se o rastro de como ele será gasto ou transferido no futuro.

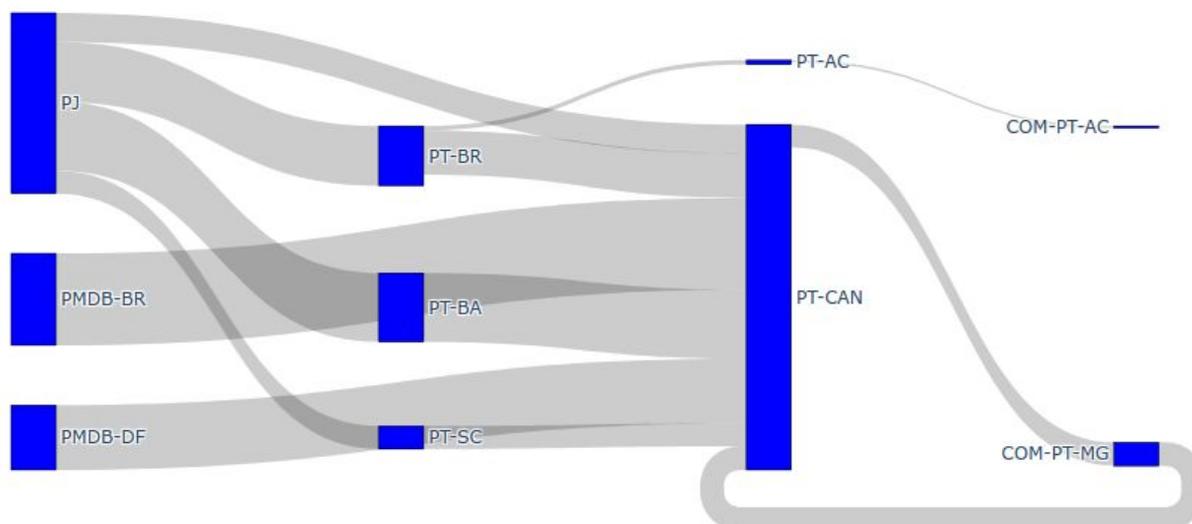
Para tentar rastrear o caminho do dinheiro, o TSE procurou incluir o registro das doações feitas indiretamente ao candidato a partir das doações feitas pelo AGP a um candidato, criando um atributo com a identificação do “Doador Original”, ou seja, a doação veio do partido, mas o doador original foi uma pessoa física ou jurídica. Entretanto, fica a critério do AGP definir como será feito o rateio entre os candidatos de uma doação específica e parece que não existe nenhum tipo de validação por parte do TSE destas movimentações, pois existem diversas inconsistências nas contas correntes dos AGPs. Peixoto (2016) confirma essa argumentação:

a regulamentação do sistema de financiamento brasileiro possui uma série de “pontos cegos” que dificultam tanto o controle e a fiscalização das contribuições financeiras a candidatos e partidos quanto se tornam inócuas no que se refere aos limites aos gastos. No primeiro quesito, as transações entre comitês financeiros e candidatos impossibilitam qualquer rastreamento das reais fontes de doação empregadas nas campanhas (p.171)

Para a comprovação deste argumento, foram construídos alguns estudos de casos comprovando tais inconsistências. O primeiro estudo selecionou aleatoriamente um CNPJ de uma empresa que tivesse doado diretamente para o partido, para algum AGP e também diretamente para os candidatos. O objetivo da análise era verificar como se dava a alocação das doações ao partido e como as doações eram distribuídas entre os candidatos. O resultado pode ser visto no gráfico abaixo⁵⁵:

⁵⁵ O gráfico Sankey permite a visualização de movimentação de valores entre contas. Procura mostrar as movimentações a partir da largura das setas que trafegam entre uma conta e outra, permitindo assim que o leitor tenha uma melhor noção de proporcionalidade entre os valores envolvidos. https://en.wikipedia.org/wiki/Sankey_diagram

Gráfico 4 – Exemplo de Fluxo Financeiro das doações de uma empresa



Fonte: Elaborado pelo autor

Neste gráfico, a sigla PJ (Pessoa Jurídica) representa a empresa que foi tomada como exemplo; PT-BR, o Diretório Nacional do partido selecionado, no caso o PT; SC, BA, AC, os diretórios estaduais com a sigla correspondente das unidades da federação; e os de prefixo PT-COM são os comitês eleitorais de cada estado. A partir do gráfico é possível visualizar várias movimentações interessantes. A que chama mais a atenção é a origem das doações de outro partido, ou seja, a empresa doa para o PMDB, que em seguida doa para um candidato do PT.

Um ponto peculiar é a movimentação entre candidatos que “devolvem” a doação recebida. Essa movimentação pode ser vista no comitê indicado como COM-PT-MG, onde o comitê transfere 500 mil reais para o candidato, que, em seguida, transfere o mesmo valor para o comitê. Neste caso específico, o valor foi exatamente o mesmo, mas nada impede que o candidato devolva apenas parte da doação ou devolva mais do que recebeu.

Um outro recurso interessante desse tipo de gráfico é a possibilidade de identificar rapidamente os “saldos” existentes nos fluxos financeiros. Se uma barra tem uma saída menor que a entrada, o gráfico indica que parte das doações não foi disponibilizada para os candidatos. Esse saldo está presente, por exemplo, na barra do Diretório Nacional (indicado pela sigla PT-BR), ou seja, o partido destinou 1,3 milhões ao Diretório Nacional, mas este só repassou 968 mil reais aos candidatos e

mais uma pequena parcela ao Diretório do Acre que, por sua vez, não doou o valor para nenhum candidato.

O detalhamento dos valores não fica indicado no gráfico “impresso” para não poluir a imagem com muitos números, mas é possível visualizá-los na sua versão dinâmica que fica disponível na web. Ou seja, em sua versão online, o leitor pode “navegar” e posicionar o mouse sobre cada um dos fluxos para visualizar os valores envolvidos.

Ao fim das análises realizadas, foi possível obter algumas regras para o estudo da base de financiamento a citar: para calcular o total doado para um **candidato específico**, basta somar as doações na base de candidatos, sem considerar as bases de comitês e de partidos. Os subtotais por tipo de doação também são elementos importantes de análise, pois é possível identificar os candidatos que foram mais beneficiados pelo partido, os que receberam mais doações de pessoas físicas e os que receberam um aporte maior de empresas.

Já para calcular o total de doações de pessoas jurídicas ou físicas a um determinado **partido**, não se pode levar em conta apenas a soma das doações aos candidatos, mas deve-se considerar também a soma das doações feitas diretamente ao partido e aos comitês. Em resumo, o total doado ao partido deve ser calculado pela fórmula: $PART + COM + CAND - DDO - DC$ onde:

- PART: o valor total recebido pelo partido;
- COM: valor total recebido pelos comitês;
- CAND: valor total recebido pelo candidato na planilha de doação aos candidatos;
- DDO: valor das doações marcadas por doadores originais na base de candidatos pois estas doações já foram contabilizadas em PART e COM.
- DC: devoluções de valores feitas por candidatos aos comitês ou doações feitas entre candidatos.

Inclusive, se esta subtração não é realizada, as totalizações das receitas são infladas, pois os valores de transferência entram em dobro na totalização. O próprio painel de estatísticas das eleições do TSE em 2018 mostra um valor errado em sua totalização!⁵⁶

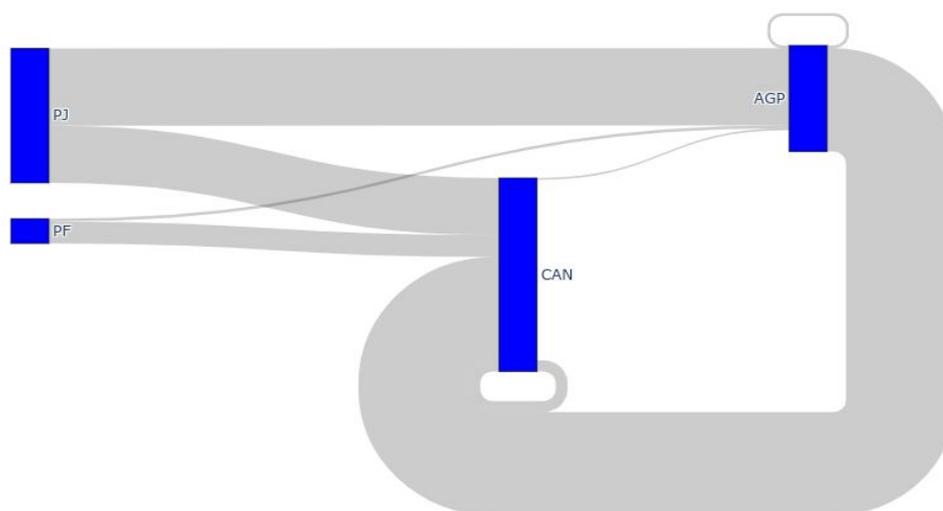
⁵⁶ <https://www.tse.jus.br/eleicoes/estatisticas/estatisticas-eleitorais>

Além disso, o campo de “Doador Original” não deve ser levado em conta no cálculo das doações feitas aos candidatos pois esta alocação tem apenas efeito contábil para o partido, que tem que prestar contas ao TSE como o dinheiro doado por um ente privado foi alocado ou gasto. Ou seja, não se pode garantir que o dinheiro recebido pelo candidato veio necessariamente do doador original indicado no registro da doação.

Por fim, se agruparmos as doações recebidas por cada partido por tipo de doação, ou seja, agrupando nas quatro variáveis definidas (PF, PJ, RP, AGP e CAND) é possível identificar durante uma eleição o peso de cada ator no processo eleitoral tal como foi construído por HOROCHOVSKI et al (2016), só que utilizando um gráfico sankey, pois este permite que se visualize o volume de transações entre os atores do processo de forma mais clara.

Gráfico 5 – Fluxo Financeiro das doações em 2014

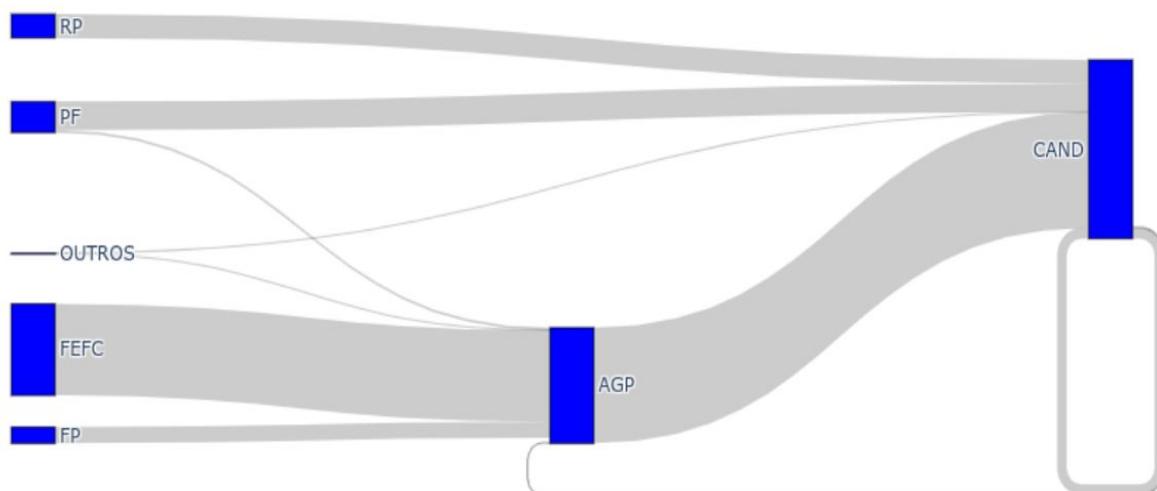
Doações em 2014



Fonte: Elaborado pelo autor

Já no ano de 2018, com o fim do financiamento empresarial, a maior parte das doações se originam do FEFC (R\$ 1,717 bilhões) e a participação de entes privados (R\$ 1,03 bilhões) não chegou sequer a um terço do montante total de 3,13 bilhões de reais de receita de toda a campanha. O fluxo financeiro de 2018 pode ser visualizado no gráfico 6 abaixo.

Gráfico 6 – Fluxo Financeiro de doações em 2018



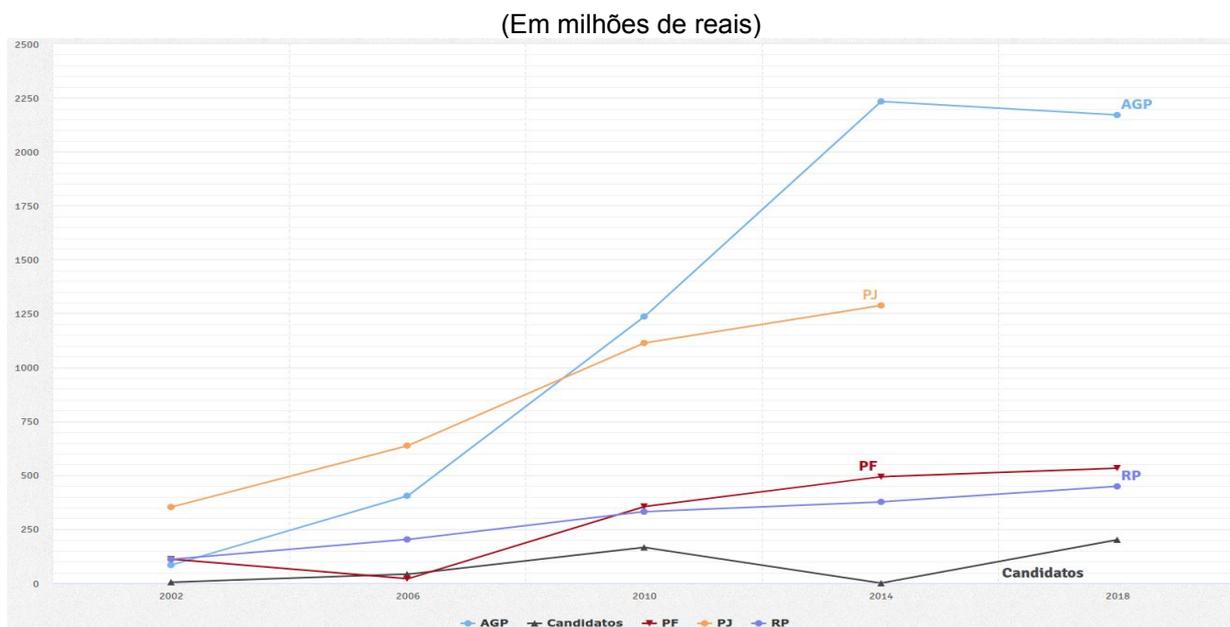
Fonte: Elaborado pelo autor

Ao final da análise dos tipos de doações que ocorreram de 2002 até 2018 nas eleições federais, também foi possível notar como as doações deixaram de ser feitas diretamente aos candidatos e passaram a ser feitas diretamente aos partidos, ou seja, os agentes partidários passaram a intermediar as doações. Tal análise sugere que as direções nacionais dos partidos passaram a ter um maior poder sobre a distribuição das doações para os candidatos de suas preferências mesmo antes da reforma eleitoral de 2015. O histórico das doações feitas aos candidatos pode ser visto no gráfico 7.

Tal fato corrobora a importância do partido político no processo eleitoral e também no financiamento de campanha. Peixoto (2016) resume bem a questão:

A filiação partidária diferencia os desempenhos dos candidatos. A explicação teórica possível é que o sistema eleitoral pode não dar incentivos para os eleitores escolherem e controlarem seus representantes, mas os partidos afetam indiretamente as chances dos seus candidatos. As formas mais prováveis são os recursos organizacionais, o tempo de propaganda eleitoral e as estratégias de coligação (...) A conclusão mais adequada é que pertencer a determinados partidos aumenta o desempenho eleitoral dos candidatos. Se isso afeta ou não o comportamento legislativo é outra história (p.116)

Gráfico 7 – Doações aos candidatos por Tipo de doador



5.4 FINANCIAMENTO POR SETORES ECONÔMICOS

Uma das diversas formas utilizadas para analisar o financiamento de campanha é verificar quais os setores econômicos financiam quais tipos de partidos e candidatos. Como já foi visto acima, desde o início das eleições em 1989 até o ano de 2014, o financiamento empresarial foi a maior rubrica dentre todos os tipos de financiamento existentes e entender quais os setores econômicos que mais investiram nas eleições foi matéria de análise na Ciência Política.

Samuels (2001) foi um dos primeiros pesquisadores que buscou construir uma divisão por setores econômicos e ainda dividiu as doações pelos diferentes cargos que estavam sendo disputados nas eleições de 1994 e 1998. Para tal, utilizou uma técnica para definir os setores de cada doador a partir do nome da empresa. Entretanto, o pesquisador não explicita como definiu os setores econômicos e tampouco utilizou os setores utilizados pelo IBGE⁵⁷.

Kasahara (2009) segue na mesma linha que Samuels, indicando que existe uma preponderância de doações por parte do setor financeiro em função da

⁵⁷ Desde 1995, o IBGE utiliza o Cadastro Nacional de Atividades Econômicas (CNAE) como forma de categorizar e sistematizar todas as atividades econômicas existentes e tal codificação também passou a ser utilizada pela Receita Federal obrigando que toda empresa informe em quais atividades econômicas irá operar.

influência deste na definição das políticas econômicas que afetam diretamente o sistema financeiro nacional. Boito Jr (2007) vai além e trabalha com o conceito de frações de classes capitalistas que se beneficiaram no modelo neoliberal, tal como os grandes bancos, mas também o agronegócio exportador e as empresas que se beneficiaram com as privatizações e com vultosos financiamentos disponibilizados pelo BNDES.

Santos (2020) também trabalha na hipótese de que “os fatores estruturais do capitalismo brasileiro, expressos pelo próprio setor de atividade, e o porte econômico das companhias determinariam a probabilidade de engajamento eleitoral”. Segundo o autor:

ao menos no caso de uma elite econômica, as vantagens obtidas pelo ato de financiar campanhas tiveram menor peso político do que normalmente se supõe. Valeria aqui, portanto, atentar previamente à posição dos agentes no sistema econômico e político, e não necessariamente à “utilidade” ou à “satisfação” de financiadores e financiados. Até mesmo porque a interação entre esses atores não se limitaria às campanhas eleitorais, pois está consolidada inclusive no arcabouço institucional que estrutura o processo decisório nacional — comitês setoriais constitutivos de política industrial (por parte do Executivo) e participação em audiências e fóruns estabelecidos pelas comissões e frentes parlamentares (no âmbito legislativo). (p.122).

Já Bandeira de Mello et al (2008) fez um contraponto indicando que as características organizacionais das empresas (como lucratividade, tamanho e estrutura societária) são mais importantes do que os setores a que pertencem para explicar como são decididas as doações das empresas (não só as eleitorais) e sugere que os interesses dos grandes investidores não se resumem à vantagem econômica obtida (contratos públicos e empréstimos do BNDES) mas à orientação direta das partes interessadas da empresa (os *stakeholders*).

Sob o ponto de vista empírico e pragmático, a identificação do setor econômico ao qual uma determinada empresa pertence pode ser obtida através do CNPJ da mesma. O cadastro de pessoa jurídica tem esse dado como obrigatório, seja no ato da criação da empresa ou sempre que o contrato social da mesma é alterado. Entretanto, como já foi visto no capítulo 2, a base de CNPJs não era disponibilizada em formato aberto, o que impossibilitava que pesquisadores fizessem uso desta informação de forma automatizada. Com a abertura desta fonte

de dados em 2017, foi possível baixar essa base e agregar os setores econômicos das empresas que financiavam as campanhas eleitorais.

Um dos problemas centrais para que fosse possível categorizar as doações seria definir quais setores econômicos utilizar. Para tal buscou-se a categorização utilizada pelo IBGE, o CNAE - Cadastro Nacional de Atividades Econômicas⁵⁸. O CNAE tem uma categorização complexa, dividida em seções, divisões, grupos, classes, sub-classes e atividade econômica e desta forma, muito fragmentada. Mesmo que se agrupasse as doações pelas seções (que é o grupamento mais geral), a totalização ficaria fragmentada. Ou seja, se houvessem muitas categorias, perderíamos a visualização dos grandes setores.

Dessa forma, foi criada uma classificação mais abrangente com apenas 6 setores: Indústria, Infraestrutura, Agronegócio, Comércio, Serviços e Instituições Financeiras, e a correlação de cada um deles com as divisões do IBGE pode ser visualizadas na tabela abaixo:

Tabela 13 - Setores Econômicos

Setor	Seções CNAE
Indústria	Indústrias Extrativas (B) e de Transformação (C)
Infraestrutura	Eletricidade e Gás (D); Água e Esgoto (E); Construção (F)
Agricultura	Agricultura, Pecuária, Produção Florestal, Pesca e Aquicultura (A)
Financeiro	Atividades Financeiras, De Seguros e Serviços Relacionados (K), Atividades Imobiliárias (L)
Comércio	Comércio (G)
Serviços	Os demais com exceção de Governo e Terceiro Setor

Fonte: O autor. Baseado na tabela de atividades econômicas do IBGE

Um outro ponto que complexifica ainda mais a divisão por setores econômicos é a existência dos chamados “grupos econômicos”, um conjunto de empresas interligadas, porém juridicamente interdependentes, cujo controle é centralizado pelos proprietários e que atuam em diversas atividades (GONÇALVES, 1991). Não se pode garantir que doações feitas por um dos CNPJs do

⁵⁸ https://cnae.ibge.gov.br/images/concla/documentacao/CNAE20_Introducao.pdf

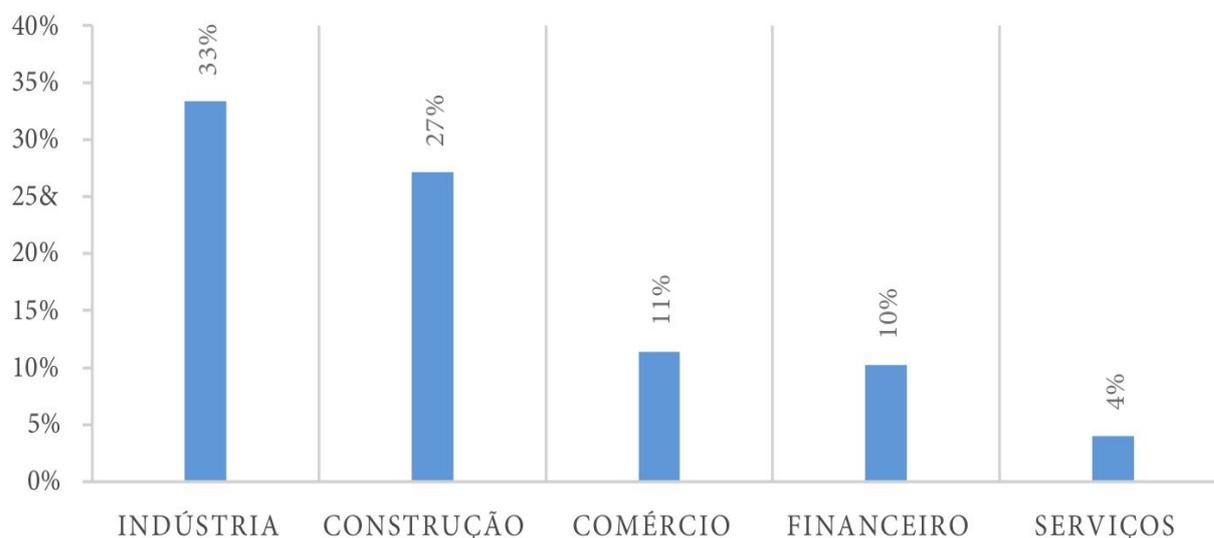
conglomerado terão o objetivo de privilegiar o setor a que ele pertence. Ou seja, a doação pode ter sido realizada por uma empresa do setor de comércio, mas com o objetivo de financiar um candidato que irá beneficiar o setor financeiro.

Santos (2020) realizou estudo semelhante entre setores econômicos⁵⁹ e o financiamento de campanha. Em sua análise, o autor sugere que “a decisão empresarial de participar das campanhas se explica para além da observação de ganhos e vantagens diretas obtidas junto à esfera pública” (p.31). O autor ainda reforça a utilidade de tal análise:

Independentemente do segmento, a elite empresarial de todos os setores, considerando os seus maiores doadores, financiaram de maneira semelhante, privilegiando campanhas nacionais e estruturas partidárias; enquanto os pequenos doadores, também de todos os segmentos, voltaram-se às candidaturas a deputado estadual e federal (SANTOS, 2020, p.33).

Nesta pesquisa, Santos demonstra que para toda a campanha, o setor que mais doou foi o da Indústria, seguida do ramo de Construção. Interessante notar que o setor agropecuário não foi computado. O resultado pode ser visto no gráfico 8.

Gráfico 8 – Doações em 2014 por Setor Econômico



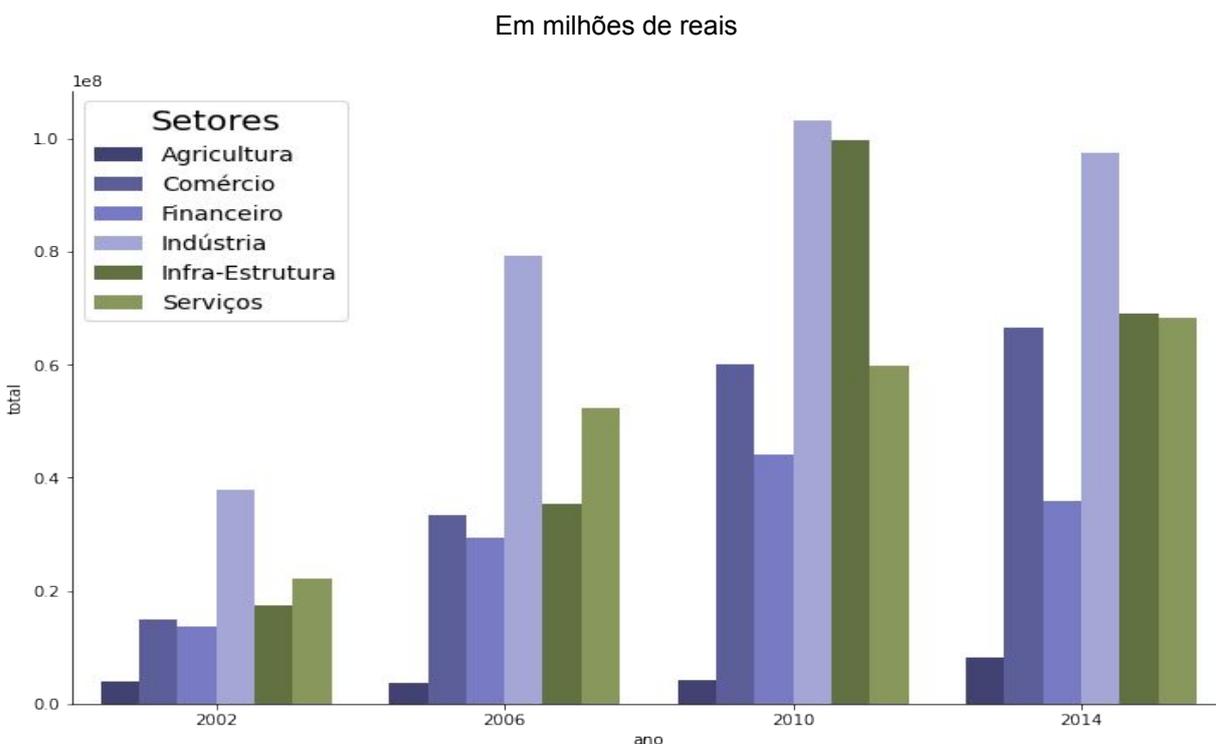
Fonte: SANTOS (2020)

Como o foco da presente pesquisa são os deputados federais, buscou-se realizar o mesmo processamento e verificar também se havia variação ao longo dos

⁵⁹ A categorização de Santos difere da apresentada neste trabalho apenas na seção de Construção onde achamos por bem incluir englobar os setores de Água, Esgoto e Eletricidade em um setor denominado Infraestrutura. Essa decisão foi tomada pois tais áreas estão intimamente ligada a obras públicas, tanto quanto o ramo de Construção.

anos. O gráfico abaixo mostra as doações por setor especificamente para os deputados federais nos anos de 2002 até 2014 (último ano em que foi permitido o financiamento empresarial).

Gráfico 9 – Doações por Setor Econômico – Somente para Deputados Federais



A partir do gráfico acima, é possível notar que a distribuição dos setores é diferente quando se trata de doações para os deputados federais: a presença de empresas do setor agropecuário é tímida dentre as demais e o setor de indústria se mantém em primeiro lugar durante todos os anos.

Seguindo ainda neste mesmo tipo de análise, buscou-se uma forma de comparar os anos anteriores com 2018. Apesar do fim do financiamento empresarial, algumas matérias jornalísticas apontavam que determinados setores econômicos continuaram a financiar as eleições através de doações realizadas pelos sócios de grandes conglomerados. Ou seja, como a legislação vigente impôs um teto muito elevado em função dos rendimentos individuais das pessoas físicas, muitos sócios de empresas, sendo eles candidatos ou não, fizeram vultuosas doações a partir das eleições de 2016. Apesar do volume ter sido bem menor que os dos anos anteriores,

as doações advindas de pessoas físicas ainda representavam um valor considerável.

Nesse ponto, surge um entrave metodológico de difícil solução: como identificar a que setor econômico uma pessoa pertence? Doadores que são pessoas famosas, como o ex-candidato a presidente e ex-ministro da Fazenda Henrique Meirelles (o maior doador individual na campanha de 2018), são casos excepcionais. E obviamente, não é possível buscar manualmente cada um dos doadores em um volume de 560 mil doadores únicos desde 2002.

Assim, buscou-se adotar uma abordagem que verificasse, a partir da base de CNPJs, quais os doadores são sócios de empresas e, a partir daí, assumir que o sócio de tal empresa é um representante daquele setor econômico. Ou seja, assume-se aqui que as mesmas motivações que faziam com que uma empresa fizesse doações a uma campanha política também se aplicam a seus sócios.

Mesmo com essa abordagem, algumas questões ainda podem ser levantadas: como classificar uma pessoa que seja sócia de mais de uma empresa, que atuem em setores diferentes? Ou então: é válido classificar uma pessoa que seja sócia de uma microempresa e faça uma pequena doação em função de suas convicções políticas no mesmo balaio que um grande investidor?

Para a primeira pergunta, optou-se por classificar o sócio pelo setor onde ele tivesse o maior número de participações e, em caso de desempate, priorizar os setores de agricultura, indústria e infraestrutura, ou seja, os setores produtivos ditos primário e secundário. Partiu-se da premissa que estes setores seriam a fonte primária da receita dos sócios e, se houvesse participação deste em outros setores, estas seriam de menor monta.

Um outro ajuste também foi realizado em função da forma como grandes empresários movimentam seus bens. Geralmente, boa parte dos capitalistas criam empresas específicas para administrar seus bens, ou seja, além de serem sócios da indústria, comércio etc. que são sua principal fonte de renda, também são proprietários de administradoras de imóveis ou de outras atividades do ramo financeiro. Dessa forma, o processamento só considerou a pessoa sendo do ramo financeiro se ela só tiver empresas deste setor.

Para a segunda pergunta, optou-se por só classificar uma pessoa física como sócio dentro de um setor específico se o total doado fosse maior que um salário-

mínimo. Ou seja, considerou-se que doações acima desse valor não seriam feitas por um trabalhador ou pessoa assalariada e sim por alguém que tivesse algum empreendimento ou vivesse de renda.

Uma outra limitação que deve ser considerada é a temporal. Como a base de CNPJ utilizada na pesquisa é um retrato do ano de 2020, não seria possível realizar a associação de um sócio que tivesse se retirado da sociedade, mas que tivesse feito doações antes de sua saída, ou seja, este ex-sócio não seria associado à empresa. Ou seja, é possível que uma parcela dos sócios não entre nesse processamento em função deste entrave técnico.

Mesmo contando com essas limitações, buscou-se realizar esse processamento para que fosse possível realizar uma visualização aproximada do montante de sócios que tinham efetivamente doado para as campanhas dos deputados federais em 2018. O resultado pode ser visto na tabela abaixo:

Tabela 14 - Doadores por Setores Econômicos em 2018

Setor do Doador	Total de Doadores	Total Doado	%
Agricultura	533	12.703.019	7,61%
Comércio	2.265	24.967.899	14,95%
Financeiro	767	7.748.496	4,64%
Serviços	6.151	43.158.023	25,84%
Indústria	793	13.093.972	7,84%
Infra-Estrutura	1.093	15.916.920	9,53%
Pequeno Doador	31.057	37.037.996	22,18%
Grande Doador	1.416	12.372.758	7,41%
Total	44.075	166.999.083	100,00%

Fonte: Dados da pesquisa a partir dos dados do TSE e do CNPJ

A partir do processamento, verificou-se que mais de 70% do total doado por pessoas físicas foram doados por sócios de empresas, ou seja, apesar do elevado número de doadores únicos que não tiveram nenhuma associação na base de CNPJs (representados nas linhas Pequeno e Grande Doador), o montante doado por empresários foi muito elevado.

Dentre os não sócios, houve o cuidado de dividi-los em dois grupos: os grandes doadores (que tivessem doado acima de 10.000,00) e as demais pessoas. Ou seja, existiam 1.416 pessoas que fizeram doações de grande monta, mas não se

encontrou nenhum registro societário na base de CNPJs. Pode-se conjecturar que tais pessoas vivem de renda, são funcionários públicos ou pensionistas do alto escalão da administração pública. Efetivamente, através dos dados existentes não é possível chegar a uma conclusão. O total doado por faixa de doação pode ser visualizado na tabela 15.

Tabela 15 - Faixas de Doação em 2018

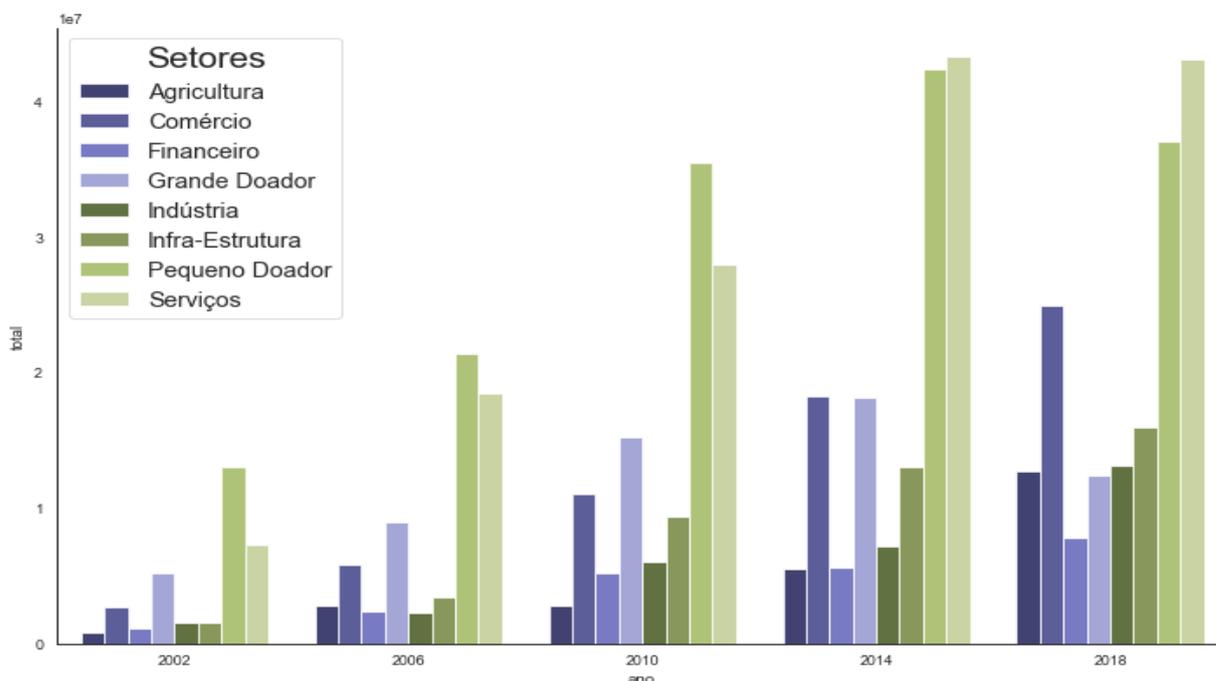
Tipo de Doador	Nº Doadores	Total Doado (R\$)	%
Não Sócio ou quem doou menos de 1k	19.657	11.384.510	6,82%
Sócio que doou entre 1k e 10k	18.635	46.866.636	28,06%
Sócio que doou entre 10 e 50k	4.143	39.615.764	23,72%
Sócio que doou 50 e 100k	687	14.388.956	8,62%
Sócio que doou acima de 100k	953	54.743.217	32,78%
Total	44.075	166.999.083	100,00%

Fonte: Dados da pesquisa a partir dos dados do TSE e do CNPJ

Ao analisar a divisão por setores, foi possível verificar que os números destoam bastante do resultado apresentado por SANTOS (2020), pois mostram que a maioria das doações veio de sócios dos setores de Serviços e Comércio, ou seja, o conjunto de doações para deputados federais em 2018 tem uma dinâmica bem diferente da apresentada nas eleições de 2014. Entretanto, ao executar o mesmo processamento para eleições anteriores, verifica-se que o mesmo padrão de financiamento por sócios se repete, tal como demonstrado no gráfico 10.

Gráfico 10 – Doações por Setor Econômico – Somente para Deputados Federais

Em milhões de reais

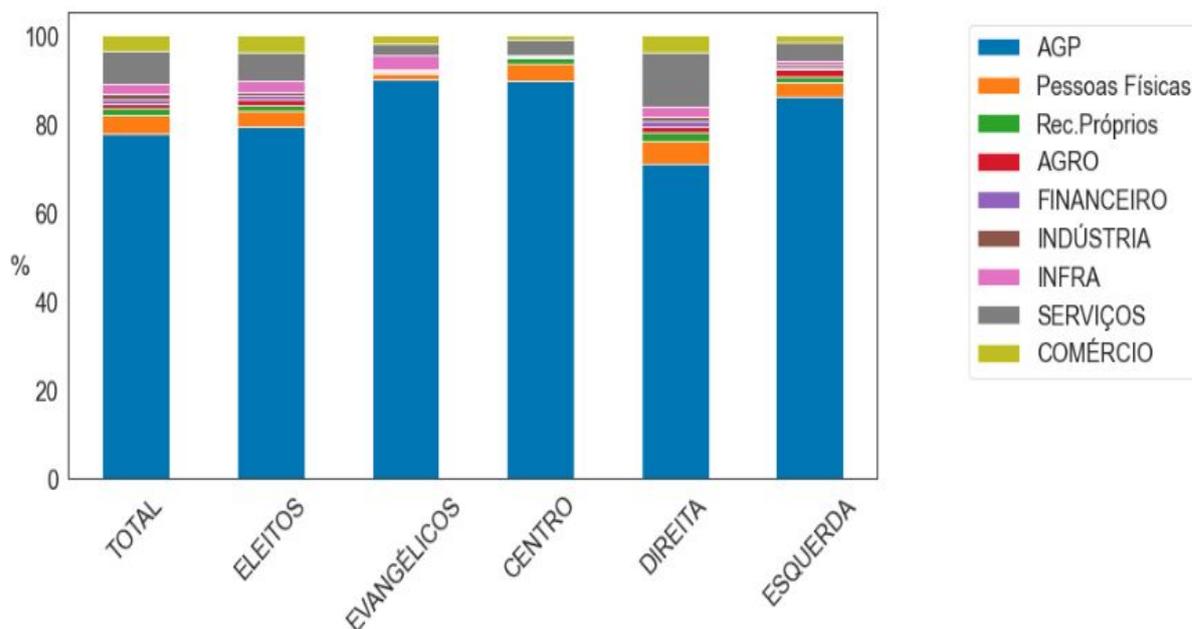


Fonte: O autor. Baseado nos dados do TSE

Ou seja, com o fim do financiamento empresarial, houve apenas um pequeno aumento nas doações de sócios nas campanhas para deputado federal, com uma variação maior no Agronegócio e no Comércio. A partir do processamento realizado, foi possível utilizar os dados agregados como base comparativa entre os deputados, ou seja, foi possível verificar se existe variação dos tipos de doação por partido, por tipologia ideológica (Esquerda, Centro e Direita) e também para os grupos dos deputados evangélicos, que é o foco da pesquisa.

Como é possível verificar no gráfico 11, não houve uma variação considerável entre os deputados de cada partido se agruparmos por bloco ideológico, ou seja, todos os blocos apresentam percentuais semelhantes. O único setor que se destaca um pouco dos demais é o setor de serviços, pois o percentual doado aos partidos de Direita é 10% superior ao compararmos com os blocos de Esquerda e Centro.

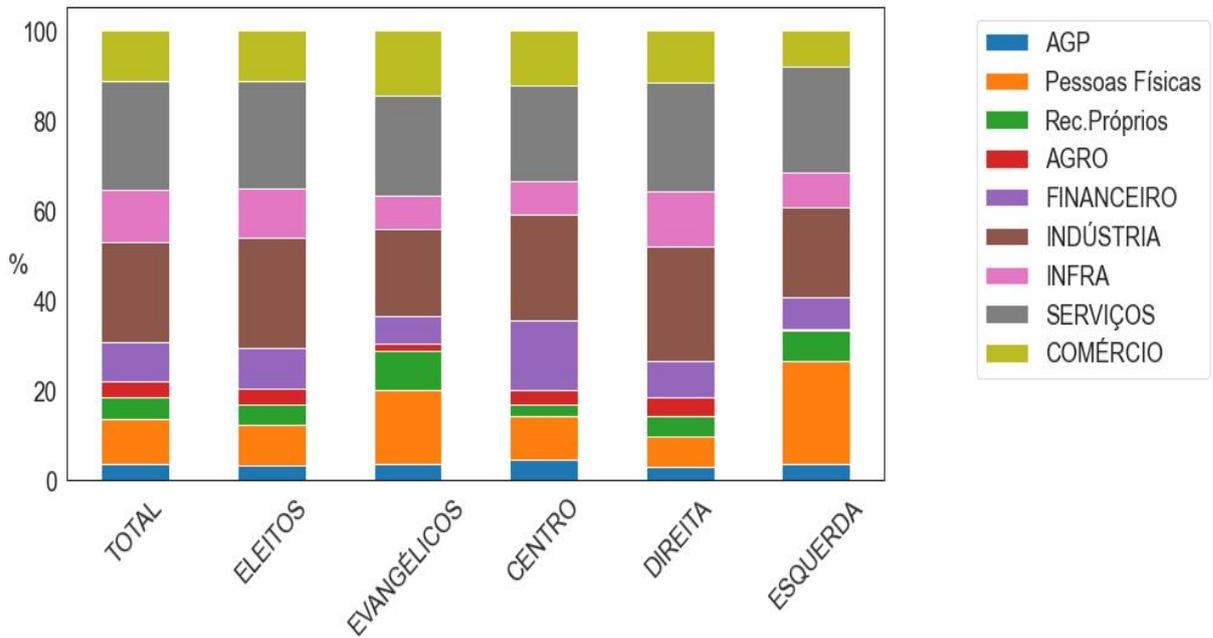
Gráfico 11 – Doação por Setor Econômico e por Bloco Ideológico em 2018



Fonte: O autor. Baseado nos dados do TSE

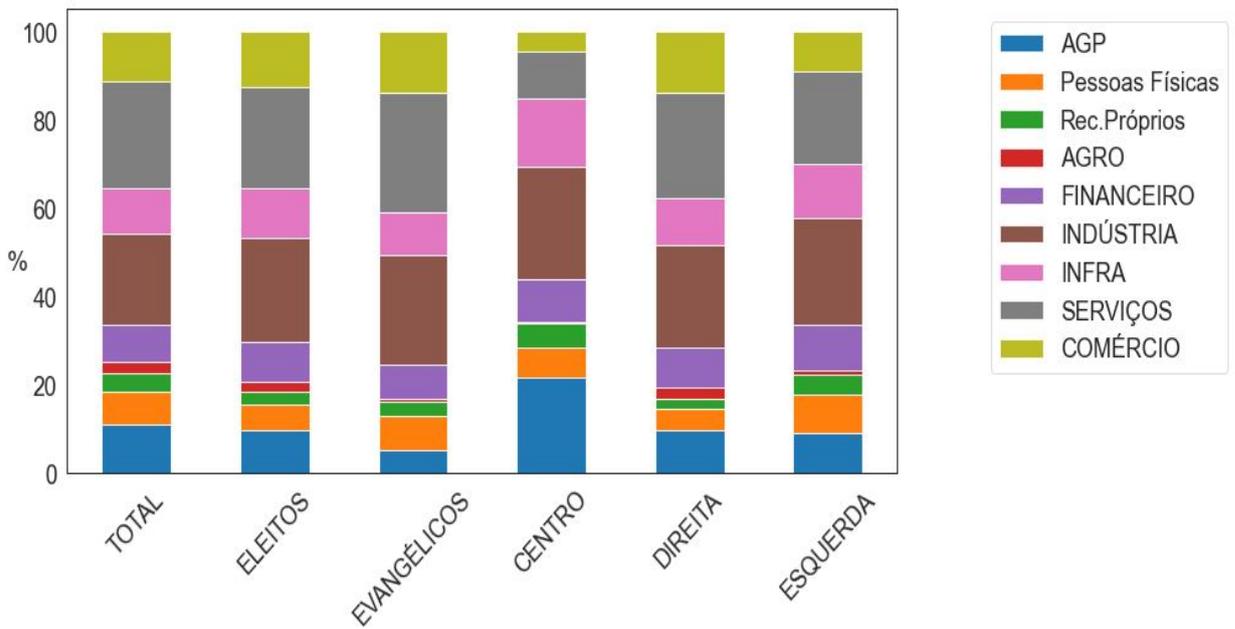
De forma análoga, os deputados evangélicos eleitos também apresentam o mesmo padrão de financiamento se compararmos com os demais deputados eleitos, apresentando apenas uma pequena variação de dez por cento no montante recebido pelos agentes partidários (AGPs). O mesmo padrão de financiamento ocorreu nos anos anteriores tal como pode-se observar nos gráficos de 12 a 16, referentes aos anos de 2002, 2006, 2010 e 2014.

Gráfico 12 – Doação por Setor Econômico e por Bloco Ideológico em 2002



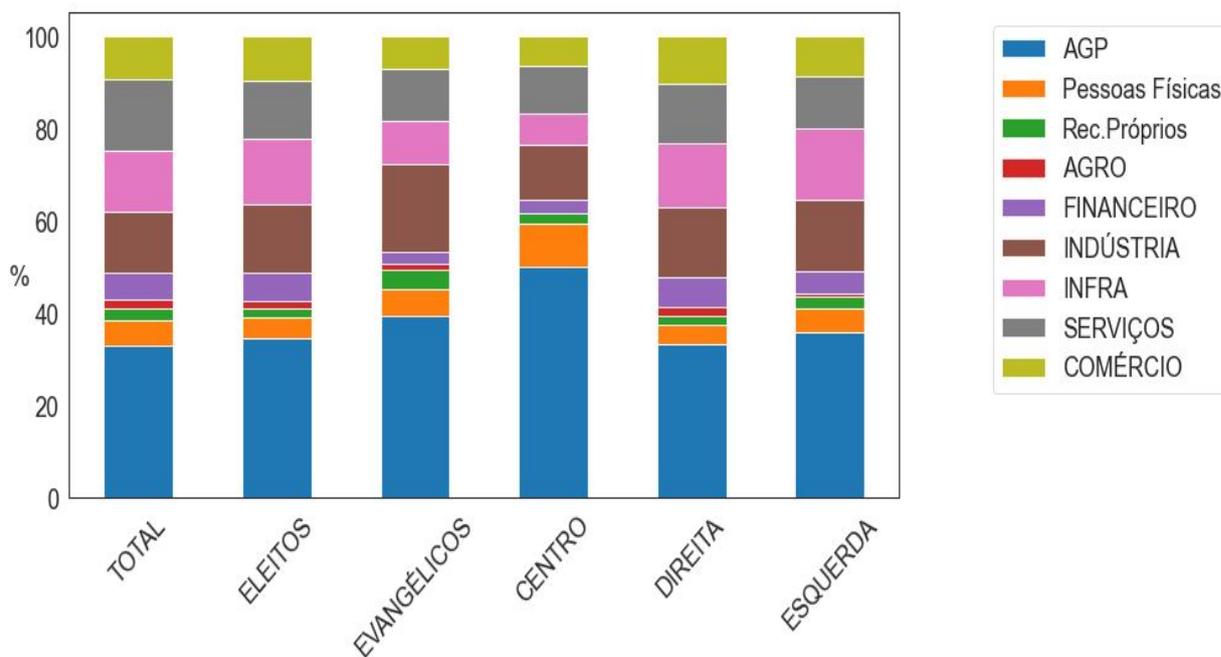
Fonte: O autor. Baseado nos dados do TSE

Gráfico 13 – Doação por Setor Econômico e por Bloco Ideológico em 2006



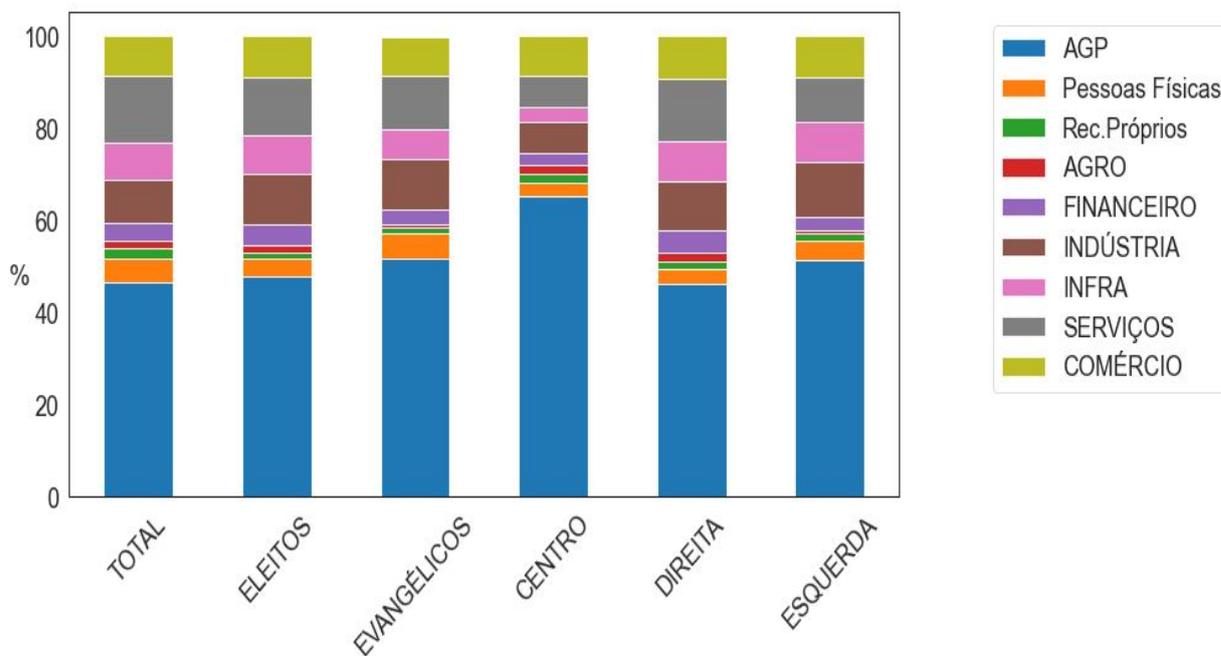
Fonte: O autor. Baseado nos dados do TSE

Gráfico 14 – Doação por Setor Econômico e por Bloco Ideológico em 2010



Fonte: O autor. Baseado nos dados do TSE

Gráfico 15 – Doação por Setor Econômico e por Bloco Ideológico em 2014



Fonte: O autor. Baseado nos dados do TSE

Alguns pontos são interessantes notar nos gráficos. O principal é que a participação dos agentes políticos (diretório nacional e outros candidatos) foi crescendo ao longo do tempo. Esse crescimento segue o mesmo padrão do

financiamento geral (que engloba todos os cargos) e que foi demonstrado no gráfico 7. Além disso, é possível notar que a participação por parte de pessoas físicas foi maior em 2002 se comparada aos anos seguintes. Entretanto, esse número pode ter sido inflado pela falta de informações sobre sócios, ou seja, quanto mais antiga é a eleição, maior a probabilidade do sócio não ter sido encontrado na base de CNPJ e, dessa forma, o seu CPF não foi associado a nenhum setor econômico.

Vale ressaltar também uma maior participação da indústria e do setor de serviços nas eleições de 2002 e 2006 enquanto que o setor de infraestrutura passou a ser preponderante em 2010 e 2014. Um outro ponto que merece destaque é o maior uso dos recursos partidários por partidos de Centro, ou seja, ao longo de todas as eleições, os candidatos de Centro receberam um maior aporte de recursos por parte dos agentes partidários.

Entretanto, o ponto central que interessa a esta pesquisa foi a verificação de que não houve uma grande diferença entre o padrão de financiamento dos deputados evangélicos eleitos em relação aos demais e que houve pouca variação nos tipos de doações entre os blocos ideológicos. Desta forma, a montagem e validação destas informações abre caminho para a análise principal da tese, onde se tentará categorizar os deputados federais evangélicos a partir do padrão de financiamento dos mesmos utilizando técnicas de mineração de dados.

6 RESULTADOS

Inicialmente, é importante relembrar a proposta da pesquisa definida na Introdução da tese onde foram levantadas duas hipóteses: a primeira de que o financiamento de deputados evangélicos tem um padrão diferente dos demais candidatos eleitos e a segunda é que os deputados evangélicos não compõem um grupo homogêneo e que, a partir do uso de técnicas de mineração de dados, é possível classificar os deputados em duas ou mais categorias a partir dos padrões de financiamento de suas campanhas.

Ao longo do capítulo, buscou-se um texto o mais didático possível para leitores menos familiarizados com os termos estatísticos ou computacionais pudessem acompanhar o raciocínio geral sem a necessidade de buscar uma bibliografia externa.

6.1 MONTAGEM DO BANCO DE DADOS

O primeiro passo para a obtenção dos atributos dos deputados evangélicos foi montar a base de resultados eleitorais de 1998 a 2018. A montagem do banco de dados que deu suporte à pesquisa foi resultado de 5 anos de pesquisa e desenvolvimento. Desde o início de 2017, buscou-se agregar dados não só para esta tese como para outras pesquisas que pudesse fazer uso dos dados agregados das eleições e do financiamento de campanha eleitoral. As 4 principais fontes de dados que compõe o banco de dados foram a base de candidatos que concorreram às eleições, a base de resultados eleitorais (com os votos de cada candidato - eleito ou não), a base de financiamento de campanha e por fim a base de candidatos evangélicos.

Especificamente para os dados de financiamento de campanha, a montagem do banco de dados foi realizada para os candidatos a deputados federais a partir de 2002 já que os dados de financiamento de campanha anteriores a este período não continham o detalhamento completo por agente partidário e o número de inconsistências da base do TSE foi considerável, com muitos CPFs e CNPJs inválidos e muitos registros sem a indicação da origem da receita.

Para que se possa ter uma visão mais completa do processo eleitoral, algumas variáveis que influenciam o processo também foram adicionadas. Conforme foi visto no capítulo 2, o objetivo da adição destas variáveis foi tentar aumentar a

precisão do processo de mineração de dados pois quanto mais variáveis e atributos fossem adicionados na análise (tendo o cuidado de não utilizar variáveis com alta correlação) melhor será a qualidade do processamento.

Dessa forma, buscou-se adicionar as seguintes bases no escopo de pesquisa:

- Base do CNPJ da Receita Federal que contém os dados de todas as empresas no Brasil, seus respectivos sócios e em que setores econômicos pertencem;
- Número de Filiados por partido ao longo dos anos e o número de deputados eleitos por partido ao longo dos anos. Estas duas variáveis têm o objetivo de quantificar minimamente o tamanho do partido.

Em especial para a base de CNPJs, o objetivo foi buscar se o doador era sócio de alguma empresa. Tal como foi visto no capítulo anterior, utilizou-se a base do CNPJ para associar o CPF do doador com a empresa da qual é sócio e assim tentar encontrar o setor econômico ligado ao doador. Dos 25 milhões de sócios da base da receita, foram encontrados 1.211.274 que fizeram alguma doação ou que participaram das eleições como candidatos.

6.2 VARIÁVEIS UTILIZADAS

As seguintes variáveis foram utilizadas:

- SEXO: se o candidato é do sexo masculino (0) ou feminino (1), considerando que o TSE não abre a possibilidade da seleção de gênero;
- EVANGÉLICO: 0 ou 1. Se o candidato se autodenomina pertencente a uma alguma igreja ou denominação evangélica a partir do resultado obtido no capítulo 4;
- INCUMBÊNCIA: (0) se está tentando um novo mandato ou (1) se é um novo entrante no cargo pretendido;
- CAPILARIDADE: Índice de 0 a 1 que indica quão ampla (sob o ponto de vista geográfico) foi a votação do candidato. Quanto maior o número de zonas eleitorais onde o candidato recebeu votos, maior o índice. Este indicador é a média dos percentis⁶⁰ dos votos em cada zona de cada um dos candidatos.

⁶⁰ Percentil é uma pontuação que varia entre 0 a 100 e que indica em que posição um elemento se encontra em uma distribuição de frequência. Desta forma, é possível identificar em que posição

- VOTOS PROPORCIONAIS: Índice de 0 a 1 indicando em que percentil o candidato está posicionado em função dos votos recebidos no seu Estado. A opção pelo uso do percentil teve como objetivo nivelar os candidatos em relação aos estados em que concorrem ao pleito. Desta forma, o candidato mais votado de SP, com por exemplo 200 mil votos, estaria no mesmo percentil que o mais votado de Roraima que obteve 10 mil votos já que ambos estão entre os mais votados em seus respectivos estados.
- RECEITA TOTAL: Valor Total Recebido pela campanha do candidato (em Reais);
- RECEITA PARTIDÁRIA: Total de Doações realizadas pelo Agente Partidário;
- RECEITA PRÓPRIA: Doação que o próprio candidato investiu em sua campanha;
- RECEITA INDÚSTRIA: Doações feitas por sócios de empresas de setores da indústria tal como documentado no capítulo 5;
- RECEITA INFRA-ESTRUTURA: Idem para o setor de infra-estrutura;
- RECEITA AGRICULTURA: Idem para o setor agrícola;
- RECEITA COMÉRCIO: Idem para o setor de comércio;
- RECEITA SERVIÇOS: Idem para o setor de serviços;
- RECEITA INST.FINANCEIRAS: Idem para o setor específico dos bancos e outras instituições financeiras e imobiliárias;
- IDEOLOGIA do Partido: será utilizado o índice de Power e Silveira-Rodrigues (2018) que varia de -1 a 1.
- Tamanho do Partido: índice variando de 0 a 1 indicando o decil⁶¹ ao qual o partido pertence em relação ao número de deputados federais eleitos no início de cada legislatura. Esse é um dos fatores utilizado pelo TSE para o cálculo do fundo partidário.

(ranking) um determinado objeto está em relação aos demais. Por exemplo, o percentil 50 (a mediana) é a pontuação abaixo da qual (exclusiva) ou na qual ou abaixo da qual (inclusive) 50% das pontuações na distribuição podem ser encontradas.

⁶¹ O decil segue a mesma lógica que o percentil só que a distribuição é feita sobre 10 partes ao invés de 100. O uso de decil é utilizando no lugar do percentil quando o número de elementos é pequeno, como no caso dos partidos existentes.

- Número de Filiados: índice variando de 0 a 1 indicando o decil ao qual o partido pertence em função do número de filiados no ano da eleição em que os deputados foram eleitos⁶².

Vale ressaltar que algumas variáveis comumente utilizadas em outros estudos não foram contempladas no processamento. São elas: RAÇA e OCUPAÇÃO. A ocupação não foi utilizada pois seguimos o estudo de Peixoto (2016) que indicou que tal variável não tem influência no pleito. Já a RAÇA não foi utilizada pois foram encontradas muitas inconsistências na base do TSE em relação a esta variável o que demandaria um estudo mais aprofundado sobre este tema específico.

A partir da seleção das variáveis, foi calculada a matriz de correlação entre todas as variáveis para verificar se não havia uma alta correlação entre duas ou mais variáveis fato este que causaria uma distorção na análise das clusterizações.

6.3 ANÁLISE ESTATÍSTICA E REGRESSÃO LOGÍSTICA

Para a primeira hipótese, foi utilizado o método de seleção estatística ou redução dimensional de variáveis para selecionar as variáveis que têm maior contribuição na explicação da variância observada nos dados, e, a partir daí, foi utilizado um modelo de regressão logística para quantificar os efeitos das variáveis incluídas no modelo final.⁶³

O modelo de regressão logística foi escolhido pois a variável dependente é binária, isto é, busca-se um modelo explicativo para quantificar o efeito do financiamento recebido durante a campanha por deputados federais eleitos em 2018 na probabilidade de um(a) deputado(a) federal ser ou não evangélico(a). O uso de tal modelo baseados em regressão logística tem sido frequentemente utilizados em pesquisas na área de Ciência Política (FERNANDES et al, 2020).

O primeiro passo foi a seleção das variáveis com maior contribuição para a explicação da variância. A técnica de seleção estatística de variáveis “para a frente” (*stepwise forward-selection*) é utilizado para reduzir o número de variáveis que

⁶² O número de filiados pode ser obtido no link: <https://www.tse.jus.br/eleitor/estatisticas-de-eleitorado/filiados>

⁶³ A análise estatística e a regressão logística contou com a consultoria do Professor e Estatístico Felipe Figueiredo que realizou o processamento e validação do modelo. O código fonte e o relatório técnico pode ser obtido em <https://github.com/philsf-biostat/SAR-2021-012-JG/>

contribuam para a análise (DRAPER e SMITH, 1981). Foi usado o critério de $p < 0,20$ para inclusão no modelo multivariado.

Os modelos criados foram comparados com o Critério de informação de Akaike (AIC, do inglês, *Akaike's Information Criterion*) que tem o objetivo de medir a qualidade de um modelo estatístico visando também a sua simplicidade. Este índice tem o objetivo de encontrar o modelo que se ajusta melhor aos dados, ao mesmo tempo em que penaliza modelos que utilizam um número excessivo de parâmetros. De forma objetiva, quanto menor o índice, melhor será o modelo.

Tabela 16 - Características dos deputados federais eleitos em 2018

Características	Outros, N = 398	Evangélico, N = 116
Receita total (milhão R\$)	1.12 (0.68)	1.08 (0.69)
Capilaridade (10%)	7.70 (1.57)	8.05 (1.60)
Reeleição vs primeiro mandato		
Primeiro mandato	346 (87%)	94 (81%)
Reeleito	52 (13%)	22 (19%)
Sexo		
Masculino	346 (87%)	91 (78%)
Feminino	52 (13%)	25 (22%)
Votos (100k)	0.97 (0.58)	1.27 (1.99)
Índice de Power e Silveira-Rodrigues	0.17 (0.50)	0.42 (0.30)
Decil do núm. de filiados	0.79 (0.22)	0.70 (0.20)
Decil do núm. de deputados	0.78 (0.22)	0.74 (0.22)

Fonte: O autor

Após o processamento, a seleção *stepwise* de variáveis incluiu no modelo o partido, o sexo, a capilaridade e o número de votos dos deputados, ajustando a estimativa do efeito da receita total e descartou a UF e a incumbência (se foi a primeira eleição do candidato para o cargo). A seleção das variáveis causou uma redução do AIC de 584,99 para 489,63, o que indica um ajuste substancialmente melhor do modelo aos dados.

O próximo passo foi processar os modelos de regressão⁶⁴. A tabela 17 mostra o resultado com as estimativas de efeito de cada variável que compõe o teste da hipótese e nela é possível constatar de que **há associação entre a receita total**

⁶⁴ Os modelos foram criados sem intercepto e as análises foram realizadas ao nível de significância de 5% e todos os testes de hipóteses e intervalos de confiança calculados foram bicaudais.

recebida durante a campanha pelos deputados federais e a classificação como evangélico. A estimativa bruta (modelo univariado) está significativamente associada com a razão de chances de ser evangélico quando comparado aos não evangélicos e indica que os evangélicos tendem a receber 60% menos verba que os outros, tendo o *odds ratio* (OR) de 0,41 (95% CI 0,35, 0,49; $p < 0,001$).

Tabela 17 - Efeito da receita total recebida por deputados federais

Características	OR ¹	95% CI ¹	valor p
Receita Bruta (Não ajustada)	0.41	0.35, 0.49	<0.001
Receita Bruta (milhão R\$)	0.77	0.46, 1.27	0.3
Sexo			
Masculino	—	—	
Feminino	3.18	1.61, 6.28	<0.001
Capilaridade (10%)	1.25	1.05, 1.49	0.012
Votos (100k)	1.26	1.01, 1.73	0.10

A razão de possibilidades (em inglês: *odds ratio*; abreviatura OR) é definida como a razão entre a chance de um evento ocorrer em um grupo e a chance de ocorrer em outro grupo, ou seja, quanto maior a razão, maior a probabilidade.

A estimativa ajustada do efeito da receita total indica um efeito menor que a estimativa bruta, onde a receita dos deputados evangélicos é aproximadamente 23% menor que os outros deputados (OR: 0,77 (95% CI 0,46-1,27; $p=0.3$)). A OR ajustada não foi significativamente diferente de 1.

A interpretação conjunta destes dois resultados é que, na média, a receita total dos deputados evangélicos é menor que a receita total dos outros deputados. No entanto, o tamanho do efeito da receita, ajustado pelas outras variáveis, não pode ser estimado com precisão suficiente para quantificar sua magnitude.

Um outro resultado expressivo foi em relação à variável sexo. **As deputadas mulheres pareciam ser maioria entre os evangélicos, sendo três vezes mais prováveis de serem evangélicas quando comparadas aos homens** (OR: 3.18 (95% CI 1.61, 6.28; $p < 0.001$)).

Outro ponto a ser considerado foi a capilaridade onde **os deputados evangélicos atingem uma capilaridade 25% maior quando comparados aos outros deputados** (OR: 1.25 (95% CI 1.05, 1.49; $p=0.012$)).

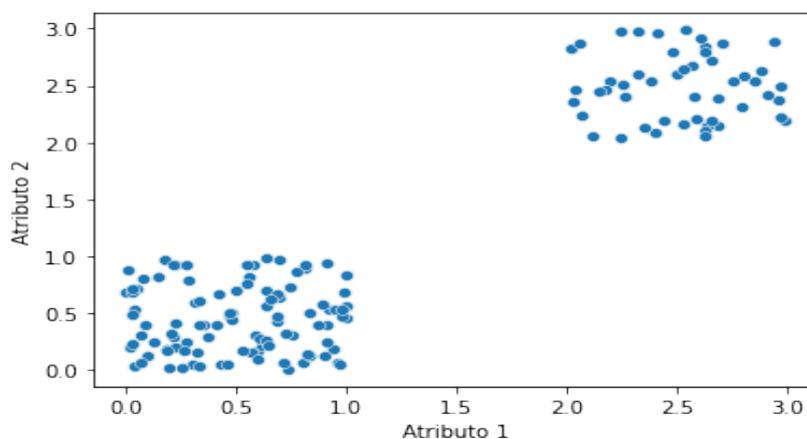
Por fim, o número de votos parece ser um fator de confundimento. O número de votos não está na cadeia causal entre a receita e a classificação como evangélico, e também está associado tanto com a classificação como evangélico (OR: 0.49 (95% CI 0.41, 0.59; $p < 0.001$)), como com a receita total (0.50 (95% CI 0.44, 0.56; $p < 0.001$)). Como os critérios de confundimento foram preenchidos, esta variável foi mantida no modelo final para corrigir a estimativa do efeito da receita no desfecho do estudo.

6.4 SEGMENTAÇÃO

Como já foi visto no capítulo dois, uma das técnicas utilizadas para segmentar uma população é denominada clusterização (*Clustering*). O processo de clusterização clássico tem 2 entradas: o conjunto de objetos que se quer classificar e uma indicação de quantos grupos se quer formar. O resultado do processamento é uma nova informação que indica em qual grupo cada um dos objetos está contido.

Um exemplo que pode facilitar a visualização deste tipo de processamento é o desenho de um grupo de objetos hipotéticos que tenham 2 atributos numéricos (por exemplo, a altura e o peso) e estes dois atributos são representadas em um plano cartesiano, onde o eixo X representa um atributo e o eixo Y representa o segundo atributo. Dessa forma, cada ponto no gráfico representa um objeto que queremos classificar:

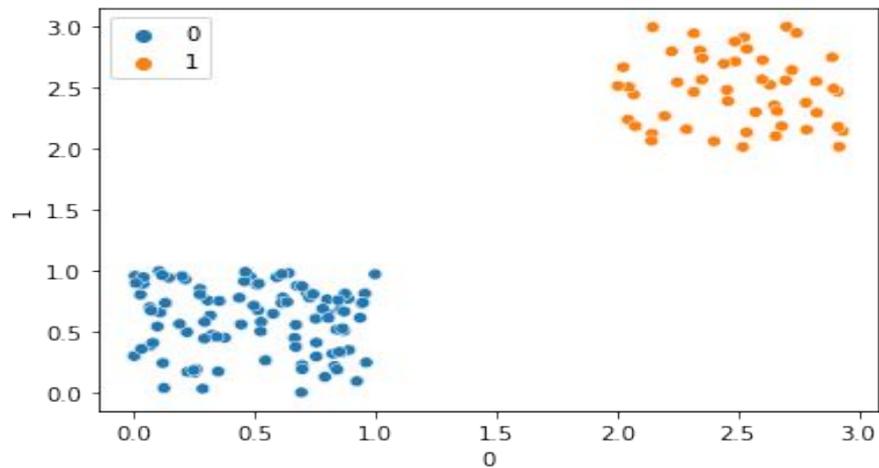
Gráfico 16 – Conjunto de Dados a Segmentar



Fonte: O autor.

Neste exemplo, não seria necessário processar nenhum algoritmo, pois a visão humana consegue perceber que existem 2 grupos distintos no conjunto apresentado. E ao aplicar um algoritmo de clusterização para separar os elementos, o resultado sai conforme o esperado, tal como podemos visualizar na figura a seguir.

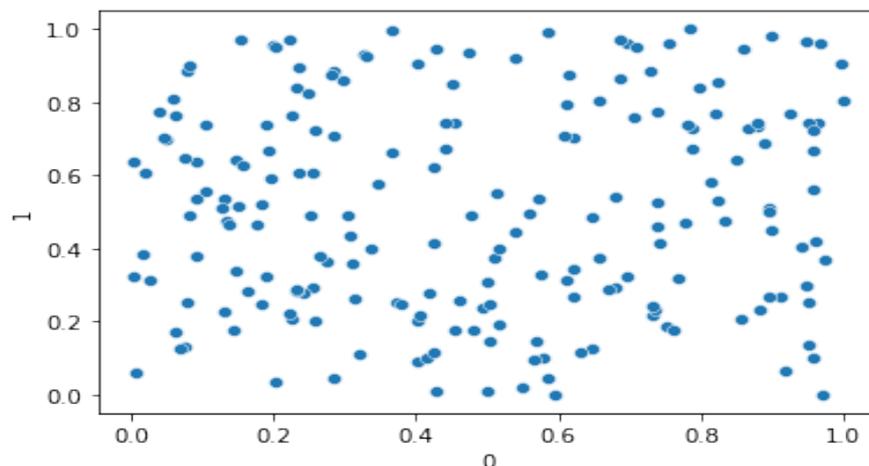
Gráfico 17 – **Resultado da segmentação**



Fonte: O autor.

Entretanto, dependendo da complexidade do conjunto de dados, a visão humana já não consegue identificar grupos tão facilmente como o exemplo do gráfico 18 onde os objetos são dispostos em outra configuração já que seus atributos são diferentes dos objetos apresentados no exemplo do gráfico 17. Além disso, surge outro problema: será que existem apenas 2 categorias ou seria possível classificar o conjunto de dados em 3 ou 4 categorias? Como calcular qual seria a melhor quantidade de grupos que pudesse trazer uma melhor interpretação?

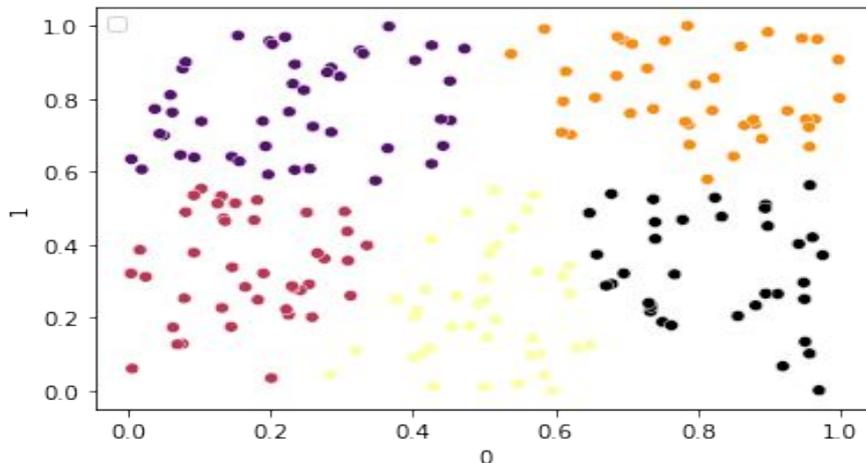
Gráfico 18 – **Exemplo de segmentação Complexa**



Fonte: O autor.

Ao aplicar a clusterização para o exemplo acima, o algoritmo irá sugerir uma possível segmentação do conjunto, tal como indicado no gráfico 19.

Gráfico 19 – Exemplo de segmentação Complexa



Fonte: O autor.

O processo pode ficar ainda mais complexo se pensarmos na quantidade de atributos que existem para cada objeto. Nos exemplos acima, utilizamos apenas 2 atributos para fins didáticos para que fosse possível representá-los em um plano cartesiano. Em conjuntos de dados reais, os objetos tem dezenas ou centenas de atributos. E nem precisava ter tantos pois acima de 3 dimensões já não se conseguiria desenhar uma representação do objeto em um gráfico.

Por fim, pode-se problematizar ainda mais a questão: será que a classificação encontrada tem alguma relação com a realidade? Efetivamente, qualquer clusterização tem como resultado uma tipologia com ao menos 2 *clusters* mas como medir se os *clusters* resultantes não só têm boas propriedades estatísticas (compactos, bem separados e conectados) mas se também indicam resultados relevantes. Como medir a qualidade destes processamentos?

Para resolver a questão quantitativa, existem diversos índices que buscam medir a qualidade de uma clusterização (JAIN, 2010). O índice Silhoutte (ROUSSEEUW, 1987) é um dos mais utilizados por ser de fácil interpretação. É um índice que varia de -1 a 1 onde 1 seria uma clusterização de ótima qualidade, isto é, os objetos de cada cluster terão características bem semelhantes e haverá pouca ou nenhuma dúvida se um objeto pertence a um cluster ou a outro.

Por exemplo, ao calcular o índice silhouette do modelo “ideal” apresentado no gráfico 16, obtem-se um valor de 0,81, ou seja, bem próximo de 1. E se os 2 grupos fossem ainda mais coesos, o índice ficaria ainda mais próximo de 1.

Dessa forma, se o resultado da clusterização sobre o conjunto de deputados evangélicos obtiver um índice de boa qualidade e se os clusters encontrados fizerem algum sentido sob a ótica do que já foi estudado no referencial teórico que embasa a pesquisa, poderemos considerar que a hipótese foi confirmada.

6.5 PROCESSAMENTO

Em função da natureza dos algoritmos de clusterização, as variáveis utilizadas devem ser numéricas e normalizadas para que estejam dentro de uma mesma ordem de grandeza. Ou seja, não se pode comparar variáveis na casa de milhões com variáveis que variam de 0 a 1 pois tal comparação tornaria o cálculo dos algoritmos inconsistente. Dessa forma, para as variáveis que estavam fora dessa ordem de grandeza, optou-se por criar novas variáveis que representassem seus valores originais “equalizados” para que ficassem dentro da faixa de 0 a 1 (com 2 ou 3 casas decimais).

Para realizar todo o processamento foram utilizadas as bibliotecas em Python SciKitLearn (PEDREGOSA, 2011) e HDbScan (MCINNES et al, 2017) e os seguintes algoritmos foram utilizados: k-means, Agglomerative Clustering, HdbSCAN.

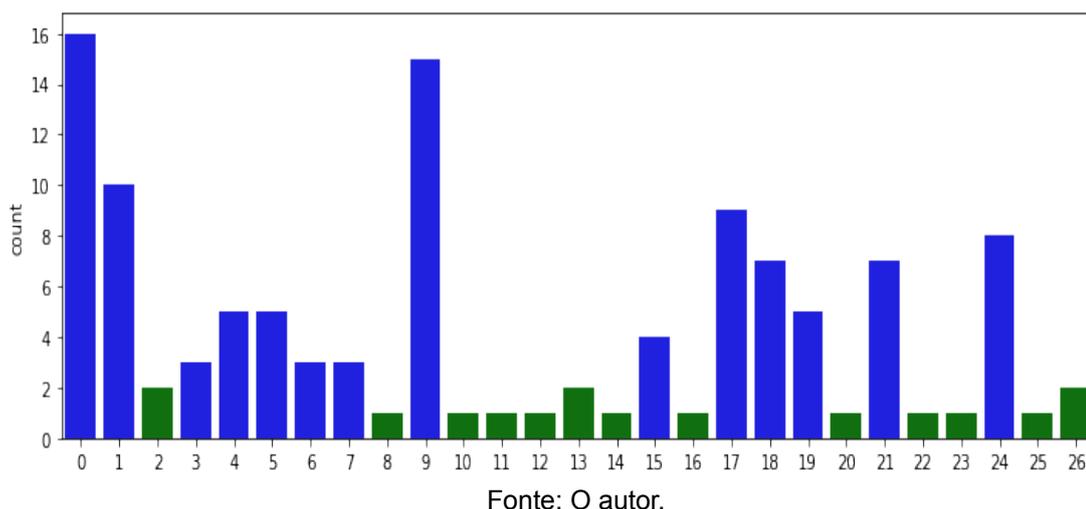
O primeiro processamento foi realizado com o método K-Means que é um dos mais comuns para se realizar o processo de clusterização. Nesta tentativa, buscou-se montar de 2 a 20 clusters e o índice silhouette foi calculado para cada processamento. O resultado não foi muito animador, pois em todos os processamentos, o índice silhouette ficou abaixo de 0,3. O resultado foi ainda mais sofrível ao utilizar os algoritmos Agglomerative Clustering e HdbSCAN.

O processamento que obteve o melhor índice indicava 27 clusters. A grande quantidade de clusters formados podia sugerir que o número de variáveis estava prejudicando o algoritmo, ou seja, algumas variáveis estavam embutindo uma complexidade ao modelo e por isso o número calculado de clusters se tornava muito grande. Isso indicava que um processamento de redução de variáveis pudesse melhorar o modelo. O processo de redução de variáveis é um processo matemático

que consiste em encontrar as características que são as mais influentes no modelo e descartar as demais.

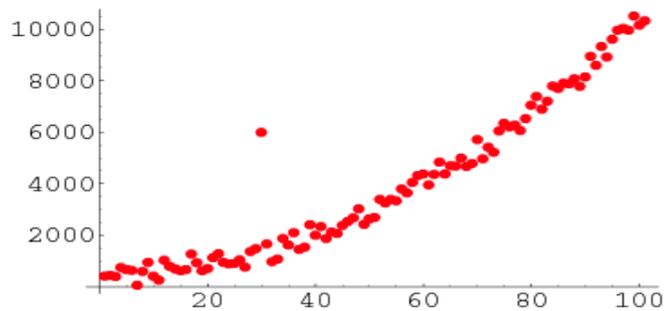
Um segundo fator importante foi ressaltado ao se analisar o histograma da clusterização: verificou-se que a grande maioria dos clusters tinha apenas um único elemento. Essa análise poderia sugerir que o conjunto de dados continha muitos *outliers*, ou seja, havia muitos deputados que tinham um padrão de financiamento muito distinto dos demais. O gráfico 20 mostra o total de deputados que foram alocados em cada cluster e as barras em verde indicam os clusters com dois ou menos deputados.

Gráfico 20 – Histograma da primeira tentativa de segmentação



Os outliers são elementos de um conjunto que contém uma ou mais características que distoam dos demais elementos e, neste caso, seriam deputados que teriam características, seja de financiamento ou de perfil de votação muito diferente dos demais. O gráfico 21 é um exemplo de identificação de outliers quando trabalhamos com 2 ou 3 atributos mas quando o conjunto a ser estudado contém mais de 3 casos, são necessárias técnicas específicas para identificar quem são os possíveis outliers do conjunto de deputados evangélicos.

Gráfico 21 – Exemplo de Outlier

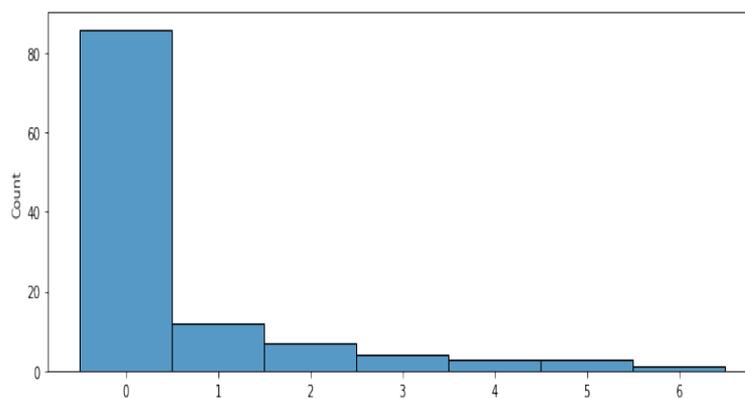


Fonte: O autor.

Assim, realizou-se um processamento de redução de variáveis que indicasse quais as variáveis que mais contribuíam para uma clusterização que apresentasse um índice de qualidade melhor que o encontrado no primeiro processamento. O processamento que obteve o melhor índice foi o que mantinha 10 variáveis. Dessa forma, as seguintes variáveis foram removidas do processamento: Capilaridade, Receita Total, Receita AGP, Filiados e Percentil de Deputados.

A partir daí, foi realizado um segundo processamento tentando encontrar o modelo com um índice silhouette melhor que o anterior. O segundo processamento obteve um resultado bem animador com um índice silhouette com o valor de 0,53 o que indicava que a clusterização já tinha uma boa qualidade. Entretanto, o processamento retornou 7 clusters ainda com uma configuração bem disforme tal como podemos visualizar no histograma representado pelo gráfico 22.

Gráfico 22 – Histograma do segundo processamento

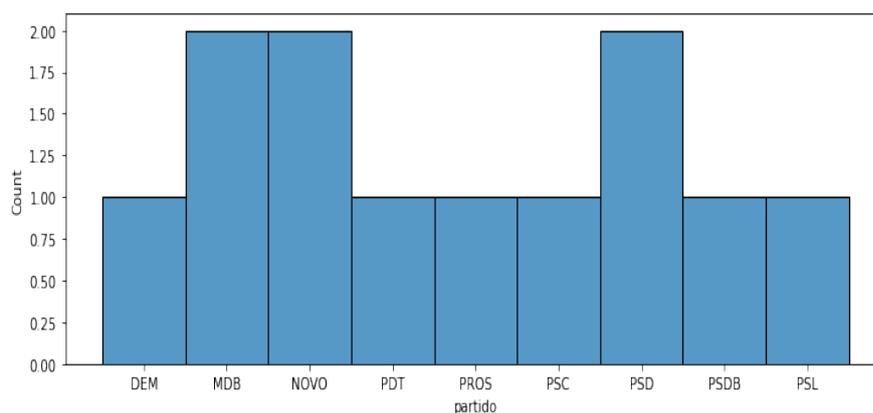


Fonte: O autor.

Novamente, o processamento indicava que deveriam haver outliers no conjunto de dados. Desta forma, utilizou-se o algoritmo *Isolation Forest* (LIU et al, 2008) para determinar os outliers do conjunto, considerando que 10% dos deputados seriam outliers. Este percentual foi calculado a partir do total de deputados pertencentes aos clusters que continham até dois elementos. Foram encontrados 12 deputados de 9 partidos tal como se pode observar no gráfico 23.

Ao analisar as características deste deputados que foram marcados como outliers, percebeu-se que todos tiveram um financiamento de empresários de diversos setores muito acima da média e receberam, também em média, menos recursos dos agentes partidários⁶⁵. Vale notar também que dois dos deputados Onyx Lorenzoni (DEM-RS) e Fábio Faria (PSD-RN) galgaram o posto de ministro do governo Bolsonaro.

Gráfico 23 – Outliers por Partido

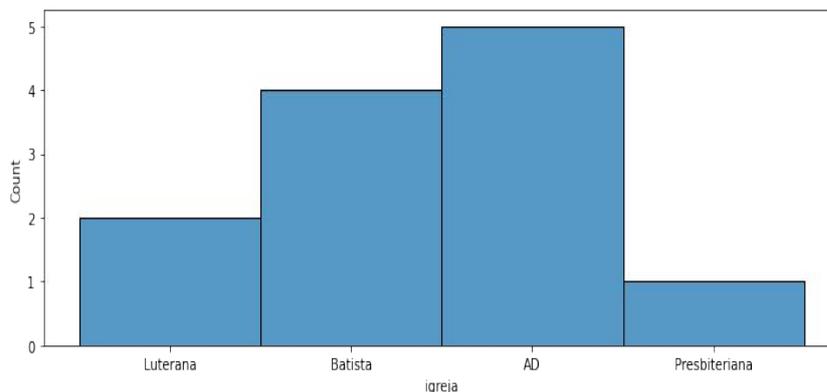


Fonte: O autor.

Já no histograma apresentado no gráfico 24, que detalha de quais denominações religiosas estes candidatos estavam associados, é possível notar um ponto interessante: com exceção dos deputados da Assembleia de Deus, nenhum deputado associado às corporações pentecostais está presente.

⁶⁵ A relação de deputados que foram selecionados como outliers pelo algoritmo estão descritos no Anexo I

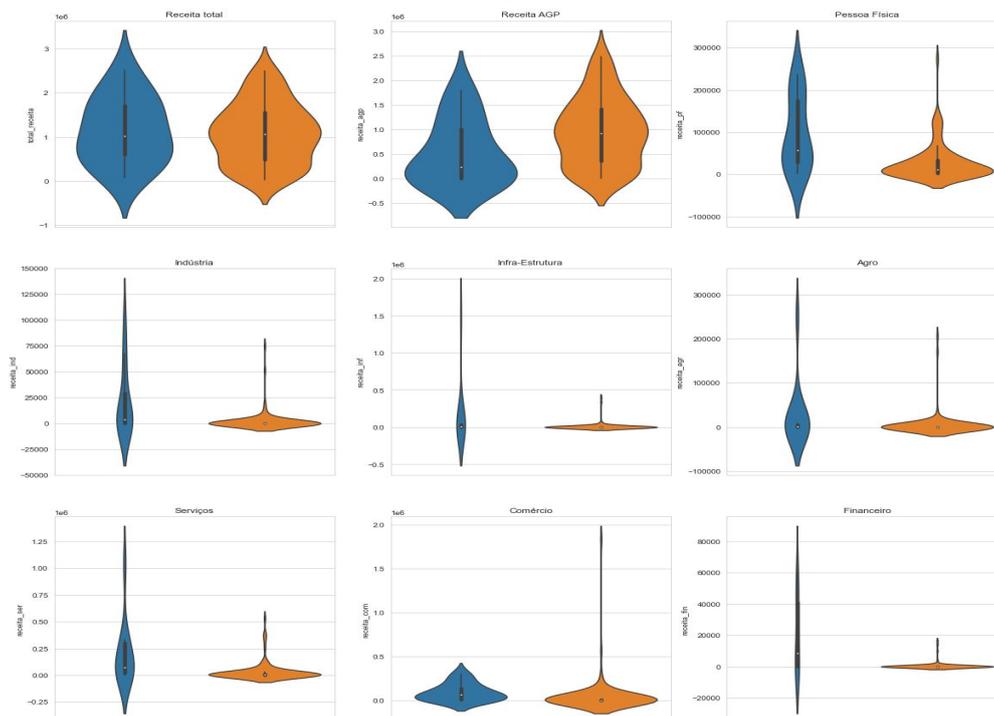
Gráfico 24 – Outliers por Denominação



Fonte: O autor.

Já o gráfico 25 busca comparar a distribuição do financiamento a partir de cada variável analisada. O gráfico escolhido para representar esta comparação foi o tipo violino. Este gráfico mostra em azul o grupo dos outliers e em laranja o restante dos deputados evangélicos. A largura do “violino” indica a distribuição dos deputados, ou seja, quanto mais largo é o “violino”, mais elementos o grupo contém e, dessa forma, é possível analisar a distribuição graficamente de todas as variáveis de forma compacta. A partir deste gráfico, é possível visualizar que o grupo dos outliers tem uma configuração semelhante na receita total mas tem deputados recebendo financiamento em todos os setores econômicos.

Gráfico 25 – Comparação dos Outliers



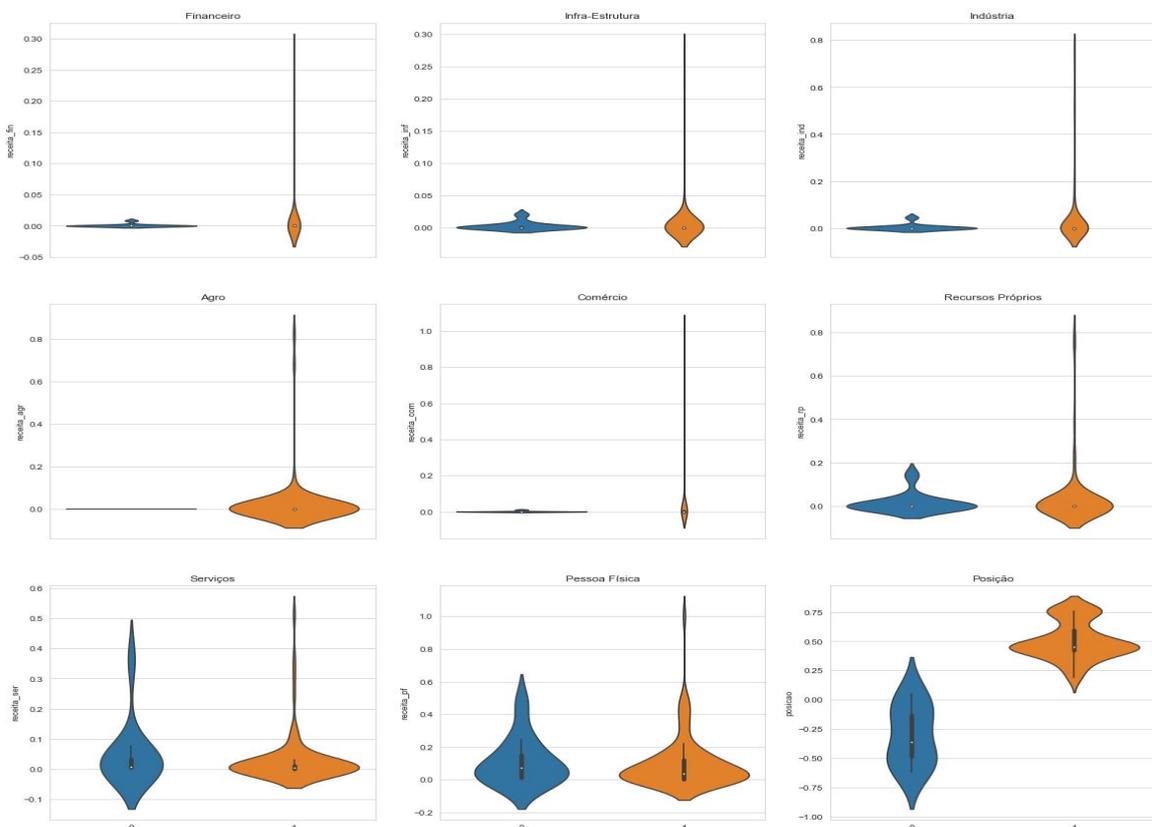
Fonte: O autor.

O próximo passo foi realizar uma nova tentativa de clusterização sem considerar os outliers. Ao analisar os números, pode-se visualizar resultados bem consistentes. Com o novo processamento, obteve-se um modelo com 2 clusters e com um índice silhouette de 0,558. O primeiro cluster (cluster 0) continha apenas 11 deputados e o segundo 93 (cluster 1).

Os deputados do cluster (0) pertenciam a um espectro político mais ao centro e à esquerda, contaram mais com um financiamento maior por parte de pessoas físicas, continham uma capilaridade ligeiramente menor que a média, pertenciam a partidos maiores, recebiam mais recursos do partido e a receita do setor de serviços era bem mais expressiva que a do segundo cluster.

Já os deputados do segundo cluster (cluster 1) pertencem quase que exclusivamente a partidos de Direita e recebem doações elevadas do Agronegócio. Contém todos os deputados ligados a IURD e quase todos ligados às outras corporações pentecostais como AD, Maranata e IEQ. A comparação de cada variável pode ser visualizada nos gráficos 26 e 27.

Gráfico 26 – Clusterização Final



Fonte: O autor.

Após esta clusterização final, ainda foram realizadas 4 outros tipos de clusterização para verificar se alguma outra configuração de clusters com uma qualidade maior seria obtida mas a relatada aqui se manteve com o melhor resultado. Também foram utilizados outros algoritmos e também foram testados outros índices que indicam a qualidade da clusterização, como o índice de Dunn, Hopkins e Davis e o método *Elbow*. Em todos os casos, os índices que apresentaram a melhor qualidade foram para 2 clusters.

Efetivamente, **pode-se considerar que a segunda hipótese da tese foi confirmada** já que o índice de silhouette atingiu um número satisfatório e também foi possível analisar o resultado obtido com critérios objetivos que condizem com a realidade. Entretanto, é importante ressaltar que, diferentemente da primeira hipótese, o resultado não tem efeito estatístico e tampouco é possível indicar um intervalo de confiança já que o número de deputados estudados é muito pequeno.

Não foi possível afirmar que existe um padrão de financiamento que diferencia deputados ligados às corporações pentecostais em relação aos demais deputados já que diversos deputados que não se identificam com tais corporações também se encaixaram no cluster 1. Entretanto todos os deputados da IURD, AD-MM, AD-VEC e Maranata se encaixaram no cluster 1 tal como podemos visualizar nas tabelas 18 a 21.

Como o posicionamento político foi uma das variáveis independentes e o resultado final do processamento se correlacionou com esta mesma variável, houve a preocupação de realizar um novo processamento, removendo todas as variáveis relativas ao partido político (posição ideológica, número de deputados do partido e número de filiados) para verificar se o resultado da clusterização se alterava. O novo processamento não alterou significativamente o resultado encontrado, trazendo apenas um número maior de outliers e um índice *silhouette* de qualidade inferior.

Por fim, outro ponto interessante diz respeito à descoberta de outliers já que foi possível identificar deputados evangélicos que tem um padrão distinto dos demais e que contam com um padrão de financiamento com um apoio acentuado por parte do empresariado.

Interessante notar que o único deputado de um partido de esquerda que entrou na lista de outliers (ALEX SANTANA PDT-BA) migrou para um partido de Direita no mesmo ano que assumiu o mandato, indo para o Republicanos.

Tabela 18 - Distribuição das Denominações na clusterização

Denominação	Cluster 0	Cluster 1	Outliers
AD	2	27	5
AD-MM		2	
AD-VEC		2	
Anglicana		1	
Batista	3	7	4
Brasil para Cristo		1	
IEQ	1	4	
IIGD	2	1	
IURD		22	
Luterana		1	2
Maranata		2	
Metodista		2	
Outras	2	15	
Presbiteriana	1	5	1
Sara Nossa Terra		1	
Total	11	93	12

Fonte: O autor

Tabela 19 - Distribuição das Denominações por Ideologia

Denominação	Cluster 0	Cluster 1	Outliers
Esquerda	3		1
Centro	5	9	1
Direta	3	84	10
Total	11	93	12

Fonte: O autor

Tabela 20 - Distribuição por Corporações Pentecostais

Denominação	Cluster 0	Cluster 1	Outliers
Corporações Pentecostais	5	62	5
Outros	6	31	7
Total	11	93	12

Fonte: O autor

Tabela 21 - Distribuição das Denominações por Partido

Denominação	Cluster 0	Cluster 1	Outliers
AVANTE	-	4	-
DEM	-	5	1
MDB	-	2	2
NOVO	-	-	2
PATRIOTA	-	1	-
PDT	1	-	1
PHS	3	-	-
PMN	-	1	-
PODE	-	4	-
PP	-	5	-
PR	-	10	-
PRB	-	23	-
PROS	2	-	1
PRP	-	1	-
PSB	3	-	-
PSC	-	4	1
PSD	-	6	2
PSDB	-	6	1
PSL	-	15	1
PT	2	-	-
PTB	-	2	-
PTC	-	1	-
SOLIDARIEDADE	-	3	-
Total	11	93	12

Fonte: O autor

7 CONCLUSÕES

Iniciei esta pesquisa em 2017 logo após o golpe parlamentar e antes das eleições de 2018. Apesar do crescimento do campo da Direita e da forte reação de vários setores conservadores frente aos diversos avanços sociais promovidos nos 20 anos anteriores, não podia imaginar que o retrocesso sobre os direitos humanos e sociais seria tão grande. As tentativas do Palácio do Planalto de intimidar o STF, o aparelhamento da Polícia Federal e da PGR e, principalmente, o aumento da miséria e a volta da inflação com dois dígitos mergulharam o país em um cenário tenebroso.

Em meio à crise política e econômica, a pandemia atingiu o ânimo e a disposição de seguir estudando um assunto que talvez não fosse mais relevante frente a todo sofrimento que assolava o mundo e, de forma ainda mais contundente, o Brasil. Seguir com a pesquisa, assistindo diariamente uma CPI que trazia esquemas de corrupção, negacionista e a desfaçatez de vários deputados que apareciam em várias etapas da pesquisa foi muito difícil. Mas, no final, o que me moveu a continuar foi a ideia de que a tese poderia ser um pequeno ato de resistência e, também, que poderia fazer alguma diferença para quem estuda o universo político brasileiro.

A pesquisa partiu da premissa que a análise do financiamento de campanha através do uso de técnicas de mineração de dados poderia trazer novos conhecimentos sobre o fenômeno eleitoral e, a partir de uma perspectiva marxista, seria possível relacionar quais frações de classe e setores econômicos financiavam os políticos evangélicos.

A tese se iniciou a partir da busca de referências sobre mineração de dados no campo da Ciência da Informação e das Ciências Sociais e Políticas. O escasso número de publicações que fizessem referência ao tema “Mineração de Dados” em ambas as áreas foi um ânimo extra para utilizar este tipo de metodologia para analisar os dados eleitorais e de financiamento de campanha.

Ainda no segundo capítulo, procurei trazer o debate sobre o uso de dados abertos (tema tão caro à Ciência da Informação) já que tal tema perpassa todo o levantamento e processamento dos dados. Também procurei mostrar que o ecossistema baseado em software livre, no qual modestamente milito há mais de 30

anos, é parte importante no processo do fazer científico e, ao elaborar toda a pesquisa utilizando apenas software livre, tive a intenção deliberada de construir uma “vitrine” para esta filosofia de trabalho. Busquei mostrar que não só é possível trabalhar com software livre, como este caminho pode trazer inúmeras vantagens para a pesquisa acadêmica.

Já no capítulo três, o levantamento bibliográfico da classificação das religiões abriu caminho para um entendimento que me fez ficar bem mais atento em relação ao preconceito contra os evangélicos e à urgência em “separar o joio do trigo”. A criação do termo “corporação pentecostal” veio com este intuito: de deixar bem clara a diferença entre a religiosidade de uma parcela enorme da população brasileira e os interesses financeiros de um grupo de indivíduos que encontrou na fé uma forma de auferir lucro.

A construção da base de dados com o histórico dos deputados evangélicos desde 1933 é efetivamente um capítulo à parte e se tornou uma pesquisa independente dentro da tese. Acredito que esta base poderá facilitar o trabalho de vários pesquisadores pois procurei sistematizar todo o processo, tornando bem clara a origem de cada registro. Além disso, procurei deixar a listagem o mais amigável possível, publicando o resultado tanto em formato legível por máquina como em formato de planilha eletrônica para que vários públicos pudessem fazer uso dos dados.

Já a base de financiamento é, de longe, a parte mais sensível da base de dados que compõe a pesquisa. Devido à falta de documentação por parte do TSE e pelas inconsistências encontradas na literatura, a base de dados pode ser de grande valia, mas também podem apresentar algumas divergências conceituais, uma vez que esta é a primeira vez que este tipo de consolidação foi realizado. O que eu procurei fazer foi minimizar tais divergências, deixando todos os passos registrados nos mínimos detalhes para que outros pesquisadores possam aprimorar ou corrigir os resultados obtidos.

A inovação de agrupar os doadores a partir dos setores econômicos foi uma aposta que deu certo pois foi possível ter uma ideia de como tais setores financiavam os deputados. A quantidade de pessoas físicas que doaram valores elevados, mas que não foram associadas a nenhum setor, pode ser um indicativo que este processamento pode ser melhorado utilizando bases do CNPJ antigas ou

então, através de outras técnicas que associem os doadores a um determinado setor. De qualquer forma, a base com o resultado agregado por setor pode constituir uma fonte de pesquisas futuras sobre o tema.

Para o problema em si que motivou a pesquisa, o resultado ultrapassou as expectativas do início da pesquisa. Através do método de regressão logística foi possível confirmar a primeira hipótese: os deputados federais evangélicos eleitos em 2018 tiveram um padrão de financiamento distinto dos demais. Apesar de utilizarmos um método estatístico “tradicional”, este só foi possível pois a base de dados que foi utilizada para o processo de mineração de dados já estava disponível. Desta forma, a construção dos modelos estatísticos foi facilitada em função da base de dados consolidada já existente.

A confirmação da segunda hipótese mostrou que existe um padrão de financiamento que diferencia os deputados evangélicos de centro-esquerda dos demais. Apesar do número de deputados evangélicos progressistas eleitos ter sido muito pequeno na legislatura de 2018, foi possível verificar que havia uma diferença marcante entre eles, pelo menos sob o ponto de vista de financiamento de campanha. Havia também uma expectativa de que fosse possível diferenciar os deputados ligados às corporações pentecostais ou então que outras classificações pudessem surgir do processamento mas os números não confirmaram esta expectativa.

A metodologia adotada abre caminho para novas pesquisas. Utilizando a mesma metodologia é possível encontrar padrões de financiamento, por exemplo, para os deputados da bancada católica, ou então para os deputados da bancada da segurança pública, comumente chamada pela imprensa de “bancada da bala”. A metodologia também pode ser empregada para os deputados estaduais ou vereadores evangélicos e também para os deputados federais anteriores a 2018. Especificamente para os deputados federais eleitos em 2014 e nos anos anteriores, anos em que o financiamento privado teve um peso maior sobre o total financiado, talvez a variância dos valores investidos para cada setor econômico possam gerar uma clusterização que traga agrupamentos mais interessantes.

Ainda sobre os anos anteriores, é importante ressaltar que o planejamento inicial tinha como meta a comparação dos padrões de financiamento entre as diversas eleições em anos diferentes, mas, ao longo do estudo da bibliografia,

percebi que a comparação seria muito complexa ou até mesmo equivocada pois as eleições seguiam comportamentos e regras diferentes, ou seja, não haveria um critério para que se pudesse comparar tais financiamentos, mesmo que se buscasse uma normalização dos dados.

Ao longo da pesquisa, surgiram ideias para o uso de novas variáveis que não estavam no planejamento inicial. Uma delas seria o percentual de doações intrapartidária, ou seja, ao invés de incluir o valor em reais doados pelos agentes partidários, a ideia seria utilizar o percentual ou o percentil de doações em relação aos candidatos dentro do mesmo partido. Dessa forma, seria possível verificar se determinados candidatos foram privilegiados em relação aos seus colegas de chapa.

A criação do índice de capilaridade pode trazer bons frutos não só para tentar identificar os deputados que se elegem em função dos chamados “currais eleitorais”, ou seja, localidades onde existe uma grande concentração de votos para um determinado candidato ou partido, como também para identificar se esse é um fator determinante para o sucesso eleitoral do candidato. Talvez, o índice possa trazer informações ainda mais relevantes se for aplicado para os casos de deputados estaduais ou para os vereadores.

Uma outra frente que pode ser estudada é a possibilidade de análise preditiva a partir dos clusters gerados. Talvez seja possível encontrar uma correlação entre os deputados agrupados em cada cluster e as votações no Congresso sobre um determinado setor econômico.

Considero que os três objetivos da pesquisa foram alcançados: a base de dados com os dados de financiamento de campanha foi criada e está disponível para o reuso; o uso de técnicas de mineração de dados auxiliaram a “visualização” dos dados eleitorais sob uma nova perspectiva e a metodologia apresentada para a classificação de agentes políticos em função do tipo de financiamento de campanha abre caminho para novas pesquisas tanto na área da ciência política como também para a CI, que tem entre um dos seus pilares os sistemas de classificação e organização do conhecimento.

Ao longo de todo o percurso investigativo, houve um trabalho extra considerável para montar um ambiente computacional que permitisse que outros pesquisadores pudessem reproduzir exatamente os dados e gráficos aqui apresentados. Principalmente, na seção de financiamento de campanha foi

necessário revisar as importações realizadas, pois diversas vezes os valores eram incompatíveis com vários outros artigos científicos que apresentavam números distintos dos que eu havia encontrado.

Com 30 anos de desenvolvimento de sistemas de informação, não tenho a pretensão de ter certeza de que os dados agregados do financiamento de campanha aqui apresentados estão completos e sintetizam a complexidade deste domínio, ainda mais porque a documentação do TSE não foi muito útil e das duas vezes que tentei o suporte do tribunal, não obtive nenhuma resposta. Na verdade, se algum erro for encontrado, ficarei muito feliz pois tal fato indicaria que algum outro pesquisador utilizou o ambiente computacional que eu preparei com afinco.

A certeza maior que se buscou foi a de que outros pesquisadores conseguirão reproduzir os resultados aqui apresentados, sem nenhum tipo de bloqueio aos dados ou ao código fonte utilizado. Mais do que esperar uma grande descoberta sobre a bancada evangélica, o objetivo maior da pesquisa sempre foi construir um método baseado em dados e em código que pudesse servir de orientação para outros pesquisadores da área das ciências políticas. Sempre houve também uma vontade de demonstrar como a Ciência da Informação pode ser um suporte para a análise de dados de pesquisa das mais diversas áreas de interesse.

Tal como a Matemática é um suporte para todas as ciências (sem deixar de ser um campo do conhecimento próprio), a Ciência da Informação também pode ter essa função: de oferecer alternativas metodológicas para as outras ciências, oferecendo caminhos na gestão de dados, na organização do conhecimento e, principalmente, fazer uso de uma de suas características mais marcante: o pensamento crítico sobre como os dados foram construídos, quais as motivações por trás desta construção e quais as implicações éticas e políticas sobre seus usos.

8 REFERÊNCIAS BIBLIOGRÁFICAS

- ADAMS, C. J. **Classification of religions**. Encyclopedia Britannica. Disponível em: <<https://www.britannica.com/topic/classification-of-religions>>. 2018. Acesso em: 31 jan. 2019.
- AHLQUIST, J. S.; BREUNIG, C. Model-based Clustering and Typologies in the Social Sciences. **Political Analysis**, v. 20, n. 01, p. 92–112, 2012.
- ALMEIDA, A. J. S. **“Pelo Senhor, marchamos”**: os evangélicos e a ditadura militar no Brasil (1964-1985). 2016. Tese (Doutorado em História). UFF, Niterói, RJ.
- ALMEIDA, R. D. Bolsonaro presidente: conservadorismo, evangelismo e a crise brasileira. **Novos estudos CEBRAP**, v. 38, n. 1, p. 185–213, abr. 2019.
- ALTMANN, W. Censo IBGE 2010 e Religião (IBGE 2010 Census and Religion). **Horizonte**, v. 10, n. 28, p. 1122–1129. <http://doi.org/10.5752/P.2175-5841.2012v10n28p1122>
- AMMERMAN, Nancy T. **Pillars of faith: American congregations and their partners**. Berkeley: University of California Press, 2005.
- AQUINO, J. A. DE. Software livre e desenvolvimento de trabalhos científicos: o R como exemplo a ser seguido. **Revista Política Hoje**, v. 24, n. 2, p. 75–86, 27 dez. 2015.
- ASSMANN, H. **A Igreja Eletrônica e Seu Impacto na América Latina**. Petrópolis: Editora Vozes, 1986.
- BAKAN, J. **The corporation: the pathological pursuit of profit and power**. Free Press. New York, NY. 2005.
- BANDEIRA DE MELLO, R.; MARCON, R.; ALBERTON, A. Drivers of discretionary firm donations in Brazil. **BAR - Brazilian Administration Review**, v. 5, p. 275–288, dez. 2008.
- BAPTISTA, S. Os evangélicos e o processo republicano brasileiro. **Revista Brasileira de História das Religiões**, v. 13, n. 39, 2021.
- BAPTISTA, S. de T. C. **Cultura Política Brasileira, Práticas Pentecostais e Neopentecostais**: A presença da Assembléia de Deus e da Igreja Universal do Reino de Deus no Congresso Nacional (1999-2006). 2007. Tese (Doutorado em Ciências da Religião). Universidade Metodista de Sao Paulo, São Bernardo do Campo.
- BASTOS, P. P. Z. Ascensão e crise do governo Dilma Rousseff e o golpe de 2016: poder estrutural, contradição e ideologia. **Rev. Econ. Contemp.** Rio de Janeiro, v. 21, n. 2, ago. 2017. <http://dx.doi.org/10.1590/198055272129>
- BATISTA, M. T. **Adaptações da fé: análise antropológica da transnacionalização da Igreja Universal entre Brasil e Argentina**. Doutorado em Antropologia—Porto Alegre: UFRS, 2013.
- BAUER, J. A. **Politique et religion**. Paris: PUFF, 1999.
- BELO, C. D. A. M. O uso de candidaturas laranjas dentro do contexto do financiamento

público de campanhas no estado do Acre: um estudo de caso. **Revista Brasileira de Ciências Policiais**, v. 12, n. 5, p. 207, 31 maio 2021.
<http://dx.doi.org/10.31412/rbcp.v12i5.706>

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web. **Scientific American**, v. 284, n. 5, p. 34–43, 2001.

BEZERRA, A. C.; SCHNEIDER, M.; SALDANHA, G. **Ascensão e queda da utopia tecnoliberal: a dialética da liberdade sociotécnica**. In: Encontro Nacional de Pesquisa em Ciência da Informação, 14., Florianópolis, 2013.

BOHN, Simone R. **Evangélicos no Brasil: perfil socioeconômico, afinidades ideológicas e determinantes do comportamento eleitoral**. Opinião Pública 10.2 (2004): 288-338.

BOITO JR., A. Estado e burguesia no capitalismo neoliberal. **Revista de Sociologia e Política**, p. 57–73, jun. 2007.

BOLOGNESI, B. et al. Como os partidos distribuem o dinheiro. Estrutura organizacional e recursos eleitorais em 2014 no Brasil. **Colombia Internacional**, n. 104, p. 33–62, out. 2020.
<https://doi.org/10.7440/colombiaint104.2020.02>

BOLOGNESI, B.; BABIRESKI, F. R.; MACIEL, A. P. Conhecendo o vazio: congruência ideológica e partidos políticos no Brasil. **Política & Sociedade**, v. 18, n. 42, p. 86–116, 16 dez. 2019.

BORGMAN, C. L. **Big data, little data, no data: scholarship in the networked world**. MIT Press. Cambridge. 2016.

BOURDIEU, P. **O poder simbólico**. Rio de Janeiro: Ed. Bertrand Brasil, 1989.

BOWKER, C.G; STAR, S.L. **Sorting Things Out: Classification and its consequences**. Cambridge: MIT Press, 2000.

BRAGA, C.F. TUZZO, S. A. Dados abertos à brasileira: aspectos de uma cidadania denegada. **Comunicação & Inovação**. PPGCOM/USCS v. 18, n. 37 (48-65) maio-ago 2017

BRAMAN, Sandra. **The emergent global information policy regime**. Basingstoke, Hampshire; New York: Palgrave Macmillan, 2004.

BRÄSCHER, Maria. Semantic Relations in Knowledge Organization Systems. **Knowledge Organization** v. 41, n. 412, 2014.

BRASIL. PROGRAMA DE GOVERNO ELETRÔNICO. Programa de governo, surgido através do Decreto Presidencial de 3 de abril de 2000. Disponível em:
 <<https://www.governodigital.gov.br/EGD/historico-1/historico>>. Acesso em: 10 jul. 2019.

BRISOLA, Anna Cristina; GOMES, J. C.; SCHNEIDER, M. A. F. **Hackeando dados abertos: competência crítica em informação e cidadania**. XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (XVIII ENANCIB). **Anais...** In: XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (XVIII ENANCIB). 24 set. 2017.

- BROUGHTON, Vanda. A new classification for the literature of religion. **A new classification for the literature of religion**, v. 29, n. 4, p. 59–61, 2000.
- BUCKLAND, M. K. Information as Thing. **Journal of the American Society of Information Science** v.42, n.5, p.351-360. 1991.
- BURLE, C. et al. **Os degraus da implementação efetiva no Brasil: como as regulamentações locais de acesso à informação impactam na implementação de portais de dados abertos e transparência**. ConDatos. Anais...Santiago, Chile: 2015.
- CAMPOS, L. S. **Os “políticos de Cristo” – uma análise do comportamento político de protestantes históricos e pentecostais no Brasil**. GT Religião e Sociedade. Anais... In: XXVI ANPOCS. Caxambu: 2002
- CAMPOS, M. M. **Democracia, partidos e eleições: os custos do sistema partidário-eleitoral no Brasil**. 2009. Tese (Doutorado em Ciência Política) - Faculdade de Filosofia e Ciências Humanas, Universidade Federal de Minas Gerais, Belo Horizonte.
- CAMURÇA, M. A religião e o censo: enfoques metodológicos uma reflexão a partir das consultorias do ISER ao IBGE sobre o dado religioso nos censos. In: CUNHA, C.; MENEZES, R. DE C. (Org.). **Religiões em conexão: números, direitos, pessoas**. COMUNICAÇÕES DO ISER. 69. ISER, 2014.
- CAPURRO, R.; HJØRLAND, B. O conceito de informação. **Perspectivas em Ciência da Informação**. Belo Horizonte, v.12, n.1, p.148-207, 2007.
- CAPURRO, R. **Ethics and Information in the Digital Age**. Annual Course and Conference: Libraries in the Digital Age, Dubrovnik, Croatia, 23-27 May, 2001. Disponível em: <<http://www.capurro.de/lida.htm>> Acesso em 02 jun 2016.
- _____. Knowledge map of Information Science. **Rafael Capurro's responses to Chaim Zins**. Disponível em: <<http://www.capurro.de/zins.html>>. Acesso em 02 jun 2017.
- CARLAN, E.; BRÄSCHER, M. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **Revista Ibero-Americana de Ciência da Informação**, v. 4, n. 2, p. 53–73, 31 dez. 1969.
- CARVALHO, A.C.F. Interdisciplinaridade da Ciência de Dados. **Computação Brasil**. v.31, p.62-65. 2016.
- CARVALHO, M. B. DE; TSUNODA, D. F. Análise de dados em artigos recuperados da Web of Science (WoS). **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 0, n. 0, p. 112–125, 20 jun. 2018.
- CHAVES, M. **American religion: contemporary trends**. Princeton, N.J: Princeton University Press, 2011.
- CHAVES, M. All Creatures Great and Small: Megachurches in Context. **Review of Religious Research**, v. 47, n. 4, p. 329–346, 2006.
- CHEN, S.Y.; LIU, X. **The contribution of data mining to information science**. Journal of Information Science December 2004 30: 550-558.

CODATO, A.; BERLATTO, F.; BOLOGNESI, B. Tipologia dos políticos de direita no Brasil: uma classificação empírica. **Análise Social**, v. LIII, n. 229, p. 870–897, 15 dez. 2018.

COENEN, F. Data mining: past, present and future. **The Knowledge Engineering Review**, v. 26, n. 1, p. 25–29, 2011.

COSTA, D.; MIRANDA, M. L. C. DE. A Organização do Conhecimento sobre Umbanda e sua representação bibliográfica: uma análise exploratória a partir de registros bibliográficos. **Informação & Informação**, v. 24, n. 3, p. 154, 31 dez. 2019.

CUNHA, Magali do Nascimento. **Do Púlpito às Mídias Sociais. Evangélicos na política e ativismo digital**. Rio de Janeiro: Appris Editora, 2019. 244 p. ISBN: 978-85-5507-717-3

CUNHA, Christina Vital da; LOPES, Paulo Victor Leite.; LUI, Janayna. **Religião e Política: medos sociais, extremismo religioso e as eleições 2014**. Rio de Janeiro: Fundação Henrique Böll/Instituto de Estudos da Religião, 2017. 196 p. ISBN: 978-85-62669-21-7.

DA SILVA, H. **Da moral privada à ética pública: A nova elite parlamentar evangélica na América Latina**. Tese (Doutorado em Ciência Política). Universidad de Salamanca, 2017.

DAHLBERG, Ingetraut. Teoria do conceito. **Ciência da Informação**. v. 7, n. 2, p. 101-07, 1978.

DANTAS, M. The Financial Logic of Internet Platforms: The Turnover Time of Money at the Limit of Zero. **tripleC: Communication, Capitalism & Critique**. v. 17, n. 1, p. 132–158, 18 maio 2019. <https://doi.org/10.31269/triplec.v17i1.1088>

DEBORD, G. **A sociedade do espetáculo**. [s.l.] E-Book. Projeto Periferia, 2003.

DEWEY, M. Dewey decimal classification and relative index. 22nd ed. Dublin: OCLC. 2003.

DRAPER, N. SMITH, H. **Applied Regression Analysis**. 2d Edition, New York: John Wiley & Sons, Inc. 1981.

FAJARDO, A. **A atuação dos evangélicos no rádio brasileiro: origem e expansão**. Dissertação de Mestrado. São Bernardo do Campo: Universidade Metodista de São Paulo, 2011.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. **AI Magazine**, v. 17, n. 3, p. 37–37, 15 mar. 1996.

FERNANDES, A. A. T. et al. Leia este artigo se você quiser aprender regressão logística. **Revista de Sociologia e Política**, v. 28, 8 jan. 2021. <https://doi.org/10.1590/1678-987320287406en>

FILHO, D. B. F. et al. Cluster Analysis for Political Scientists. **Applied Mathematics**, v. 5, n. 15, p. 2408–2415, 5 ago. 2014.

FITZGERALD, T. The Ideology of Religious Studies Revisited: The Problem with Politics. In: **Method and Theory in the Study of Religion: Working Papers from Hannover**. BRILL, 2017. p. 124–152.

FONSECA, A. B. **Pluralismo Religioso e Relação Religião-Estado: Uma Análise da Presença Evangélica No Legislativo Federal**. Em: ANPOCS XXVIII. Londrina, PR: 2004. Disponível em: <<https://www.anpocs.com/index.php/encontros/papers/28-encontro-anual-da-anpocs/st-5/st25-1/4091-afonseca-pluralismo/file>>

FONSECA, A. B. Religion and Democracy in Brazil: A Study of the Leading Evangelical Politicians. In: FRESTON, P. (Ed.). **Evangelical Christianity and Democracy in Latin America**. [s.l.] Oxford University Press, 2008. p. 163–204.

FRESTON, P. Bolsonaro, o populismo, os evangélicos e América Latina. In: GUADALUPE, J. L. P.; CARRANZA, B. (Eds.). **Novo ativismo político no Brasil: os evangélicos do século XXI**. Rio de Janeiro: Konrad Adenauer Stiftung, 2020.

_____. The protestant eruption into modern Brazilian politics. **Journal of Contemporary Religion**, v. 11, n. 2, p. 147–168, maio 1996.

_____. Protestantismo e política no Brasil: da constituinte ao impeachment. 1993. 307f. **Tese (doutorado)** - Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas, Campinas, SP.

FUNDAÇÃO FHC. **Os Evangélicos na Sociedade e na Política: efeitos e significados de uma influência crescente**. Disponível em: <<https://fundacaofhc.org.br/iniciativas/debates/os-evangelicos-na-sociedade-e-na-politica-efeitos-e-significados-de-uma-influencia-crescente>>. Acesso em: 24 jan. 2021

FURNER, Jonathan. “Data”: The data. In: KELLY, Matthew; BIELBY, Jared (Orgs.). **Information Cultures in the Digital Age: A Festschrift in Honor of Rafael Capurro**. Wiesbaden: Springer Fachmedien Wiesbaden, 2016, p. 287–306. https://doi.org/10.1007/978-3-658-14681-8_17

GEERTZ, C. O beliscão do destino: a religião como experiência, sentido, identidade e poder. In: **Nova luz sobre a antropologia**. Rio de Janeiro: Zahar, 2001.

GOMES, J. C.; SCHNEIDER, M.; BEZERRA, A. C. Aplicativos cívicos: apropriação de dados abertos governamentais pela sociedade. In: **Encontro Nacional de Pesquisa em Ciência da Informação**, 18. Londrina, 2018.

GONÇALVES, R. B. **“Bancada evangélica?”: uma análise do discurso parlamentar evangélico durante a 52a Legislatura da Câmara Federal**. Dissertação de Mestrado—Pelotas: Universidade Federal de Pelotas, 2011.

GONÇALVES, R. Grupos econômicos: uma análise conceitual e teórica. **Revista Brasileira de Economia**. v. 4, p. 491-518. 1991. <http://lemate.paginas.ufsc.br/files/2019/04/534-19325-1-PB.pdf>

GREEN, R. **A New View of Religion**. **The Dewey blog**, 2006. Disponível em: <https://ddc.typepad.com/025431/2006/07/a_new_view_of_r.html>. Acesso em: 24 jan. 2021

HAN, J.; KAMBER, M. **Data mining: concepts and techniques**. San Francisco: Morgan Kaufmann Publishers, 2001.

HAN, J., KAMBER, M., PEI, J. **Data Mining: Concepts and Techniques**. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 3rd edition. 2012.

HAND, D. J.; MANNILA H.; SMYTH P. **Principles of Data Mining (Adaptive Computation and Machine Learning)**. Cambridge, MA: MIT Press, 2001.

HJØRLAND, B. Fundamentals of Knowledge Organization. **Knowledge Organization**, v. 30, p. 87–111, 1 jan. 2003.

_____. Data (with big data and database semantics). **Knowledge Organization** 45, no. 8: 685-708. ISKO Encyclopedia of Knowledge Organization, ed. Birger Hjørland, coed. Claudio Gnoli, <http://www.isko.org/cyclo/data>. 2018.

HOROCHOVSKI, R. R. et al. Estruturas de poder nas redes de financiamento político nas eleições de 2010 no Brasil. **Opinião Pública**, v. 22, n. 1, p. 28–55, abr. 2016. <http://dx.doi.org/10.1590/1807-0191201622128>

IBGE. Censo Demográfico 2010 - **Características gerais da população, religião e pessoas com deficiência**. Rio de Janeiro: IBGE, 2012. Disponível em <http://biblioteca.ibge.gov.br/visualizacao/periodicos/94/cd_2010_religiao_deficiencia.pdf>. Acesso em: 01 jun. 2016.

ICIE - INTERNATIONAL CENTER FOR INFORMATION ETHICS. **The field**. <https://www.i-c-i-e.org/copy-of-information-ethics>. Acesso em: 01 mar. 2021.

IDREES, H.; MAHMOOD, K. Devising a Classification Scheme for Islam: Opinions of LIS and Islamic Studies Scholars. **Library Philosophy and Practice (e-journal)**, 5 nov. 2009.

ISOTANI, Seiji; BITTENCOURT, Ig Ibert. Dados abertos conectados. São Paulo: Novatec, 2015. 175 p.

JAIN, A. K. Data clustering: 50 years beyond K-means. **Pattern Recognition Letters**, v. 31, n. 8, p. 651–666, jun. 2010.

JAPPE, A. **Guy Debord**. Lisboa: Ed. Antígona, 2008.

JUNCKES, I. J. et al. Poder e democracia: uma análise da rede de financiamento eleitoral em 2014 no Brasil. **Revista Brasileira de Ciências Sociais**, v. 34, n. 100, 2019.

KASAHARA, Y. **Estado, empresariado financeiro e representação de interesses: uma análise das reformas dos mercados bancário e acionário brasileiros**. Tese de Doutorado—Rio de Janeiro: IUPERJ, 2009.

KOURIS, I. N.; MAKRIS, C. H.; TSAKALIDIS, A. K. Using Information Retrieval techniques for supporting data mining. **Data & Knowledge Engineering**, v. 52, n. 3, p. 353–383, 1 mar. 2005.

KRAUSE, S.; MACHADO, C.; MIGUEL L.F. **Coligações e Disputas Eleitorais na Nova República: Aportes teórico-metodológicos, tendências e estudos de caso**. Rio de Janeiro, RJ: UNESP, 2017.

KRAUSE, S.; REBELLO, M. M.; DA SILVA, J. G. O perfil do financiamento dos partidos brasileiros (2006-2012): o que as tipologias dizem? **Revista Brasileira de Ciência Política**,

n. 16, p. 247–272, abr. 2015.

KRUSE, K. M. **One nation under God: how corporate America invented Christian America**. New York, NY: Basic Books, 2016.

LACERDA, F. Assessing the Strength of Pentecostal Churches' Electoral Support: Evidence from Brazil. **Journal of Politics in Latin America**, v. 10, n. 2, p. 3–40, 22 ago. 2018.

_____. Evangelicals, pentecostals and political representation in brazilian legislative elections (1998-2010). **Revista Brasileira de Ciências Sociais**, v. 32, n. 93, 2017.

LENIN, Vladimir I. **Cadernos Sobre a Dialética de Hegel**. Rio de Janeiro: Editora UFRJ, 2011.

LESKOVEC, J., RAJARAMAN, A., ULLMAN, J. D. **Mining of massive datasets**. Cambridge University Press, 2014.

LIEBMAN, R. C.; SUTTON, J. R.; WUTHNOW, R. Exploring the Social Sources of Denominationalism: Schisms in American Protestant Denominations, 1890-1980. **American Sociological Review**, v. 53, n. 3, p. 343, jun. 1988.

LIU, F. T.; TING, K. M.; ZHOU, Z.-H. **Isolation Forest**. 2008 Eighth IEEE International Conference on Data Mining. **Anais...** In: 2008 Eighth IEEE International Conference on Data Mining. Dez. 2008.

MACHADO, C. Coligações Revisitadas: metodologias e encruzilhadas. In: KRAUSE, S.; MACHADO, CARLOS; MIGUEL, L. F. (Eds.). **Coligações e Disputas Eleitorais na Nova República: Aportes Teórico-Metodológicos e estudos de caso**. São Paulo: Fundação Konrad Adenauer, 2017.

MAFRA, Carla. O que os homens e as mulheres podem fazer com números que fazem coisas. In: TEIXEIRA, F.; MENEZES, R. (Eds.). **As religiões no Censo de 2010. Petrópolis: Vozes, 2013**. [s.l.] Editora Vozes, 2013.

MAFRA, Carla. **Os evangélicos**. Editora Zahar, 2001.

MAI, J.-E. Ethics, Values and Morality in Contemporary Library Classifications. **Knowledge organization**, v. 40, n. 4, p. 242–253, 2013.

MANCUSO, W. P. Investimento eleitoral no Brasil: balanço da literatura (2001–2012) e agenda de pesquisa. **Revista de Sociologia e Política**, v. 23, n. 54, p. 155–183, jun. 2015.

MANCUSO, W. P.; HOROCHOVSKI, R. R.; CAMARGO, N. F. Empresários e financiamento de campanhas na eleição presidencial brasileira de 2014. **Teoria e Pesquisa**, v. 25, n. 3, p. 38–64, 2016.

MANCUSO, W. P. ; SPECK, B. W. Financiamento empresarial na eleição para deputado federal (2002-2010): determinantes e consequências. **Teoria & Sociedade (UFMG)**, v. 23, p. 103-125, 2015.

MANNING, C. D. Computational Linguistics and Deep Learning. **Computational Linguistics**, v. 41, n. 4, p. 701–707, dez. 2015.

MARIANO, R. Expansão pentecostal no Brasil: o caso da Igreja Universal. **Estudos Avançados**, v. 18, n. 52, p. 121–138, dez. 2004.

_____. Análise Sociológica do Crescimento Pentecostal no Brasil. São Paulo: **Tese de doutorado em Sociologia**. Universidade de São Paulo, 2001.

_____. Os neopentecostais e a teologia da prosperidade. **Novos estudos**, nº 44, p. 24-44, 1996.

MARIANO, R.; HOFF, M.; DANTAS, T. Y. DE S. Evangélicos sanguessugas, presidenciáveis e candidatos gaúchos: a disputa pelo voto dos grupos religiosos. **Debates do NER**, v. 2, n. 10, 21 dez. 2007.

MARIANO, R.; PIERUCCI, A. F. O envolvimento dos pentecostais na eleição de Collor. **Novos estudos CEBRAP**, v. 3, n. 34, p. 92–106, 1992.

MARIN, J. R. A Assembleia de Deus nos anos 1990: a “Década da Colheita” (Assembly of God in the 1990s: the “Decade of Harvest”). **HORIZONTE**, v. 12, n. 34, p. 436–464, 28 jun. 2014. <https://doi.org/10.5752/P.2175-5841.2014v12n34p436>

MARTELETO, R.; NÓBREGA, N.; MORADO, D. Cultura informacional: Demarcações de uma linha de estudos de cultura, informação e sociedade. In: ALBAGLI, S. (Ed.). **Fronteiras da Ciência da Informação**. Brasília: IBICT, 2013. p. 80–108.

MARX, Karl. **O 18 de brumário de Luís Bonaparte**. São Paulo: Boitempo, 2011.

MATOS, A. S. DE. Breve história do Protestantismo no Brasil. **Vox Faifae: Revista de Teologia da Faculdade FASSEB**, v. 3, n. 1, 2011.

MCILWAINE, I.; MITCHELL, J. S. **The new ecumenism: Exploration of a DDC/UDC view of religion**. Knowledge organization for a global learning society. **Anais...** In: Proceedings of the ninth international ISKO conference. Vienna, Austria: Würzburg: Ergon Verlag, 2006

MCINNES, L.; HEALY, J.; ASTELS, S. hdbscan: Hierarchical density based clustering. **The Journal of Open Source Software**, v. 2, n. 11, p. 205, 2017.

MEDEIROS, J. D. S.; CAREGNATO, S. E. Compartilhamento de dados e e-Science: explorando um novo conceito para a comunicação científica. **Liinc em Revista**, v. 8, n. 2, 18 dez. 2012.

MICHALAKIDIS, G. Appreciation of structured and unstructured content to aid decision making - from web scraping to ontologies and data dictionaries in healthcare. **Tese de Doutorado**. University of Surrey. 2016.

MIRANDA, M. L. C. DE; SILVA, F. G. DA. Religião e cultura periféricas: a representação do islamismo na Classificação Decimal de Dewey. **Logeion: Filosofia da Informação**, v. 5, n. 2, p. 86–120, 20 mar. 2019.

MIRANDA, M. L. C. de. A organização do etnoconhecimento: a representação do conhecimento afrodescendente em Religião na CDD. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 8., 2007, Salvador. Anais... Salvador: UFBA, 2007. 12 p. 2018.

NHACUONGUE, J. A. **O campo da Ciência da Informação: contribuições, desafios e perspectivas da mineração de dados para o conhecimento pós-moderno**. 194 f. Tese (Doutorado) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências, 2015.

NISO. **Understanding Metadata**. [s.l.] NISO Press, 2004.

NUNES, B. P.; LOPES, G. R.; CASANOVA, M. A. **Automatic Classification and Taxonomy Generation for Semi-structured Data**. In: 2015 IEEE international conference on computer and information technology; ubiquitous computing and communications; dependable, autonomic and secure computing; pervasive intelligence and computing. 2015

OLIVEIRA, Cilas Ferraz de. Nunca, na história deste país... A contribuição de Guaracy Silveira ao metodismo do Brasil. 2008. **Tese (Doutorado em Educação)**. Universidade Metodista de Piracicaba. Piracicaba, 2008, 161p.

OLIVEIRA, T. B. DE. O golpe de 2016: Breve ensaio de história imediata sobre democracia e autoritarismo. **Historiæ**, v. 7, n. 2, p. 191–232, 2016.

ORO, A. P. A laicidade no Brasil e no Ocidente. Algumas considerações. **Civitas - Revista de Ciências Sociais**, v. 11, n. 2, 16 set. 2011. <https://doi.org/10.15448/1984-7289.2011.2.9646>

PEDREGOSA, F. et al. Scikit-learn: Machine Learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PEIXOTO, V. M. **Eleições Financiamento de Campanhas no Brasil**. Editora Garamond. 2016.

Pew Research Center (PRC). **A Brief History of Religion and the U.S. Census. Pew Research Center's Religion & Public Life Project (blog)**. 2010. <<https://www.pewforum.org/2010/01/26/a-brief-history-of-religion-and-the-u-s-census/>>. Acesso em 10 jan. 2017.

Pew Research Center (PRC). **Global Christianity: A Report on the Size and Distribution of the World's Christian Population**. 2011. Disponível em <<http://www.pewforum.org/2011/12/19/global-christianity-exec/>>. Acesso em 05 ago. 2018.

PIMENTA, R. M. Os Objetos Técnicos e seus papéis no horizonte das Humanidades Digitais: um caso para a Ciência da Informação. **Revista Conhecimento em Ação**, v. 1, n. 2, p. 33, 2016.

PINTO, André Moreira. **Mineração de textos e gestão do conhecimento: aplicação na experiência operacional em geração de energia nuclear nas Usinas de Angra I e II**. 369 f. Tese (Doutorado) –Escola de Ciências da Informação, Universidade Federal de Minas Gerais, 2007.

POMBO, Olga. **Da classificação dos seres à classificação dos saberes**. [S.l.: s.n.]. Disponível em:<<http://cfc.ul.pt/textos/OP%20-%20Da%20Classificacao%20dos%20Seres%20a%20Classificacao%20dos%20Saberes.pdf>>. Acesso em: 10 mai. 2016.

POMERANTZ, Jeffrey. **Metadata**. Cambridge, Massachusetts ; London, England: The MIT Press, 2015.

POWER, T. J.; RODRIGUES-SILVEIRA, R. **Replication Data for: Mapping Ideological Preferences in Brazilian Elections, 1994-2018: A Municipal-Level Study**. Harvard Dataverse. <https://doi.org/10.7910/DVN/5P03UL>

RANGANATHAN, S. R. **Colon Classification**. 6. ed. Delhi, India: Ess Ess Publications. 1960.

REICH, G.; DOS SANTOS, P. The Rise (and Frequent Fall) of Evangelical Politicians: Organization, Theology, and Church Politics. **Latin American Politics and Society**, v. 55, n. 4, p. 1–22, 2013.

RICHEY, R. E. **Denominationalism: illustrated and explained**. Eugene, Oregon: CASCADE Books, 2013.

RODRIGUES-SILVEIRA, R.; CERVI, E. U. Evangélicos e voto legislativo: Diversidade confessional e voto em deputados da bancada evangélica no Brasil. **Latin American Research Review**, v. 54, n. 3, p. 560–573, 17 set. 2019.

ROUSSEEUW, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. **Journal of Computational and Applied Mathematics**, v. 20, p. 53–65, nov. 1987.

SALES, L. F. ; SAYÃO, L. F. ; SOUZA, R. F. . **Publicações ampliadas: um novo modelo de publicação acadêmica para o ambiente de e-science**. In: XIV ENANCIB - Encontro Nacional de Pesquisa em Ciência da Informação, 2013, Florianópolis – SC.

SAMUELS, D. Money, Elections, and Democracy in Brazil. **Latin American Politics and Society**, v. 43, n. 2, p. 27–48, 2001.

SANCHIS, P. AS CIÊNCIAS SOCIAIS DA RELIGIÃO NO BRASIL. **Debates do NER**, v. 1, n. 11, 22 dez. 2007.

SANTOS, Bruno Carazza dos. Interesses econômicos, representação política e produção legislativa no Brasil sob a ótica do financiamento de campanhas eleitorais. 2016. **Tese**. Faculdade de Direito da Universidade Federal de Minas Gerais.

SANTOS, R. D. Segmentos econômicos e determinantes do financiamento político no Brasil*. **Colombia Internacional**, n. 101, 15 jan. 2020.

SARACEVIC, T. Interdisciplinary nature of information science. **Ciência da Informação**, v. 24, n.1, p. 36-41, 1995.

SEJNOWSKI, T. J. **The deep learning revolution**. Cambridge, Massachusetts: The MIT Press, 2018.

SEMELER, A. R.; PINTO, A. L. Os diferentes conceitos de dados de pesquisa na abordagem da biblioteconomia de dados. **Ciência da Informação**, v. 48, n. 1, 3 maio 2019.

SHMUELI, G. To Explain or to Predict? **Statistical Science**, v. 25, n. 3, p. 289–310, ago. 2010.

SIEPIRSKI, Paulo D. Pós-pentecostalismo e política no Brasil. **Estudos Teológicos** 37.1. p. 47-61. 1997.

SMITH, J. Z. A Matter of Class: Taxonomies of Religion. **The Harvard Theological Review**, v. 89, n. 4, p. 387–403, 1996.

SOTTANI, Silvânia M.R. **Biopolítica dos Afetos: Alteridades Ressentidas e a Circulação da Intolerância**. Tese (Doutorado em Comunicação e Cultura). Escola de Comunicação - Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2016, 221 p.

SOUSA, Bertone de Oliveira. A teologia da prosperidade e a redefinição do protestantismo brasileiro: uma abordagem à luz da análise do discurso. **Revista Brasileira de História das Religiões** 4, no 11. 2011. DOI: 10.4025/rbhranpuh.v4i11.30407

SOUZA, Renato Rocha; TUDHOPE, D.; ALMEIDA, M. B. Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems. **Knowledge Organization**, v. 39, n. 3, p. 179–192, 2012.

SOUZA, Renato Rocha; ALMEIDA, Maurício Barcellos; BARACHO, Renata Maria Abrantes. Ciência da Informação em transformação: Big Data, Nuvens, Redes Sociais e Web Semântica. **Ciência da Informação**, v. 42, n. 2, 2013.

SOUZA, R. R.; CAFÉ, L. M. A. Análise de sentimento aplicada ao estudo de letras de música. **Informação & Sociedade: Estudos**, v. 28, n. 3, 28 dez. 2018.

SPANG-HANSEN, H. How to Teach About Information as Related to Documentation? **Human IT: Journal for Information Technology Studies as a Human Science**, v. 5, n. 1, 2001.

STEENSLAND, B. et al. The Measure of American Religion: Toward Improving the State of the Art. **Social Forces**, v. 79, n. 1, p. 291, set. 2000.

STEINER, Tobias. **Metadaten und OER: Geschichte einer Beziehung. Synergie: Fachmagazin für Digitalisierung in der Lehre** v. 4, S. 51-55. 2017.

STUDWELL, W. E.; WU, H.; WANG, R. Ideological Influences on Book Classification Schemes in the People's Republic of China. **Cataloging & Classification Quarterly**, v. 19, n. 1, p. 61–74, 26 ago. 1994.

TAROUCO, G. D. S.; MADEIRA, R. M. Os partidos brasileiros segundo seus estudiosos: análise de um expert survey. **Civitas - Revista de Ciências Sociais**, v. 15, n. 1, p. 24, 27 abr. 2015. <https://doi.org/10.15448/1984-7289.2015.1.18077>

TEIXEIRA, F. L. C.; MENEZES, R. DE C. (EDS.). **Religiões em movimento: o censo de 2010**. Petrópolis: Editora Vozes, 2013.

THATCAMP. Manifesto das Humanidades Digitais. ThatCamp [The Humanities and Technology Camp]. Paris: 2011. Disponível em: <https://tcp.hypotheses.org/497>. Acesso em: 24 jun. 2017.

TREVISAN, J. B. Evangélicos pentecostais na política partidária brasileira: de 1989 a 2010. IV Encontro Nacional do GT história das religiões e das religiosidades – ANPUH - Memória e Narrativas nas Religiões e nas Religiosidades. Anais - **Revista Brasileira de História das Religiões**, Maringá/PR, v. V, n.15, jan. 2013.

TRYBULA, W. J. Text Mining. **Annual Review of Information Science and Technology (ARIST)**, v. 34, p. 385–419, 1999.

TSIEN, T.-H. A History of Bibliographic Classification in China. **The Library Quarterly**, v. 22, n. 4. 1958. <https://doi.org/10.1086/617916>

VAZ, J. C.; RIBEIRO, Manuella M.; MATHEUS, Ricardo. **Desafios para a Governança Eletrônica e Dados Governamentais Abertos em Governos Locais**. In: WTRANS13 - Workshop de Transparência em Sistemas. Brasília. 2013.

VIANA, B. Journalism in the context of the Semantic Web. **Media & Jornalismo**, v. 18, n. 32, p. 187–199, 2018.

WEBER, Max. **Ética protestante e o espírito do capitalismo**. São Paulo: Ed. Pioneira, 1999.

WILSON, R. The New Ecclesiology: Mega-Church, Denominational Church, and No Church. **Review & Expositor**, v. 107, n. 1, p. 61–72, fev. 2010.

WUTHNOW, R. **God and mammon in america**. New York: Free Press, 2014.

YANG, F. **Atlas of religion in China: social and geographical contexts**. Leiden Boston: Brill, 2018.

ZHANG, W. Classification for Chinese Libraries (CCL): Histories, Accomplishments, Problems and Its Comparisons. **Journal of Educational Media & Library Sciences**, v. 41, 2003.

9 OBJETOS DIGITAIS

Relação de Deputados Federais Evangélicos de 1933 a 2018:

<https://doi.org/10.5281/zenodo.5766213>

Espectro ideológico dos Partidos Políticos de 1986 a 2018:

<https://doi.org/10.5281/zenodo.5515647>

Rotinas de Clusterização:

https://colab.research.google.com/drive/1xPN4XqzhBFi67N_f1oEcm03_34VyoWg7?usp=sharing

Referências Bibliográficas no Zotero:

https://www.zotero.org/groups/4537759/tese_josir/library

10 ANEXO 1

Deputados federais eleitos em 2018 considerados como Outliers pela Clusterização

Nome do Deputado	UF	Partido	Posicionamento	Igreja
ONYX DORNELLES LORENZONI	RS	DEM	DIREITA	Luterana
LUCIO ANTONIO MOSQUINI	RO	MDB	DIREITA	Batista
GUTEMBERG REIS DE OLIVEIRA	RJ	MDB	DIREITA	AD
MARCEL VAN HATTEM	RS	NOVO	DIREITA	Luterana
LUCAS DE VASCONCELOS GONZALEZ	MG	NOVO	DIREITA	Batista
ALEX MARCO SANTANA SOUSA	BA	PDT	ESQUERDA	AD
ANTONIO WANDSCHEER	PR	PROS	CENTRO	AD
JOSÉ VALDEVAN DE JESUS SANTOS	SE	PSC	DIREITA	AD
HAROLDO ALVES CAMPOS	RR	PSD	DIREITA	Batista
FÁBIO SALUSTINO MESQUITA DE FARIA	RN	PSD	DIREITA	Batista
BRUNA DIAS FURLAN	SP	PSDB	DIREITA	AD
FILIPE BARROS BAPTISTA DE TOLEDO RIBEIRO	PR	PSL	DIREITA	Presbiteriana

Deputados Federais eleitos em 2018 no Cluster 1

Nome do Deputado	UF	Partido	Posicionamento	Igreja
ANTÔNIO SERGIO ALVES VIDIGAL	ES	PDT	ESQUERDA	Batista
JOSE ABILIO SILVA DE SANTANA	BA	PHS	DIREITA	AD
FERNANDO VASCONCELOS	PE	PHS	DIREITA	Outras
GELSON DE AZEVEDO ALMEIDA	RJ	PHS	DIREITA	Outras
CLARISSA GAROTINHO	RJ	PROS	CENTRO	Presbiteriana
WELITON FERNANDES PRADO	MG	PROS	CENTRO	Batista
JOÃO HENRIQUE HOLANDA CALDAS	AL	PSB	CENTRO	IIGD
JEFFERSON ALVES DE CAMPOS	SP	PSB	CENTRO	IEQ
LIZIANE BAYER DA COSTA	RS	PSB	CENTRO	IIGD
BENEDITA SOUZA DA SILVA	RJ	PT	ESQUERDA	AD
REJANE RIBEIRO SOUSA DIAS	PI	PT	ESQUERDA	Batista

11 GLOSSÁRIO

Linked Data conjuntos de dados na rede mundial que estão interligados por meio de RDF usando identificadores de recursos uniformes baseados na web (URIs) para identificar tanto os objetos descritos bem como os termos usados para descrevê-los.

General Public License (GPL)

Termo de designação, do Inglês, dado à Licença Pública Geral, criada pela *Free Software Foundation (FSF)*. Esta licença encerra em si um acordo legal de uso de software no qual o usuário, tem acesso garantido ao código fonte, sendo convidado a alterá-lo e assim contribuir para seu desenvolvimento. A principal garantia é de que esta licença tem propriedade “virótica”, ou seja, ao se integrar a outros códigos, todos obrigatoriamente devem passar a serem licenciados sob a GPL, e portanto o autor deve disponibilizar os códigos fonte de tudo.

Open Source É uma classificação que se dá a softwares que têm seu código fonte público. Difere do Freeware que é gratuito, mas o código não é disponível ao público. Já o Open Source, em geral é licenciado sob a *General Public License (GPL)*, ou outras licenças, que se caracterizam por deixar livre a sua utilização, alteração sem custos.