

Universidade Federal do Rio Grande do Sul
Faculdade de Biblioteconomia e Comunicação
Programa de Pós-Graduação em Comunicação e Informação

Lisiane Braga Ferreira

**ARQUIVAMENTO DA *WEB* E MÍDIAS SOCIAIS:
Preservação Digital de Vídeos da Campanha Presidencial Brasileira de 2018**

**Porto Alegre
2019**

Lisiane Braga Ferreira

**ARQUIVAMENTO DA *WEB* E MÍDIAS SOCIAIS:
Preservação Digital de Vídeos da Campanha Presidencial Brasileira de 2018**

Dissertação apresentada ao Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul como requisito parcial para obtenção do título de Mestre em Comunicação e Informação.

Orientador: Prof. Dr. Moisés Rockembach

**Porto Alegre
2019**

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. Rui Vicente Oppermann

Vice-Reitora: Prof.^a Dr.^a Jane Fraga Tutikian

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Diretora: Prof.^a Dr.^a Karla Maria Müller

Vice-Diretora: Prof.^a Dr.^a Ilza Maria Tourinho Girardi

PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E INFORMAÇÃO

Coordenador: Prof.^a Dr.^a Ana Tais Martins Portanova Barros

Coordenadora-substituta: Prof.^a Dr.^a Nísia Martins do Rosário

CIP - Catalogação na Publicação

Ferreira, Lisiane Braga
ARQUIVAMENTO DA WEB E MÍDIAS SOCIAIS: Preservação
Digital de Vídeos da Campanha Presidencial Brasileira
de 2018 / Lisiane Braga Ferreira. -- 2019.
107 f.
Orientador: Moisés Rockembach.

Dissertação (Mestrado) -- Universidade Federal do
Rio Grande do Sul, Faculdade de Biblioteconomia e
Comunicação, Programa de Pós-Graduação em Comunicação
e Informação, Porto Alegre, BR-RS, 2019.

1. Arquivamento da Web e mídias sociais. . 2.
Candidatos presidenciais de 2018. . 3. Eleições
presidenciais de 2018. . 4. Facebook.. 5. Vídeos.. I.
Rockembach, Moisés, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os
dados fornecidos pelo(a) autor(a).

PPGCOM/UFRGS

Rua Ramiro Barcelos, 2705

CEP 90035-007 – Porto Alegre, RS

Telefone: (51) 3308-5116

E-mail: ppgcom@ufrgs.br

Lisiane Braga Ferreira

**ARQUIVAMENTO DA WEB E MÍDIAS SOCIAIS:
Preservação Digital de Vídeos da Campanha Presidencial Brasileira de 2018**

Dissertação apresentada ao Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul como requisito parcial para obtenção do título de Mestre em Comunicação e Informação.

Aprovada em 07 de maio de 2019.

Banca Examinadora

Prof. Dr. Moisés Rockembach (Orientador)

Prof.^a Dr.^a Ana Maria Mielniczuk de Moura

Prof. Dr. Armando Malheiro da Silva

Prof.^a Dr.^a Caterina Marta Groposo Pavão

Prof. Dr. Valdir José Morigi (Suplente)

RESUMO

O arquivamento da *Web* e mídias sociais constituiu o tema central desta pesquisa, e objetivou, a partir da personalização dos candidatos à Presidência de 2018 na plataforma *Facebook*, preservar os conteúdos audiovisuais publicados durante a campanha eleitoral. Adotou a metodologia de revisão bibliográfica e análise documental, de desenvolvimento de códigos de programação e estruturação de base de dados, sob uma abordagem exploratória-descritiva de análise quanti-qualitativa. O referencial teórico versou, predominantemente, sobre o fenômeno da personalização dos políticos nas mídias sociais e sobre o arquivamento da *web* e mídias sociais. Como objetivo geral compreendeu a necessidade de preservação digital dos vídeos publicados pelos candidatos à Presidência no *Facebook*, por meio do arquivamento da *Web*. E como objetivos específicos, identificou todos os candidatos oficiais à Presidência, bem como suas respectivas páginas no *Facebook*; coletou todos os registros audiovisuais publicados pelos candidatos em suas páginas durante a campanha eleitoral do 1º e 2º turnos nos períodos: de 16 de agosto de 2018 a 04 de outubro de 2018 e de 08 de outubro de 2018 a 26 de outubro de 2018; verificou a disponibilidade do formato original dos registros publicados pelos candidatos após o período eleitoral; verificou os conteúdos de fontes de informações dos candidatos preservados em plataformas de arquivamento da *Web* e disponibilizou o protótipo com os registros coletados e preservados. Observou que, dos 2821 vídeos coletados durante a campanha eleitoral, 32 não estavam mais disponíveis no formato original ou por restrições legais ou por indisponibilidades diversas. Constatou que a iniciativa *Internet Archive* realizou a maior parte das coletas das fontes de informação dos candidatos, mas não de maneira intensiva e nem homogênea. Observou que a maior parte dos vídeos das fontes de informação dos candidatos não foram preservados pelas três iniciativas analisadas. Conclui que existe uma carência geral de arquivamento da *Web* e mídias sociais de conteúdos brasileiros, sendo os conteúdos preservados por iniciativas internacionais fragmentado e limitado e, ainda, que a preservação de publicações públicas em mídias sociais deve ser fomentada no Brasil a fim de minimizar a efemeridade inerente aos registros.

Palavras-chave: Arquivamento da *Web* e mídias sociais. Candidatos presidenciais de 2018. Eleições presidenciais de 2018. *Facebook*. Vídeos.

ABSTRACT

The main subject of this paper is the archiving of the Web and social media and its objective was to preserve, based on the personalization of the 2018 Presidential candidates on the Facebook platform, audiovisual content published during the electoral campaign. The methodology consisted of a literature review and documentary analysis and the development of programming codes and database structure, under an exploratory-descriptive approach of quanti-qualitative analysis. The theoretical background was, predominantly, on the phenomenon of personalization of politicians in social media and on the archiving of the Web and social media. As a general objective, it understood the need for digital preservation of videos published by Presidential candidates on Facebook through archiving the Web. And as a specific goal, it identified all the official presidential candidates as well as their respective Facebook pages; collected all the audiovisual records published by the candidates on their pages during the electoral campaign of the first and second rounds in the periods: from August 16, 2018 to October 4, 2018 and from October 8, 2018 to October 26, 2018; the availability of the original format of the collected records published by the candidates after the electoral period was verified ; verified the contents of source information of the candidates preserved in archival platforms of the Web and made available the prototype with the collected and preserved records. It observed that the 2821 videos collected during the electoral campaign, 32 were no longer available in the original format or by legal restrictions or other unavailability. It was found that most of the collection of the source information of the candidates was done by the Internet Archive initiative, but not in an extensive or homogenous way. It was also observed that most videos from the candidates' source information were not collected by the three initiatives that were analyzed. It was concluded that there is a general lack of Web and social media archiving of Brazilian content; that the content archived by international initiatives is incomplete and limited; and that the archiving of public publications on social media should be fostered in Brazil in order to decrease the inherent ephemerality of the records.

Keywords: *Web archiving and social media. Presidential Candidates 2018. Presidential Elections 2018. Facebook. Videos.*

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1 - Tela de transmissão ao vivo no dispositivo móvel e no computador. | 20 |
| Figura 2 - Representação do Fenômeno Infocomunicacional nos Lives..... | 22 |
| Figura 3 - Avatar dos candidatos à Presidência em suas fanpages..... | 43 |
| Figura 4 - Interface Web de inclusão dos registros audiovisuais na base de dados. | 44 |
| Figura 5 - Pastas dos candidatos com os registros audiovisuais no servidor. | 45 |
| Figura 6 - Interface Web de inclusão após coleta de vídeo, exemplo. | 46 |
| Figura 7 - Página de consulta com exemplo..... | 47 |
| Figura 8 - Telas de atualização e histórico. | 47 |
| Figura 9 - Exemplo de identificação de registro indisponível a partir do iframe..... | 58 |
| Figura 10 - Mensagens padrões de conteúdo indisponível na Página do Facebook. | 58 |
| Figura 11 - Comparativo de vídeo excluído e repostado com acréscimos. | 59 |
| Figura 12 - Snapshot de vídeo do Facebook preservado no Archive-It. | 61 |
| Figura 13 - Snapshot da página do candidato Fernando Haddad no Instagram preservado pelo Internet Archive..... | 66 |
| Figura 14 - Snapshot da página do candidato Ciro Gomes no Instagram preservado pelo Internet Archive..... | 66 |
| Figura 15 - Lista de escolha com os nomes dos candidatos às eleições presidenciais de 2018 e do ex-presidente Lula do Protótipo. | 68 |
| Figura 16 - Campo de período para consulta do Protótipo..... | 69 |
| Figura 17 - Seleção de tipo de conteúdo, Live ou vídeo. | 69 |
| Figura 18 - Seleção de disponível no formato original, ativo ou não..... | 69 |
| Figura 19 - Consulta pelo <i>link</i> da publicação conforme página do candidato..... | 69 |
| Figura 20 - Visualização do link da publicação extraída da página do candidato. | 70 |
| Figura 21 - Apresentação de consulta por candidato..... | 70 |
| Figura 22 - Execução do link armazenado. | 71 |
| Figura 23 - Lista de histórico de interações da publicação. | 71 |
| Figura 24 - Visualização do iframe da publicação na origem. | 72 |

LISTA DE GRÁFICOS E TABELAS

| | |
|--|----|
| Gráfico 1 - Comparativo dos Lives com maior interação do candidato Jair Bolsonaro..... | 48 |
| Gráfico 2 - Quantitativo de registros audiovisuais publicados e coletados no 1º turno das eleições presidenciais de 2018..... | 49 |
| Gráfico 3 - Quantitativo de vídeos publicados no Facebook no 1º turno por candidato..... | 50 |
| Gráfico 4 - Quantitativo de Lives publicados no Facebook no 1º turno por candidato..... | 51 |
| Gráfico 5 - Quantitativo de vídeos publicados no Facebook no 2º turno por candidato..... | 52 |
| Gráfico 6 - Quantitativo de Lives publicados no Facebook no 2º turno por candidato..... | 52 |
| Gráfico 7 - Comparativo de publicações no Facebook do 1º e do 2º turno..... | 53 |
| Gráfico 8 - Quantitativo diário de publicações no Facebook no 1º turno..... | 53 |
| Gráfico 9 - Comparativo de publicações no Facebook no 1º turno dos dois principais candidatos..... | 54 |
| Gráfico 10 - Comparativo de publicações no Facebook no 2º turno dos dois principais candidatos..... | 54 |
| Gráfico 11 - Análise de disponibilidade do formato original das publicações dos candidatos no Facebook - novembro..... | 57 |
| | |
| Tabela 1 - Lista dos candidatos à Presidência - informações gerais..... | 42 |
| Tabela 2 - Número de curtidas das fanpages dos candidatos no Facebook e número de seguidores por ordem crescente de curtidas..... | 43 |
| Tabela 3 - Tabela de controle manual da coleta dos registros audiovisuais..... | 46 |
| Tabela 4 - Quantitativo em segundos do tempo de aparição no Facebook e TV durante a campanha à Presidência de 2018..... | 55 |
| Tabela 5 - Tabela de persistência dos conteúdos audiovisuais nas fanpages dos candidatos no Facebook - novembro..... | 56 |
| Tabela 6 - Tabela de plataformas de arquivamento da Web - registros do Facebook..... | 62 |
| Tabela 7 - Tabela de plataformas de arquivamento da Web - registros do Twitter..... | 63 |
| Tabela 8 - Tabela de plataformas de arquivamento da Web - registros do Youtube..... | 64 |
| Tabela 9 - Tabela de plataformas de arquivamento da Web - registros do Instagram..... | 65 |
| Tabela 10 - Tabela de plataformas de arquivamento da Web - registros dos sites..... | 67 |

LISTA DE SIGLAS E ABREVIACÕES

| | |
|--------|---|
| CI | Ciência da Informação |
| DeltCI | Dicionário eletrônico de terminologia em Ciência da Informação |
| WWW | <i>Word Wide Web</i> |
| HTTP | <i>HyperText Transfer Protocol</i> |
| URL | <i>Uniform Resource Locator</i> |
| MB | <i>Megabyte</i> |
| GB | <i>Gigabyte</i> |
| TSE | Tribunal Superior Eleitoral |
| TRE | Tribunal Regional Eleitoral |
| UNESCO | <i>United Nations Educational, Scientific and Cultural Organization</i> |
| IIPC | <i>International Internet Preservation Consortium</i> |
| LiWA | <i>Living Web Archives</i> |
| HTML | <i>Hypertext Markup Language</i> |
| WAS | <i>Web Archiving Service</i> |
| API | <i>Application Programming Interface</i> |
| OAIS | <i>Open Archival Information System</i> |
| ID | <i>Identification</i> |
| MySQL | <i>My Structured Query Language</i> |
| AWEB | Arquivamento das Eleições Brasileiras |
| NUAWEB | Núcleo de Pesquisa em Arquivamento da <i>Web</i> e Preservação Digital |
| PT | Partido dos Trabalhadores |
| SD | <i>Standard Definition</i> |
| HTTPS | <i>Hypertext Transfer Protocol Secure</i> |

SUMÁRIO

| | |
|--|-----------|
| 1 INTRODUÇÃO | 11 |
| 1.1 Objetivo Geral..... | 13 |
| 1.2 Objetivos Específicos | 13 |
| 1.3 Justificativa..... | 14 |
| | |
| 2 A PERSONALIZAÇÃO DA POLÍTICA EM MÍDIAS SOCIAIS E O ARQUIVAMENTO DA WEB..... | 17 |
| 2.1 O Ambiente do <i>Facebook</i> | 18 |
| 2.1.1 Os Conteúdos Audiovisuais no <i>Facebook</i> | 19 |
| 2.1.2 O Fenômeno Infocomunicacional no <i>Facebook</i> | 21 |
| 2.1.3 Campanha Eleitoral Brasileira no <i>Facebook</i> | 22 |
| 2.2 A Personalização da Política nas Redes Sociais | 23 |
| 2.3 O Arquivamento da <i>Web</i> | 26 |
| 2.3.1 Políticas de Arquivamento/Seleção | 28 |
| 2.3.2 Tecnologias e Infraestruturas de Arquivamento da <i>Web</i> | 30 |
| 2.3.3 Preservação Digital..... | 32 |
| 2.3.4 Arquivamento da <i>Web</i> das Eleições | 34 |
| | |
| 3 METODOLOGIA..... | 37 |
| | |
| 4 OS PRESIDENCIÁVEIS E AS PUBLICAÇÕES EM MÍDIAS SOCIAIS | 41 |
| 4.1 Os Candidatos das Eleições Presidenciais de 2018 | 41 |
| 4.2 Conteúdos Audiovisuais Coletados | 44 |
| 4.2.1 Primeiro Turno das Eleições | 48 |
| 4.2.2 Segundo Turno das Eleições | 51 |
| 4.2.3 Comparativos do 1º e 2º Turnos | 52 |
| 4.3 Disponibilidade dos Vídeos Publicados pelos Candidatos | 56 |
| 4.4 Preservação em Plataformas de Arquivamento da <i>Web</i> | 60 |
| 4.5 Apresentação do Protótipo de Arquivamento..... | 67 |
| 4.5.1 Recursos de Consulta..... | 68 |
| 4.5.2 Recursos de Visualização | 70 |
| | |
| 5 CONSIDERAÇÕES FINAS | 73 |
| | |
| REFERÊNCIAS..... | 77 |
| | |
| APÊNDICE A – OBJECT INFORMATION DO DB ARCHIVE – TBL_ALVO..... | 83 |

| | |
|---|------------|
| APÊNDICE B – OBJECT INFORMATION DO DB ARCHIVE – TBL_ARCHIVE | 84 |
| APÊNDICE C – OBJECT INFORMATION DO DB ARCHIVE – TBL_LOG | 85 |
| APÊNDICE D – SCRIPT DA PÁGINA DE INCLUSÃO DO PROTÓTIPO | 86 |
| APÊNDICE E – SCRIPT DO BOT-SCRAPING DO PROTÓTIPO | 88 |
| APÊNDICE F – SCRIPT DA PÁGINA DE CONSULTA DO PROTÓTIPO | 95 |
| APÊNDICE G – LISTA DOS LINKS DAS MÍDIAS SOCIAIS E SITES DOS CANDIDATOS À PRESIDÊNCIA DE 2018 | 97 |
| APÊNDICE H – LISTA DOS LINKS DAS MÍDIAS SOCIAIS E SITES DOS CANDIDATOS À PRESIDÊNCIA DE 2018 | 99 |
| ANEXO A – RESULTADO DO PRIMEIRO TURNO DAS ELEIÇÕES PRESIDENCIAIS DE 2018 | 105 |
| ANEXO B – INFOGRÁFICO DE TEMPO DE PROPAGANDA NO RÁDIO E NA TV DOS CANDIDATOS À PRESIDÊNCIA DE 2018 | 106 |

1 INTRODUÇÃO

A constante presença das campanhas eleitorais nas mídias sociais decorre do desenvolvimento do fenômeno da personalização da política, onde há uma constante busca por novas formas de interação entre políticos e cidadãos a fim de destacar os traços pessoais dos políticos. Nossa preocupação recai sobre a ausência de políticas de preservação digital no Brasil, por instituições públicas e privadas, que permitam a busca e recuperação dos conteúdos e interações produzidos na *Internet*, algo que o arquivamento da *Web* possibilita de forma retrospectiva.

O contexto desta pesquisa desenvolveu-se a partir de um dos eventos mais importantes do Brasil, as eleições presidenciais de 2018, e trouxe para uma posição de destaque a atual realidade da produção e comunicação da informação em redes sociais.

Nosso objeto de pesquisa envolveu os conteúdos audiovisuais publicados pelos treze candidatos à Presidência, bem como do ex-presidente Lula, em suas *fanpages* oficiais no *Facebook*, durante a campanha eleitoral do primeiro e segundo turnos compreendidos entre 16 de agosto de 2018 a 04 de outubro de 2018 e 08 de outubro a 26 de outubro de 2018, composto por 2821 registros.

Segundo Silva (2006), na Ciência da Informação (CI) trans e interdisciplinar, a informação atua como objeto central a partir da interação contínua que desenvolve a capacidade humana e social de representar e conhecer o mundo, assim, as informações enriquecidas pela interação humana no *Facebook* concentram nosso objeto de pesquisa no fenômeno infocomunicacional a fim de compreender a produção, a finalidade, a ocorrência, a preservação, a transmissão, o uso e a transformação desse fluxo humano e social de signos, de símbolos, de representações de todo o tipo. (DELTICI, 2018; SILVA, 2006).

Para esta pesquisa, nos interessou, particularmente, a produção da informação nas *fanpages* dos candidatos (publicações/postagens), em especial dos conteúdos audiovisuais, a recepção destas postagens pelo público seguidor a partir das interações e visualizações, bem como o uso das publicações através do compartilhamento das mesmas e, por fim, a ocorrência da interação por meio das “reações” e dos comentários.

Este estudo almejou, como objetivo geral, compreender a necessidade de preservação digital dos vídeos publicados pelos candidatos à Presidência no *Facebook*, por meio do arquivamento da *Web*. E dividiu-se em cinco objetivos específicos, o primeiro consistiu em identificar os candidatos à Presidência na eleição de 2018, bem como suas *fanpages* oficiais na rede social *Facebook*; o segundo visou coletar todos os conteúdos audiovisuais publicados

pelos candidatos, de forma semiautomatizada, juntamente com o conteúdo descritivo textual e os respectivos metadados. O terceiro objetivo destinou-se a verificar se os vídeos publicados pelos candidatos ainda estavam disponíveis nas *fanpages* oficiais; o quarto objetivo verificou se os conteúdos dos *sites* e mídias sociais dos candidatos foram preservados em outras plataformas de arquivamento da *Web*. Já o quinto objetivo específico consistiu em disponibilizar o protótipo de consulta dos registros audiovisuais coletados e preservados.

No capítulo 3, a metodologia utilizada foi composta por revisão bibliográfica a partir do portal de periódicos CAPES sobre duas teorias em destaque: a personalização da política e o arquivamento da *Web*; e análise documental baseada nos dados coletados. Ainda, foram utilizadas referências que auxiliaram na descrição do campo de estudo proposto, ou seja, o infocomunicacional e na compreensão dos paradigmas pós-custodial (SILVA et al, 2009) e tecnológico (CASTELLS, 2013). Para a coleta dos vídeos foram utilizados recursos informatizados a partir do desenvolvimento de interfaces *Web* de inserção, consulta e um *bot*¹ que viabilizou a coleta dos metadados, também foi feita a estruturação de três tabelas em uma base de dados para armazenar as informações coletadas, permitindo a posterior análise das mesmas. As análises foram executadas por processos de extração de informações do banco de dados (*queries*), sendo que a pesquisa foi quantitativa-qualitativa, com abordagem exploratória-descritiva.

A análise dos dados coletados está dividida em cinco seções do cap. 4, onde a primeira (4.1) apresenta os candidatos das eleições presidenciais de 2018 e como estes estiveram presentes no *Facebook* a partir de informações disponíveis nas páginas públicas.

Já a seção 4.2 relata como foram coletados os registros audiovisuais dos candidatos e quais foram os recursos utilizados sendo interpretadas, nas subseções 4.2.1, 4.2.2 e 4.2.3, as informações do 1º turno das eleições, do 2º turno das eleições e o comparativo dos dois turnos.

A terceira seção (4.3) apresenta a análise de disponibilidade do formato de origem dos vídeos coletados, realizando a apuração geral e por candidato, bem como a análise de exemplos específicos de casos de indisponibilidade.

Analizamos, na seção 4.4, se houve a preservação das fontes de informações dos candidatos nas plataformas *Internet Archive*, *Archive-It* e *Archive.is*, considerando as mídias sociais *Facebook*, *Twitter*, *Youtube* e *Instagram*, bem como os *sites* oficiais de campanha.

¹ Aplicação de *software* que desempenha uma tarefa de forma autônoma.

Finalizamos o cap. 4 com a seção 4.5, onde é apresentado o protótipo de arquivamento utilizado para a coleta e que será disponibilizado para futuras pesquisas, sendo explicados nas subseções 4.5.1 e 4.5.2 os recursos de consulta e visualização.

Dessa forma, chegamos ao problema de pesquisa e seu consequente questionamento: No contexto da personalização da política em mídias sociais, de que forma o arquivamento da web, de vídeos publicados no Facebook pelos candidatos à Presidência de 2018, contribuem para a preservação digital?

1.1 Objetivo Geral

Compreender a necessidade de preservação digital dos vídeos publicados pelos candidatos à Presidência no *Facebook*, por meio do arquivamento da *Web*.

1.2 Objetivos Específicos

- a) Identificar os candidatos das eleições presidenciais de 2018 e suas respectivas *fanpages* oficiais;
- b) Coletar os conteúdos audiovisuais publicados pelos candidatos no período de propaganda eleitoral do 1º e do 2º turno;
- c) Verificar se os vídeos publicados pelos candidatos ainda estão disponíveis nas *fanpages* oficiais;
- d) Verificar se os conteúdos dos *sites* e mídias sociais dos candidatos foram preservados em outras plataformas de arquivamento da *web*.
- e) Disponibilizar o protótipo de consulta dos registros audiovisuais coletados e preservados.

1.3 Justificativa

Para justificar as principais teorias utilizadas (personalização da política e arquivamento da *Web*) buscamos aporte em referências como Manin (1995) que, já na década de 90, afirmava que a representação política estava passando por uma crise nos países ocidentais com o surgimento da personalização da política. Desta forma, a estratégia em vigor de se fazer política e campanha dita o uso, cada vez mais popular das redes sociais, onde os políticos vêm investindo em aparições espontâneas que os mostrem como pessoas comuns e, conseqüentemente, mais próximos de seus eleitores. Já o arquivamento da *Web*, como possibilidade de preservação digital, faz emergir no Brasil uma preocupação que vem sendo fonte de pesquisas, estudos e projetos internacionais desde 1996, com a criação da iniciativa *Internet Archive*. Com o aumento da produção informacional e da comunicação via *Web* é essencial a promoção de projetos que visem a preservação desses conteúdos, mesmo que se dê de forma independente e sem incentivos governamentais. É preciso demonstrar que grande parcela das informações que deveriam compor a memória nacional serão perdidas se não houver o fomento de iniciativas de arquivamento da *Web* e, em especial, das informações públicas produzidas em redes sociais.

Entendemos que a memória de um país deve contemplar as novas formas de comunicação, pois trata-se de uma forma de agir, comunicar e informar típicos de uma determinada época. A forma de comunicação atual torna-se muito importante para a memória, uma vez que desencadeia uma interação entre produtor e receptores e vice-versa. E assim, Gautier (2018) corrobora com nossa preocupação afirmando que em uma era em que um celular possibilita registrar e compartilhar cada instante de vida das pessoas, o armazenamento da informação para as futuras gerações mostra-se um grande desafio, onde a proteção dessa memória envolve um grande problema de curadoria: como filtrar e escolher o que preservar? (GAUTIER, 2018). A memória digital será um dos grandes desafios de pesquisadores, cientistas e historiadores num futuro bem próximo e, em decorrência da lenta evolução de pesquisas e projetos voltadas para a temática, sofrerá com perdas significativas. Assim, pensando em estratégias que visem mitigar as perdas de informações da *Web*, esta pesquisa voltou-se para registros de interesse público atuando como uma curadoria pertinente a muitas pessoas, bem como voltada para uma das plataformas que mais efemera suas publicações. E ainda, a partir desta preservação, é possível criar fontes de validação das informações, bem como possibilitar o uso e reuso da informação a partir de repositórios/Arquivo da *Web*, fomentando o *open data* (dados abertos).

Outro aspecto frágil que pode dificultar o desenvolvimento de iniciativas de preservação da *Web* no Brasil é a falta de clareza da legislação brasileira quanto ao direito de se informar dos usuários em contrapartida ao direito autoral ou ao direito à privacidade.

Começamos pela Lei nº 4.117, de 27 de agosto de 1962 – que institui o Código Brasileiro de Telecomunicações, que determina no art. 71, § 3º, que as gravações dos programas políticos, de debates, entrevistas, entre outros, deverão ser conservadas em arquivo pelo prazo de 20 a 30 dias – conforme o caso, depois de transmitidas. (BRASIL, 2017a). Durante esses períodos “as gravações ficarão no arquivo da emissora, mas à disposição da Justiça Eleitoral, para servir como prova sempre que requerido”. (TRIBUNAL SUPERIOR ELEITORAL, 2019). Assim, percebemos que não existe a obrigatoriedade das emissoras e da Justiça Eleitoral de preservar, nem mesmo, os registros físicos dos programas políticos transmitidos na mídia convencional (rádio e televisão). Aqui, percebemos uma lacuna significativa para que o cidadão possa exercer o seu direito de acesso à informação – sendo que os registros das propagandas eleitorais não possuem garantia de preservação. Outra preocupação consiste em o Brasil não possuir uma legislação que assegure a preservação desses registros e a Lei 12.527/Lei de Acesso à Informação regula apenas o acesso a informações dos órgãos públicos ou de economia mista, não havendo menção a produção de candidatos políticos durante os períodos de campanha eleitoral mesmo que estes sejam eleitos. (BRASIL, 1997).

Contudo, entendemos que o processo de campanha eleitoral gera registros valiosos não apenas para o período das eleições, mas também após a sua conclusão. Nesse contexto, as propagandas eleitorais e todas as suas formas de manifestações retratam como transcorreu um determinado período eleitoral, quais foram os posicionamentos e discursos que influenciaram os resultados das eleições e servirão de fonte de informação para que cidadãos fiscalizem as atitudes dos políticos eleitos. Por outro lado, não foram encontrados impedimentos quanto a preservação da produção de candidatos políticos, uma vez que só configuram informações privadas aquelas que os usuários não disponibilizam voluntariamente para o acesso público. Estas diretivas podem ser observadas tanto nas políticas de privacidade do *Facebook* (FACEBOOK, 2018a), quanto no Marco Regulatório da *Internet* (BRASIL, 1997).

Como exemplo, trazemos uma publicação usada em alguns países, como os EUA, denominada *Fair Use* (Uso Justo) que também se alinha à missão do *Internet Archive* (BAILEY, 2016), de fornecer acesso universal a todo o conhecimento humano. O *Fair Use* atua como ponto de equilíbrio entre a Lei de Direitos Autorais (*Copyright*), garantindo que a proteção concedida aos autores não sufoque a própria criatividade e inovação que foi projetada

para promover. Assim, é garantido ao público acesso à informação para a educação, a pesquisa, a criatividade e a descoberta - atividades essenciais ao funcionamento de uma sociedade democrática. O *Fair Use* também permite o acesso e a reutilização de materiais para criticá-los ou comentá-los, para fins educacionais ou de maneiras que alterem o original com uma nova mensagem ou significado.

2 A PERSONALIZAÇÃO DA POLÍTICA EM MÍDIAS SOCIAIS E O ARQUIVAMENTO DA WEB

Em Silva *et al* (2009), o processo informacional, em decorrência dos distintos processos históricos, é dividido em três fases principais, duas no paradigma custodial e uma inserida no pós-custodial. A primeira fase, denominada de sincrética e custodial, surge em decorrência da Revolução Industrial, das revoluções políticas e sociais e da valorização pela identidade nacional, e, em decorrência, os conceitos de Arquivo Nacional e fundo são estabelecidos como sinônimo de proteção da memória nacional. A primeira fase perdurou do século XVIII até 1898 – a partir deste ano, instalou-se a fase técnica e custodial que, em consequência da evolução tecnológica e científica, voltou-se a especialização profissional. (SILVA *et al*, 2009).

O paradigma custodial, a partir de 1980, dá espaço ao paradigma pós-custodial informacional e científico, proposto por Silva *et al* (2009), que traz a emergência do cientista da CI onde, em uma sociedade e globalização da informação, é potencializado o acesso pelos Sistemas de Informação. Nesta pesquisa, adotamos o paradigma pós-custodial concomitantemente ao paradigma tecnológico, proposto por Castells (2013), por entendermos que ambos se complementam neste contexto onde a era da informação é impulsionada pelas tecnologias.

O paradigma tecnológico, segundo Castells (2013), tem como primeiro aspecto o fato de a informação ser a sua matéria-prima, uma vez que as tecnologias precisam, necessariamente, agir sobre as informações. O segundo aspecto determina que somos moldados pelo novo meio tecnológico e, a terceira característica, afirma que, ao se usar essas novas tecnologias de informação, qualquer sistema ou conjunto de relações seguirá a lógica de redes. O autor acredita que o crescimento das redes, após sua difusão, é exponencial, pois as vantagens que surgem a partir de seu uso são igualmente exponenciais. E, assim, quanto mais a rede cresce mais aumenta o declínio de oportunidades de se obter outros elementos fora da rede. (CASTELLS, 2013).

A comunicação em massa, por muito tempo, limitou-se a uma comunicação unidirecional, mas, com a difusão da *Internet*, surgiu a interatividade, que possibilita a transmissão de muitas mensagens para muitas pessoas em tempo real. Castells nomeia esse processo historicamente novo como auto comunicação de massa. (CASTELLS, 2017). Este tipo de interação vem sendo fonte de pesquisa há muito tempo, no entanto, segundo Weare e Lin (2000), a *World Wide Web* (WWW) vem abrindo campos adicionais para a pesquisa analítica de conteúdo e desenvolvendo novas técnicas empíricas. Os autores afirmam que tornou-se

quase obrigatório que campanhas políticas e organizações de notícias tenham presença na *Web*. E, assim, existe grande probabilidade da mídia tradicional, acrescida da *Internet*, sofrer uma série de influências nas mensagens de campanha, discurso político, reportagem de notícias e consumo de notícias. Essas mudanças afetarão potencialmente os processos democráticos fundamentais, incluindo a socialização política, o discurso público e o papel da mídia na definição da agenda. (WEARE; LIN, 2000).

Quando nos referimos a memória de um acontecimento nacional, como as eleições brasileiras, torna-se pré-requisito considerar a preservação digital e descobrir como poderá ser capturado o conteúdo que é considerado válido preservar. Segundo Hockx-Yu et al (2010), esse tem sido um desafio significativo para o arquivamento da *Web*, especialmente para *sites* com conteúdo de mídia de streaming incorporado, que não podem ser copiados por meio de uma solicitação HTTP² simples para um URL³.

2.1 O Ambiente do *Facebook*

Para compreender os desafios e peculiaridades da presente pesquisa, é importante descrever o conjunto de aspectos que irão delimitar nosso contexto, assim, iniciamos com uma breve narrativa sobre o surgimento do *Facebook* e a evolução dos recursos ofertados.

Mark Zuckerberg registrou, em 11 de janeiro de 2004, o endereço *TheFacebook.com*, antes mesmo da criação de seu produto. Segundo Kirkpatrick (2011), Zuckerberg vislumbrou a criação de um novo serviço que fosse uma ferramenta de comunicação básica que possibilitasse acompanhar seus colegas da faculdade e o que acontecia com eles. Em 04 de fevereiro de 2004, o *TheFacebook.com* foi ao ar como um *site* e a partir desse momento foi acumulando novas funcionalidades que foram sendo desenvolvidas, incluindo a incorporação de anúncios – começando, assim, a gerar receita. Cabe salientar, como descreve o autor, que o objetivo de Zuckerberg era aumentar o número de usuários do *site* e não o lucro, sendo a receita necessária para subsidiar e impulsionar as melhorias do *site*. Em 20 de setembro de 2005, a empresa tornou-se, oficialmente, *Facebook* – sob o domínio *Facebook.com*. No entanto, no outono de 2006 é que Zuckerberg começou a construir uma plataforma que permitisse que outros pudessem criar aplicativos em cima do *Facebook*.

O *Facebook* possui diversos tipos de produtos e serviços que geram valores de receita, todavia, a principal fonte está ligada a promoção de anúncios, sendo que funciona de maneira

² *Hyper Text Transfer Protocol* ou Protocolo de Transferência de Hipertexto.

³ *Uniform Resource Locator*, ou Localizador Padrão de Recursos. Consiste em um endereço virtual.

mais ativa, enquanto em outras empresas ainda predomina a abordagem passiva. Pelo fato de o *Facebook*, conforme Kirkpatrick (2011), obter, voluntariamente, dados preciosos sobre seus usuários – a empresa consegue direcionar anúncios que sejam pertinentes ao seu público com base em gostos e interesses, ao invés de ofertar produtos aleatórios.

O perfil pessoal é o primeiro passo para incluir-se na rede social, após realizar o cadastro é possível participar de grupos e criar *fanpages* (páginas de fãs). As *fanpages* são páginas públicas que visam dar destaque para um negócio ou marca, bem como, para uma comunidade ou figura pública. Segundo as Políticas específicas de páginas do *Facebook* a visibilidade do “conteúdo publicado em uma página é público e pode ser visto por qualquer pessoa apta a acessar a página”. (FACEBOOK, 2018).

Os anúncios do *Facebook*, nominados como *Facebook Ads* ou Publicações Não-publicadas, como descreve GhFly (2018), são o formato de anúncio mais utilizado pelas empresas no *Facebook*, pois não são publicações que ficarão para sempre dentro da *fanpage*. Esses anúncios são salvos dentro do acervo de anúncios da conta de anúncios do *Facebook* da *fanpage*, na plataforma de Gerenciamento de Negócios do *Facebook*.

Além do serviço de anúncios (*Facebook Ads*), uma *fanpage* pode contratar o serviço de impulsionamento de uma publicação já existente na página. Este impulsionamento, como explica GhFly (2018) pode alcançar os seguidores da *fanpage*, os amigos dos seguidores, pessoas que curtiram determinada página do *Facebook*, entre outras estratégias de segmentação disponíveis na plataforma. Este recurso de impulsionamento tem sido bastante usado em páginas de políticos brasileiros.

Além do *Facebook* ser a mídia social com o maior número de usuários do mundo, no Brasil a empresa também é a mais popular como fonte de notícias, como relata Valente (2018), o percentual é de 66%, seguido pelos Estados Unidos (45%), Reino Unido (39%) e França (36%).

2.1.1 Os Conteúdos Audiovisuais no *Facebook*

O Facebook lançou em 2007 um serviço de vídeo que, como descreveu Cashmore (2018) no site *Mashable*, visava competir com todos os principais sites de compartilhamento de vídeo. O serviço iniciou com upload de vídeos e a gravação direta da Webcam, permitindo a marcação de nomes de amigos. O novo recurso também era integrado ao *feed* de notícias do Facebook com controles de privacidade para a seção de vídeos. Inicialmente a duração máxima

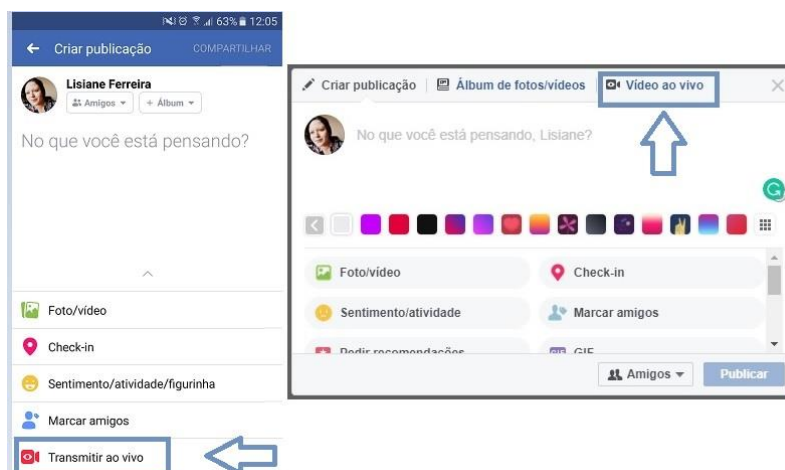
aceita, do vídeo, era de 10 minutos e tamanho de 200 MB, mas, atualmente, são aceitos vídeos de até 120 minutos com tamanho máximo do arquivo de 4.0 GB. (CASHMORE, 2018; RUTHES, 2018).

Em agosto de 2015, o *Facebook* começou a permitir que alguns usuários selecionados transmitissem *live streaming*, recurso chamado de *live* por seus criadores. Para compreender o *live*, trazemos o conceito da *live streaming* ou *live stream* que é amplamente definida como uma transmissão ao vivo de um evento pela *Internet*. (BELL, 2018; OXFORD LIVING DICTIONARIES, 2018).

Em janeiro de 2016, o recurso começou a ser liberado para o público em geral. Os *lives* aparecem no feed de notícias e os usuários podem comentar sobre eles em tempo real. As transmissões ao vivo são salvas automaticamente como uma postagem de vídeo na página do *streamer*⁴. Em março de 2017, o *Facebook* também estendeu o suporte a *streaming* ao vivo para PCs, ou seja, a transmissão pode ser gravada através da *webcam* do computador ou notebook. Assim, o *Facebook live* aproveita o público existente do *site*, o maior público do mundo, para competir com serviços ao vivo semelhantes. (BELL, 2018).

O recurso do *live* pode ser acessado pelo computador ou pelo dispositivo móvel, na parte superior do feed de notícias ou do perfil, selecionando o ícone ao vivo abaixo do local em que se insere o status, respondendo a: "No que você está pensando?", como mostra a **Figura 1**:

Figura 1 - Tela de transmissão ao vivo no dispositivo móvel e no computador.



Fonte: Perfil da autora no *Facebook*, 2019.

⁴ Usuário que transmitiu o *Live*.

O usuário da rede social *Facebook* para utilizar o *live* deverá conceder a permissão para o aplicativo usar a câmera para gravar vídeos, caso ainda não tenha feito isso. É possível escolher o nível de privacidade entre o compartilhamento público, todos os amigos, amigos, exceto os especificados, ou somente para você. Poderá ser inserida uma descrição para o vídeo, bem como a localização e a marcação de amigos. Após a inserção das informações desejadas, basta tocar no botão “Iniciar vídeo ao vivo”. Os *lives* suportam comentários ao vivo, gravação de tela, filtros, espelhamento de vídeo vertical ou horizontal e reações dos espectadores. E os usuários também podem bloquear visualizadores específicos, se desejarem. Embora o limite máximo de tempo para uma transmissão tenha sido inicialmente de 30 minutos, os usuários agora podem permanecer ativos por até quatro horas. (TECHTARGET, 2018). Em resumo, o *live* é um vídeo disponibilizado durante a sua produção e, posteriormente, mantido no perfil do produtor.

2.1.2 O Fenômeno Infocomunicacional no *Facebook*

Justificamos nossa abordagem, através do fenômeno infocomunicacional, por entendermos, assim como Lomborg (2012), que as mídias sociais se apresentam como textos disponíveis para análise textual clássica de produção e audiência, sendo possível através das práticas comunicativas dos usuários. Assim, é possível afirmar que não há “texto” pré-produzido, em nosso caso pelo *Facebook*, apenas no referencial tecnológico ou no modelo comunicativo facilitado por esta plataforma. É preciso haver o “preenchimento” pelas contribuições comunicativas dos usuários. Assim, nas mídias sociais, o texto se baseia no engajamento contínuo do usuário; sendo emergente, editável e passando por um processo contínuo de desenvolvimento. (LOMBORG, 2012). Assim, situamos os registros coletados no fenômeno infocomunicacional para compor o nosso objeto de pesquisa e representamos os três eixos na **Figura 2**:

Figura 2 - Representação do Fenômeno Infocomunicacional nos *Lives*.



Fonte: Elaborado pela autora, 2019.

O primeiro eixo da infocomunicação, no *live* por exemplo, é compreendido como a sua própria produção, mas não apenas do vídeo ao vivo – mas também da descrição que o acompanha. Já o segundo eixo, que define a recepção e o uso da informação criada – pode ser identificada através do número de visualizações, da quantidade e tipos de reações, bem como pelo número de compartilhamentos. E, por fim, o terceiro eixo, que traz a interação e a comunicação – onde podemos associar com os comentários e, também, com as reações, tendo em vista que essas últimas transmitem sentimentos diversos.

2.1.3 Campanha Eleitoral Brasileira no *Facebook*

A propaganda eleitoral brasileira é regulamentada pelo Tribunal Superior Eleitoral (TSE), que define o período que a mesma deverá ocorrer, quais meios e quais critérios devem atender as propagandas, bem como, qual será o tempo de cada candidato no horário eleitoral gratuito.

A Lei nº 9.504 (Lei das Eleições), de 30 de setembro de 1997, determina em seu art. 36 que a propaganda eleitoral inicia após o dia 15 de agosto do ano da eleição, sendo permitidos os comícios, carreatas, distribuição de material gráfico e **propaganda na Internet**, entre outras formas. Contudo, não é permitido qualquer tipo de propaganda política paga no rádio e na televisão. A mesma lei, em seu art. 47, encerra a propaganda eleitoral no dia 04 de outubro. (TRIBUNAL SUPERIOR ELEITORAL, 2018b).

O TSE e os TREs elaboraram até o dia 24 de agosto o plano de mídia para uso da parcela do horário eleitoral gratuito a que tinham direito os candidatos, garantindo a todos a participação nos horários de maior e menor audiência. A propaganda eleitoral gratuita no rádio e na televisão teve início em 31 de agosto e término no dia 4 de outubro. (TRIBUNAL SUPERIOR ELEITORAL, 2018c). No entanto, a campanha nas redes sociais iniciou imediatamente no dia 15 de agosto, sendo percebidas publicações antes desse período já com conteúdo de campanha eleitoral.

A propaganda eleitoral no segundo turno, de acordo com o art. 43 da Lei das Eleições, foi permitida depois do dia 15 de outubro, sendo encerrada, conforme o art. 49, no dia 26 de outubro. (TRIBUNAL SUPERIOR ELEITORAL, 2018b).

Cabe aqui uma breve explicação de como funciona o nosso sistema eleitoral onde, como explica Blume (2018), o presidente brasileiro é eleito pelo sistema majoritário. Sendo necessário que o candidato conquiste a maioria absoluta dos votos, ou seja, mais de 50% dos votos válidos. Assim, quando nenhum candidato atinge essa marca é preciso marcar o segundo turno, disputado pelos dois candidatos mais votados. Vencendo, no segundo turno, quem conseguir mais votos.

O TSE veda a veiculação de qualquer tipo de propaganda eleitoral paga na *Internet*, excetuado o impulsionamento de conteúdos no *Facebook*, desde que identificado de forma inequívoca como tal e contratado exclusivamente por partidos, coligações e candidatos e seus representantes. No entanto, o *Twitter* não é compreendido pelo TSE como uma plataforma que divulga a propaganda eleitoral, pois a rede social não propaga amplamente as manifestações nela publicadas. (TRIBUNAL SUPERIOR ELEITORAL, 2018c).

2.2 A Personalização da Política nas Redes Sociais

As redes sociais se tornaram parte integrante da vida de bilhões de usuários da *Internet* em todo o mundo e, como outras tecnologias da *Web 2.0*, permitem que pessoas sem qualquer conhecimento técnico específico criem um perfil on-line e se comuniquem e compartilhem informações com outras pessoas. Com estas facilidades, redes sociais como o *Facebook* e o *Twitter* expandem a arena política e aumentam o número de campanhas personalizadas. (BRONSTEIN, 2013; ENLI E SKOGERBØ, 2013).

Enli e Skogerbo (2013) afirmam que as mídias sociais contribuem para a personalização da política. Essa ideia não significa que a personalização seja um produto da presença da mídia social dos políticos, mas que as mídias sociais se encaixam em processos contínuos de longo

prazo nos quais a comunicação política se tornou cada vez mais focada em personalidades e traços pessoais dos políticos. Essas demandas estão ligadas à popularização da política e a personalização é fundamental. (ENLI E SKOGERBØ, 2013).

A personalização também pode ser considerada uma característica intrínseca dos desenvolvimentos sociais e da mídia no último século. As reivindicações de candidatos políticos para se apresentarem como pessoas multidimensionais com lados pessoais e privados, bem como imagens públicas, aumentaram, em parte, devido à crescente competição na mídia. (LILLEKER, 2010 apud ENLI E SKOGERBØ, 2013).

Já Bronstein (2013) acredita que as redes sociais têm potencial para tornarem-se uma ferramenta decisiva para a disseminação da informação política e a criação do discurso político, permitindo que informações e opiniões sejam diretamente transferidas de políticos e partidos políticos para usuários e entre os próprios usuários. (BRONSTEIN, 2013).

O estudo de Enli e Skogerbø (2013) envolveu as eleições parlamentares de 2009 e as eleições locais norueguesas de 2011, identificando três razões principais que motivaram os candidatos a usarem as redes sociais, *Facebook* e *Twitter*, as quais foram:

1. O *marketing*, através do aumento da visibilidade de suas candidaturas e partidos na esfera pública, bem como, atrair os jovens e ser atraente para os jornalistas;

A mobilização como um motivo-chave para estar presente no Facebook e no Twitter, demonstrando que as mídias sociais melhoraram o efeito de outras estratégias de campanha, on-line e off-line; e

2. A oportunidade de diálogo com os eleitores, onde vários políticos alegaram que a mídia social representava novas oportunidades de conexão com os eleitores, recebendo *feedback* sobre questões políticas, discutindo política de forma mais contínua e envolvendo mais eleitores do que a mídia anterior. (ENLI E SKOGERBØ, 2013).

Abejón, Sastre e Linares (2012) em outro estudo, envolvendo as eleições municipais e regionais de 2011 na Espanha, mostraram que os cidadãos buscavam nas redes sociais a melhor relação com seu ambiente imediato, mas excluía a política desse ambiente. Assim, os espanhóis entendiam que os políticos tinham redes sociais para se aproximarem dos cidadãos, no entanto, os cidadãos é que não queriam se aproximar dos políticos. O *Facebook*, considerado uma das maiores revoluções sociais dos últimos anos, prova ser uma plataforma onde terceiros, neste caso políticos, fazem “negócios”. Onde se originam informações, que a priori parecem transparentes, mostrando suas propostas e ideias e deixando aberta a possibilidade de que o cidadão mostre sua concordância ou seu desacordo. O “negócio” define-se na ação em que o

político lhe dá algo que antes não lhe dava buscando obter mais um voto e se aproximar da população mais jovem, que compõem a maior porcentagem de usuários dessas redes. (ABEJÓN, SASTRE E LINARES, 2012).

O estudo desenvolvido na Nova Zelândia por Ross, Fountaine e Comrie (2015) buscou entender como os deputados do 49º Parlamento, excluindo os que se encontravam nas eleições de 2011, usavam o *Facebook* como uma ferramenta para comunicação política no contexto de uma eleição geral. Nessa pesquisa, o principal objetivo consistiu em mapear a presença on-line dos políticos para determinar até que ponto eles usavam a *Internet* e/ou mídias sociais para a campanha, concentrando-se em seus *sites* pessoais, páginas do *Facebook* e contas do *Twitter*, em vez de monitorar *sites* e *blogs* do partido. No país, são conhecidas as considerações e disciplina por parte do partido, contudo, a maioria dessas páginas pessoais tem uma declaração explícita sobre a responsabilidade pessoal pelas opiniões expressas, sugerindo que os partidos políticos estão dispostos a se distanciar do conteúdo das páginas pessoais dos parlamentares. A abordagem proposta, distinguiu as tipologias de usuários em "moderados" e "ativos", de Enli e Skogerbø (2013), e ignorou o que chamaram de "invisível" e "silencioso", identificando, por exemplo, a partir dos dados coletados, que cada deputado não realizava mais do que cinco postagens por dia, sendo que a pesquisa se deu sobre todos os tipos de publicações (textos, vídeos, fotos, links, etc). (ROSS, FOUNTAINE E COMRIE, 2015).

A campanha política através do *Facebook*, nesse contexto da personalização, também introduz novas perspectivas para o estudo da hipótese de equalização e da hipótese de normalização, essa última muito mais presente em estudos de *Websites* pessoais e/ou de partidos políticos. Segundo, Larsson e Moe (2014) apud Lev-On e Haleva-Amir (2016), a hipótese de equalização sugere que a *Internet* promove a igualdade de uso entre os atores políticos dominantes e periféricos e seus públicos. Já a hipótese de normalização, defende que as lacunas de recursos e presença seriam reproduzidas e até ampliadas em novas plataformas de mídia. No entanto, o estudo apresenta inúmeras pesquisas que defendem as duas hipóteses, cada uma voltada para um país e/ou realidade distinta, restando a afirmação de que partidos periféricos, de fato, foram os primeiros a enxergarem o potencial das redes sociais e têm investido na plataforma como principal meio das campanhas.

Porém, estudos recentes apresentam um quadro com mais nuances, de acordo com o qual a hipótese da normalização também se sustenta em termos de uso do *Facebook*. Esse resultado pode se dar em decorrência do amadurecimento do *Facebook* em uma plataforma já estabelecida, e líder de uso pelo público em geral para fins políticos e não políticos. E assim,

os partidos políticos dominantes também começaram a migrar para essa plataforma. (LEV-ON E HALEVA-AMIR, 2016).

A partir dos estudos citados, entendemos que as mídias sociais ganharam um lugar de destaque na comunicação política, servindo de fonte de estudos em diversos contextos – não apenas como fonte de informação, mas como meio de comunicação – onde a interação ganha destaque. Contudo, a proposta desta pesquisa consiste em trazer reflexões sobre o que se está produzindo no *Facebook*, com especial atenção as publicações políticas, ou seja, de interesse público, que deveriam compor uma memória digital no futuro, mas que tendem a ser uma memória efêmera, quiçá resgatada parcialmente em algumas pesquisas e trabalhos.

2.3 O Arquivamento da Web

A *Web*, como descreve Masanès (2006), é uma mídia difundida e efêmera em que a cultura moderna encontra uma forma natural de expressão e, assim, muitos aspectos da sociedade estão acontecendo ou sendo refletidos na *Internet* em geral e na *Web* em particular.

Em 2003, de acordo com Toyoda e Kitsuregawa (2012), a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) considerou os materiais digitais como um patrimônio cultural e levantou a necessidade de ação para preservar esse patrimônio digital. Por esta razão, a preservação da *Web* se tornou uma necessidade cultural e histórica – todavia, para se tornar possível, exigiu uma revisão radical das práticas tradicionais de preservação.

Assim como outros autores, Masanès (2006) acredita que a legitimidade do arquivamento da *Web* está consolidada pelo fato da publicação e acesso na *Web* serem facilitados, criando uma fonte única desse tipo de conteúdo. Milligan, Ruest e Lin (2016) defendem que a *Web* se tornou parte integrante de nosso cotidiano, capturando nossa memória coletiva, a partir do registro de grandes eventos mundiais até as interações pessoais, que são valiosas para o estudo de nossa sociedade.

Brügger (2011), define o arquivamento da *Web*, de forma simplificada, como um ato deliberado e intencional, que implica consciência de que se está preservando o material e porque o material da *Web* está sendo preservado. Segundo o autor, o início da *Web* desencadeou esta necessidade de preservar o que foi criado ou encontrado no ciberespaço por usuários comuns, definidos pelo autor como amadores, os quais deram origem a uma das muitas formas do microarquivamento. Esses arquivamentos possibilitaram a única evidência existente da *Web*

antiga. Por outro lado, com os avanços das tecnologias de arquivamento da *Web* viabilizada por grandes instituições, foi possível desenvolver o macroarquivamento. (BRÜGGER, 2011).

A partir de Ferreira, Martins e Rockembach (2018), podemos dizer que o fomento da discussão sobre o registro e o armazenamento destas informações é importante para o acesso futuro, e o desenvolvimento do arquivamento da *Web* estabelece práticas que se propõem a garantir o potencial informacional e probatório das informações que dão subsídios e sustentam as pesquisas científicas.

As primeiras iniciativas de arquivamento da *Web* datam de 1996, como o *Internet Archive*, e estão ganhando mais destaque e engajamento por parte de diversas instituições, arquivos nacionais, bibliotecas e universidades do mundo. Os arquivos da *Web*, segundo Costa, Gomes e Silva (2017), da Pandora e da Tasmânia, da Austrália, e o arquivo da *Web* Kulturw3, da Suécia, também foram criados em 1996. A preservação pode, dependendo das necessidades e dos recursos, ser mais ampla ou mais específica, nacional ou internacional, por assuntos, entre outras possibilidades. O *International Internet Preservation Consortium* - IIPC (2018) é uma das principais iniciativas de arquivamento da *Web*, unindo, atualmente, 54 membros de 45 países. Formalizada em 2002, com apenas 12 membros franceses, surgiu com o objetivo de pensar na preservação da memória da *Internet* financiando seus próprios projetos.

Dados coletados em 2010 e 2014 por Costa, Gomes e Silva (2017) mostram que as iniciativas de arquivamento da *Web* são bastante heterogêneas em termos de tamanho e escopo, sendo que a maior parte dos arquivos da *Web* contém exclusivamente conteúdo relacionado ao seu país de hospedagem, região ou instituição. Havendo apenas três iniciativas, o *Internet Archive*, a *Internet Memory Foundation* e o Arquivo da *Web* Portuguesa (Arquivo.pt), que também preservam informações relacionadas a países estrangeiros.

O *Internet Archive*, como descreve Brown (2006), foi fundado por Brewster Kahle utilizando as contribuições do *Alexa Internet*, empresa de catalogação *Web* fundada por Kahle e Bruce Gilliat em 1996. Assim, as coletas realizadas pelo *Alexa* foram doadas para o *Internet Archive* e formaram a base de sua coleção principal. O *Internet Archive* é uma organização sem fins lucrativos e com o objetivo de construir uma biblioteca digital que oferece acesso para todos de coleções históricas. A coleta do material é feita por extração remota, método usado pela grande maioria dos programas de arquivamento da *Web* atualmente, que usa software especializado de rastreamento para copiar recursos da *Web* remotamente. Para realizar sua própria coleta e desenvolver técnicas aprimoradas de rastreamento, o *Internet Archive* iniciou o desenvolvimento do *Heritrix*, seu próprio rastreador de qualidade de arquivamento,

disponível em código aberto. O *Heritrix* também passou a formar a base para o desenvolvimento do rastreador da *Web* do IIPC. A iniciativa possui 14 projetos em andamento, entre eles o *Political TV Ad Archive*, *Building Libraries Together*, *Wayback Machine* e o *Archive-It*. (BROWN, 2006; INTERNET ARCHIVE, 2018).

A *Internet Memory Foundation* (anteriormente *European Archive*) foi originalmente fundada em 2004 e é uma instituição sem fins lucrativos com sedes em Amsterdã e Paris. A fundação está atualmente arquivando dezenas de *Terabytes* de dados por mês na região da Europa e está desenvolvendo várias tecnologias de arquivamento na *Web* para dar conta do crescimento e do uso da memória da *Internet*. No projeto *Living Web Archives (LiWA)*, a fundação se concentrou em melhorar a fidelidade de captura da *Web* e desenvolveu tecnologias que capturam a *Web* oculta e outros conteúdos que eram difíceis de coletar. O projeto também possibilitou filtrar o conteúdo indesejado através da detecção de *spam*. Outros aspectos do projeto abordam a incoerência temporal inerente aos métodos e ferramentas atuais de captura da *Web* e a rápida evolução semântica e tecnológica da *Web*, a fim de promover a viabilidade a longo prazo de arquivos da *Web*. É importante destacar que essas tecnologias foram lançadas como *open source*⁵ em 2010. (TOYODA E KITSUREGAWA, 2012; INTERNET MEMORY FOUNDATION, 2018).

Algumas iniciativas permitem que qualquer pessoa arquive uma página da *Web* com um URL, como é o caso do *Internet Archive*, o qual é mais conhecido, mas trazemos também a referência do *Archive.is* que possui uma interface simples e amigável que permite pesquisar URLs preservadas e suas versões ou adicionar um novo URL para ser arquivado de maneira bem intuitiva. Contudo, esses projetos atuam de forma semelhante, gerando um link exclusivo da página arquivada a fim de atender as várias necessidades dos usuários, como os acadêmicos que preservam as páginas da *Web* citadas em seu trabalho ou os tribunais superiores que preservam as citações em suas decisões publicadas. (COSTA, GOMES E SILVA, 2017).

2.3.1 Políticas de Arquivamento/Seleção

A primeira fase do arquivamento da *Web* compreende o estabelecimento de políticas de seleção, as quais irão determinar o tipo, a extensão e a qualidade da coleção resultante, e devem ser revisadas com regularidade. Quando essa prática se torna uma atividade regular,

⁵ Compartilhamento e melhoria colaborativa de código-fonte de *software*. (OPEN SOURCE INITIATIVE, 2018).

aconselha-se a criação de um documento de orientação que defina a política de desenvolvimento das coleções pela instituição. (MASANÉS, 2006).

A seleção compreende três fases consecutivas: preparação, descoberta e filtragem. Na preparação o objetivo é definir a meta de coleta, a política de captura e as ferramentas para implementá-la. Dependendo do objeto da coleta, por exemplo um evento como as eleições, é necessário o apoio técnico de profissionais como arquivistas, bibliotecários, historiadores, pesquisadores – no intuito de definir mais precisamente onde serão coletadas as informações, qual será a extensão e com qual frequência. Neste exemplo, é necessária uma busca manual dos endereços de *sites* da *Web* que sejam de interesse para a criação da coleção. Já a coleta de *sites* pelo domínio pré-estabelecido, pode ocorrer de forma totalmente automatizada. A fase de descoberta determina a lista de pontos de entrada que serão usados para a captura, bem como a frequência e o escopo dessa captura. Pode acontecer de forma endógena, onde é aproveitada a estrutura de links da *Web* para percorrer e encontrar novos materiais, considerando que os *sites* ou os recursos vinculados mereçam pertencer à mesma coleção, ou de forma heterogênea, dependendo inteiramente do tipo, qualidade e usabilidade das fontes usadas: hubs, mecanismo de busca ou fontes externas. Por fim, a fase de filtragem – que visa reduzir o espaço aberto pela fase de descoberta até os limites definidos pela política de seleção. A filtragem pode ser feita manualmente ou automaticamente, sendo que a manual é necessária quando os critérios usados para a seleção necessitam de interpretação humana. É importante definir o nível necessário de filtragem manual para otimizar tempo e recursos. (MASANÉS, 2006).

Milligan, Ruest e Lin (2016), entendem que a curadoria manual beneficia partes impopulares, candidatos marginais, entre outros que se contrapõem às seleções de massa. Neste sentido, entendemos a importância dessas práticas que, embora resultem em maiores esforços e custos mais altos, definem e viabilizam uma preservação mais ampla e igualitária de fontes de informação produzidas e disponibilizadas na *Web*.

A seleção sempre irá atender as necessidades e interesses da instituição, podendo os critérios serem baseados pelo domínio (ex.: .gov⁶ ou .fr⁷), tópico ou evento, tipo de mídia e gênero. Muitos países optam por arquivar a *Web* com base no domínio do país. Outros, como a Biblioteca do Congresso, criaram coleções baseadas em eventos, como os arquivos da *Web* de 11 de setembro de 2001 e os arquivos da *Web* das eleições. Já a seleção baseada em tipo de mídia inclui ou exclui determinados tipos de mídia. Outras formas de seleção por evento têm

⁶ Usado em endereço de *Internet* pertence a uma organização governamental.

⁷ Usado em endereço de *Internet* delegado à França.

surgido com o uso de redes sociais, sendo o *Twitter* o mais bem-sucedido a partir do rastreamento de *hashtags* específicas. (NIU, 2012; MILLIGAN, RUEST E LIN, 2016).

2.3.2 Tecnologias e Infraestruturas de Arquivamento da *Web*

Todos esses esforços de arquivamento da *Web* necessitam de tecnologias em constante atualização para tornar viável, em termos de custos e tempo – a coleta, a pesquisa e a preservação das informações digitais. Na maior parte dos casos, as iniciativas vêm trabalhando em projetos colaborativos no desenvolvimento desses recursos, que – em sua maioria, são disponibilizados como *open source*, podendo ser reutilizados por toda a comunidade de desenvolvedores, equipes e iniciativas de arquivamento da *Web*.

As principais tecnologias utilizadas nos projetos apresentam grande aderência das iniciativas, ou seja, embora existam diversos recursos tecnológicos, a maior parte opta pelas mesmas opções. Por essa razão, apresentamos as tecnologias que são referências nos arquivos da *Web*.

Para coletar os endereços da *Web*, são utilizados os recursos que baixam um *site* da *Web* para um diretório local, criando recursivamente todos os diretórios, obtendo HTML, imagens e outros arquivos do servidor para o computador de destino. O mais utilizado pelas iniciativas em 2010, de acordo com Costa, Gomes e Silva (2017), foi o *Heritrix* (2018), que é um rastreador de arquivamento da *Web*, extensível, de qualidade de arquivamento e de código aberto fornecido pelo *Internet Archive*. Em segundo lugar, aparece o *HTTrack* (2018), disponível gratuitamente com a Licença Pública Geral de *Software Livre*.

Junto com o recurso de coletar, a função de pesquisar os arquivos da *Web* compõe a engrenagem que move o arquivamento da *Web* em si e sem esses recursos, disponíveis pelo acesso aberto, muitas iniciativas provavelmente não teriam prosperado em seus projetos.

Apresentamos as duas ferramentas de pesquisa que mais são utilizadas, cada uma com suas diferenças. A proposta do *Wayback Machine*, criada pelo *Internet Archive*, consiste na pesquisa, exclusivamente, através do URL. Por outro lado, o *NutchWAX* permite realizar a pesquisa pelo texto completo. Como as duas estratégias são diferentes, ambas são muitas vezes utilizadas em conjunto. Entretanto, algumas iniciativas afirmaram que o *NutchWAX*, se mostrou difícil de implementar e que o desempenho era insatisfatório. (COSTA, GOMES E SILVA, 2017).

Por fim, das tecnologias mais importantes para o arquivamento da *Web*, temos os serviços de arquivamento da *Web* (WAS) – que consiste na incumbência de uma iniciativa, mediante contratação, realizar o arquivamento da *Web* para outra instituição com base em uma demanda apresentada. Dados de 2013 apresentados por Costa, Gomes e Silva (2017), mostraram que o *Archive-It* possuía uma aderência de 70% das iniciativas que utilizavam o WAS, já a Biblioteca Digital da Califórnia, que também possui um WAS, contou com 17% das iniciativas.

Além dos recursos tecnológicos, não podemos esquecer a contribuição dos recursos humanos envolvidos nos projetos e iniciativas de arquivamento da *Web*. Summers e Punzalan (2017) citam a necessidade de se compreender como os arquivos da *Web* são construídos, afirmando que eles são sistemas sociotécnicos nos quais os arquivistas colaboram com agentes automatizados. Os autores explicam que o estudo de algoritmos e seu papel na construção de *bots*, plataformas e infraestruturas consideram não apenas o funcionamento específico do próprio código, mas também como os algoritmos são componentes de grandes conjuntos sociotécnicos, onde o arquivamento da *Web* é realizado por agentes automatizados, ou rastreadores, que são direcionados em graus variados pelas interações das pessoas com o software de arquivamento da *Web*.

Reforçando as considerações dos autores citados, existem diversos grupos de profissionais interdisciplinares que atuam em conjunto nos projetos de arquivamento da *Web*, como é o caso do *WebArchivist.org* – um grupo de pesquisa da Universidade de Washington que desenvolve sistemas para identificar, colecionar, catalogar e analisar arquivos em grande escala de objetos da *Web*. (SCHNEIDER et al, 2003).

Em um estudo realizado a partir de dados coletados em 2010 e 2014, Costa, Gomes e Silva (2017) identificaram que as equipes de arquivamento da *Web* eram tipicamente pequenas, e o tempo de trabalho variava, sendo menor o número de pessoas que trabalhavam em tempo integral. Além dos arquivistas da *Web*, as equipes eram compostas principalmente de bibliotecários e engenheiros de computação. Os dados mostraram que houve um aumento de 278 para 305 pessoas trabalhando nesta área em 2014. Entretanto, é possível que pequeno tamanho das equipes seja em função do grande uso de serviços de arquivamento da *Web* (WAS) pelas iniciativas ao invés de executar seu próprio sistema de arquivamento na *Web*.

Por último, descrevemos brevemente os recursos que vêm sendo utilizados para o arquivamento de mídias sociais, as quais destacaremos o *Twitter* e o *Facebook*. É necessário ressaltar, contudo, que a rede social mais utilizada como fonte de arquivamento da *Web* tem

sido o *Twitter*, que disponibiliza para desenvolvedores APIs específicos para a coleta. Com isso, o arquivamento da *Web* vem permitindo que pesquisadores coletem textos e metadados relevantes da *Internet* e compilem em arquivos de dados para vários tipos de análise. Em pesquisas de mídia social, o método tem sido usado principalmente para estudos quantitativos de uso de mídias sociais. O arquivamento na *Web* também é adequado para estudos qualitativos que abordam questões-chave de pesquisa de audiência em um determinado período e em um determinado contexto. Os arquivos de mídia social, especialmente aqueles gerados por uma estratégia de arquivamento seletiva e profunda, podem reconstruir e tornar visível uma cronologia na comunicação que seria quase impossível acessar e estudar de outra forma. (LOMBORG, 2012).

A partir da análise de seis diferentes conjuntos de dados centralizados em eventos de recursos compartilhados nas mídias sociais no período de junho de 2009 a março de 2012, Salaheldeen e Nelson (2012), encontraram cerca de 11% perdidos e 20% arquivados após apenas um ano e uma média de 27% perdidos e 41% arquivados após dois anos e meio. Concluindo que, após o primeiro ano de publicação, quase 11% dos recursos compartilhados serão perdidos e, depois disso, continuará a se perder 0,02% por dia.

O *Internet.It* desenvolveu recursos de arquivamento de tuítes (*Twitter*) e páginas, perfis e/ou grupos do *Facebook*. É possível, a partir do recurso, adicionar “sementes” destas redes sociais a coleções já existentes, sendo instruído aos usuários do serviço que não se tente rastrear e arquivar todo o *Facebook*, mas se defina especificamente qual página, perfil ou grupo que se deseja coletar. O serviço possui algumas limitações como a impossibilidade de expandir as seções de comentários das postagens para ver mais comentários do que a tecnologia de rastreamento capturou, bem como a impossibilidade de expandir as postagens quando clicadas e mesmo que cada postagem mostre quantas vezes ela foi “curtida”, não poderá se navegar pelos nomes específicos dos usuários do *Facebook* que “curtiram” qualquer postagem. (SUPPORT ARCHIVE-IT, 2018).

2.3.3 Preservação Digital

O grande desafio da preservação digital, segundo Day (2006), condiciona-se a rápida obsolescência técnica, e, assim, as atividades tradicionais de preservação (manutenção e conservação) de objetos não são mais eficazes para apoiar a preservação de longo prazo da informação digital. Em vez disso, as organizações precisam desenvolver processos e sistemas

que possam ser implementados por um longo período, ou seja, é preciso desenvolver algum tipo de repositório ou arquivo.

Um dos grandes problemas da preservação, é garantir a estabilidade da mídia na qual a informação digital é armazenada. Para isso, o conteúdo precisa ser copiado periodicamente para novas mídias ou novos tipos de mídia. Esse processo é chamado de "atualização" e é uma das atividades associadas à boa prática de gerenciamento de dados, como a criação de backups regulares. O outro problema, mais difícil de resolver, consiste na obsolescência tecnológica dos tipos de hardware, software e mídia. No geral, as estratégias para lidar com os problemas de obsolescência técnica incluem a migração periódica de objetos para novos formatos e tentativas de preservar ou emular a tecnologia. (DAY, 2006).

A partir da compreensão dos desafios da preservação, Day (2006), descreve que é possível buscar no Modelo OAIS a identificação das principais funções que precisam ser realizadas pelos serviços de preservação e definir um modelo de informações para os objetos por eles detidos.

O modelo de referência OAIS (*Open Archival Information System*), segundo descreve Ferreira (2006), é um modelo conceitual voltado para a identificação dos componentes funcionais que deverão fazer parte de um sistema de informação dedicado à preservação digital, descrevendo ainda as interfaces internas e externas do sistema e os objetos de informação que são manipulados no seu interior. O modelo, que nasceu na década de 90, foi aprovado como uma norma internacional em 2003 – ISO Standard 14721:2003.

O esquema conceitual OAIS possui quatro entidades: produtores, consumidores, administração e o arquivo propriamente dito. (MÁRDERO ARELLANO, 2008; SOUZA ET AL, 2012).

Descrevendo brevemente sobre o esquema OAIS, a partir das afirmações de Ferreira (2006), iniciamos pelo Produtor que é uma entidade externa ao repositório responsável pela submissão de material (Pacote de Submissão de Informação – PIS). Já o processo de submissão ou incorporação é nomeado Ingestão e o repositório é responsável por garantir a integridade da informação recebida. Na fase de Ingestão também é produzida toda a Informação Descritiva que irá suportar a descoberta e localização do material depositado, sendo esta informação (ou meta-informação) armazenada e gerida pelo componente Gestão de Dados, sendo o material conservado no Repositório de Dados. O componente de ingestão é a interface entre o arquivo OAIS e os respectivos produtores de informação.

O componente Planejamento de Preservação é responsável pela definição de políticas de preservação, sendo incumbido de monitorar o ambiente externo ao repositório e desencadear eventos de preservação sempre que necessário. Sendo comum a este componente o controle das tendências comportamentais da sua comunidade de interesse ou a identificação dos formatos que se encontram na iminência de se tornar obsoletos. Por sua vez, o componente Acesso atua como interface entre o repositório e a sua comunidade de interesse, o conjunto de potenciais Consumidores do material preservado. Os pacotes que são entregues ao consumidor assumem a forma de Pacotes de Informação de Disseminação – PID, que são diferentes dos Pacotes de Informação de Arquivo, pois a informação que é entregue ao consumidor poderá ser apenas um subconjunto da informação arquivada ou até uma versão transformada da mesma. Por último, o componente Administração é responsável pela manutenção, pela parametrização e pela monitorização dos processos desencadeados no interior do repositório, interagindo com todos os restantes de modo a assegurar o correto funcionamento do mesmo. (FERREIRA, 2006).

Pavão, Caregnato e Rocha (2016) complementam que a preservação digital precisa lidar com os problemas de “formatos, curadoria, modelos de referência, autenticidade, políticas e planos de preservação, ferramentas, etc., enquanto que os implementadores necessitam de uma visão integrada, geral, contextualizada e aplicável”. (PAVÃO, CAREGNATO E ROCHA, 2016). Assim, os autores trabalham com um modelo conceitual de preservação e curadoria digital, “onde a curadoria digital é evidenciada na especificação dos serviços e do ciclo de vida, e a preservação digital é destacada nos recursos de preservação usados para a realização de serviços, como políticas, estratégias, planos, objeto, registro e ambiente.”. (PAVÃO, CAREGNATO E ROCHA, 2016). O modelo descrito pelos autores contempla serviços, política de preservação, estratégia de preservação, plano de preservação, objeto de preservação, ambiente, eventos e agentes de preservação. Havendo a especificação de subclasses de serviços, como planejamento, ingestão e armazenamento, e de subclasses de ambientes, como software, hardware, comunidade e produtor.

2.3.4 Arquivamento da *Web* das Eleições

A *Web* se tornou um canal de comunicação importante, desempenhando um papel fundamental no fornecimento de informações aos eleitores e membros de partidos – não apenas durante, mas também fora da temporada eleitoral. Os partidos políticos também estão cada vez mais inclinados a fornecer certas informações apenas em formato digital, não apenas porque é

muito mais rápido fazê-lo dessa maneira, mas também porque é muito mais barato. Outras razões que impulsionam essa comunicação são as possibilidades de recrutar trabalhadores de campanha e obter apoio financeiro. Além disso, a maioria dos partidos buscam através de novas formas de participação digital envolver mais as pessoas, especialmente os jovens. (VOERMAN et al, 2002).

Com o aumento do uso da *Web* como plataforma de comunicação eleitoral, surgem as mais variadas demandas de pesquisa nesses canais e fontes de informação. Percebendo a curto tempo de disponibilidade dessas fontes, diversas iniciativas internacionais têm trabalhado na criação de coleções temáticas ou por evento sobre as eleições.

A criação de uma coleção temática ou por evento requer a identificação e captura de objetos da *Web* através de um conjunto de URLs que se acredita serem relevantes para um tema ou tópico específico. Este conjunto de URLs selecionadas, que representam os *sites* ou páginas de interesse, são utilizados como “sementes” e irão rastrear, em uma periodicidade estabelecida e com regras claramente especificadas, as páginas e objetos vinculados. Essas coleções temáticas podem ser comparadas a atividades de rastreamento de base ampla, onde um programa de rastreamento da *Web* iniciará sua atividade de arquivamento a partir de um conjunto de URLs que não têm um relacionamento de conteúdo específico entre si. (SCHNEIDER et al, 2003).

A título de exemplo, trazemos a Biblioteca do Congresso dos Estados Unidos (*Library of Congress*) que criou diversas coleções sobre política em geral. A iniciativa conta com a coleção *United States Elections Web Archive* onde são arquivados, desde 2000, os *sites* associados às eleições presidenciais, parlamentares e governamentais semanalmente durante as temporadas de eleições. A Biblioteca do Congresso também criou uma coleção temática intitulada *Brazilian Presidential Election 2010 Web Archive*, composta por 48 *Websites* de partidos políticos nacionais e de candidatos das eleições presidenciais brasileiras de 2010. (LIBRARY OF CONGRESS, 2018).

O serviço do *Archive-It* também é responsável por diversas coleções temáticas envolvendo as eleições de inúmeros países. No entanto, esse serviço não é gratuito e sempre acontece por demanda de alguma instituição que contrata o *Internet Archive*, criador do serviço. Como exemplo, temos a Biblioteca Nacional da Austrália (*National Library of Australia*) que contratou o serviço do *Archive-It* para coletar e arquivar *sites* para a *Asian Collections*, fazendo parte da coleção os *sites* relacionados às eleições gerais da Tailândia em 2007; *sites* do governo, de partidos políticos e de mídia relacionados à eleição nacional do Camboja em 2008; *sites*

relacionados às eleições parlamentares de Papua Nova Guiné 2007; e *sites* relacionados com as eleições presidenciais e parlamentares de Timor-Leste em 2007, este incluindo educação para eleitores e material de propaganda política. (CROOK, 2009).

Outro projeto envolvendo coleção temática foi desenvolvido pela Biblioteca e Arquivos do Canadá (*Library and Archives of Canada*), onde a proposta consistiu em coletar todos os tuítes públicos postados na rede social *Twitter* com a hashtag #elxn42. A hashtag se referia à eleição como a 42ª eleição federal geral do Canadá e os dados começaram a ser coletados em 03 de agosto de 2015, cessando em 5 de novembro de 2015, um dia depois que Justin Trudeau foi empossado como o 42º primeiro-ministro do Canadá. Para isto, foram utilizados APIs, ferramentas de linha de comando e utilitários disponíveis, como *twarc*, *twarcreport*, desenvolvidos para trabalhar na plataforma do *Twitter* e que estão disponíveis gratuitamente para uso de desenvolvedores. Está disponível no Scholars Portal Dataverse (2018) um arquivo “.txt” com as IDs coletada, sendo necessário, para acesso as postagens, a instalação do *twarc* de Ed Summers (este disponível no GitHub⁸). (RUEST; MILLIGAN, 2016).

A *UCLA Online Campaign Literature Archive* é outro exemplo de iniciativa que tem coletado *Websites* das eleições de Los Angeles e Califórnia desde 1998, no entanto a coleta era feita manualmente. E, em 2012, a iniciativa percebeu que era necessário encontrar um modelo mais sustentável para continuar a arquivar *sites* de campanha. O objetivo era encontrar uma ferramenta automatizada que pudesse corresponder às capturas de alta qualidade produzidas pelo processo de captura manual de trabalho intensivo existente no *Archive*. A ferramenta então escolhida foi o *Web Archiving Service* (WAS) da Biblioteca Digital da Califórnia. (GRAY; MARTIN, 2013).

Contudo, como afirmam Pop, Vasile e Masanès (2010), é difícil projetar uma solução geral para lidar com todos os *sites* que hospedam conteúdo de vídeo e a técnica de coleta deve ser adaptada para cada caso particular, bem como o esforço de engenharia de rastreamento necessário para adaptar as ferramentas geralmente depende da complexidade do *site*. E quando nos referimos a plataformas privadas, como o *Facebook*, este desafio se mostra mais complexo e exige um conjunto de habilidades, as quais apenas a engenharia de rastreamento não é capaz de sanar sozinha.

⁸ Plataforma de hospedagem de código para controle de versão e colaboração. (FRIEDMAN, 2016).

3 METODOLOGIA

A metodológica consistiu em uma abordagem exploratória-descritiva de um estudo quanti-qualitativo, onde a coleta de vídeos envolveu a utilização de recursos informatizados de linguagem de programação e estruturação em base de dados. Esta abordagem se justifica a partir de Gil (2002) que descreve a análise exploratória como usual no desenvolvimento de ideias e hipóteses que poderão ser testadas em estudos posteriores, possibilitando maior flexibilidade que o dos outros tipos de pesquisa. E a pesquisa descritiva como a descrição das características do fenômeno estudado, sendo que a preocupação do pesquisador é a de descrever com precisão essas características, utilizando instrumentos que conduzem a resultados de natureza quantitativa. Contudo, também realizou-se inferências a partir das análises documentais das publicações audiovisuais realizadas pelos candidatos em suas fanpages.

Esta pesquisa foi estruturada em **cinco etapas**, onde a **primeira etapa** consistiu na identificação de todos os candidatos à Presidência e suas *fanpages* oficiais na rede social *Facebook*.

Após a identificação das *fanpages* dos candidatos, a **segunda etapa**, compreendeu no desenvolvimento do código das páginas de interface *Web* (inclusão e consulta) e estruturação da base de dados MySQL⁹.

Para iniciar a população da base de dados, foram coletados os nomes de referência dos candidatos no *Facebook* (ex. @ad.alvarodias) e, utilizando o recurso *Find your Facebook ID* (2018), a partir desse endereço coletamos o nº de identificação (ID) respectivo de cada candidato utilizada para popular a primeira tabela e atuar como elo nas demais tabelas.

Foram criadas três tabelas no MySQL, a primeira, denominada “tbl_alvo”, que listou todos os candidatos à Presidência; a segunda, denominada “tbl_archive”, armazenou todos os registros coletados e seus respectivos dados; por fim, a terceira tabela, denominada “tbl_log”, armazenou os dados e metadados atualizados diariamente no 1º turno das eleições. A estrutura das tabelas pode ser analisada no **Apêndice A, Apêndice B e Apêndice C**.

O processo de desenvolvimento das rotinas de *scripts*¹⁰ passou por várias alterações, uma vez que a maneira ideal e correta de se buscar dados de dentro de uma página, nesse caso mais específico do *Facebook*, seria através do uso de *Application Programming Interface*¹¹ (APIs) disponibilizadas para tal. Porém, depois do escândalo da *Cambridge Analytica*, onde

⁹ Versão *Software Livre*.

¹⁰ Linguagem de programação.

¹¹ Conjunto de rotinas parametrizadas que possibilitam integrações entre diferentes aplicações.

supostamente houve uso de dados pessoais que teriam sido usados pela campanha de Donald Trump, a plataforma do *Facebook* restringiu ao máximo o uso das APIs para o público “doméstico”, deixando apenas abertas possibilidades para grandes agências de notícias, ou empresas devidamente autorizadas, com ampla restrição ao tipo de dados fornecidos.

Chegou-se a construir um protótipo de *bot* que, através das APIs oficiais, coletava dados de páginas públicas, incluindo *feeds*, reações, vídeos e demais informações. Esse *bot* funcionou corretamente até final de julho, quando o *Facebook* mudou as regras de liberações e praticamente bloqueou esse tipo de aplicação “doméstica”. Mesmo com o envio de solicitação de aprovação (o produto passa por análises para poder ser ou não liberado), o produto foi negado de ir ao ar por não estar vinculado a agências de notícias ou empresa devidamente autorizada.

Passamos, então, a usar a técnica de *scraping*, que consiste em criar *scripts* que analisam as *strings*¹² retornadas das páginas em análise. Estes *scripts* utilizados estão disponíveis no **Apêndice D, Apêndice E e Apêndice F**.

Abaixo descrevemos o funcionamento do processo de *scraping*, que consistiu nas seguintes etapas:

1. Identificação da página **pública** na qual se deseja buscar informações;
2. Realização de um “*dump*”¹³ dessa página, usando técnicas de pesquisa de programação, na qual se busca internamente por palavras chave que nos darão pistas do que se deseja coletar. Em nosso caso, buscamos por metadados que informaram os nomes dos arquivos mp4 (vídeos), o assunto da postagem (*description*), número de curtidas (*likecount*), número de comentários (*commentcount*), número de visualizações (*viewcount*), número de compartilhamentos (*sharecount*), duração do vídeo (*duration*), proprietário do link (*ownerid*), dentre outras informações. A própria definição destes metadados foi realizada a partir da análise do código-fonte durante uma navegação em uma das páginas dos candidatos através do recurso “exibir código fonte da página”, onde todo o conteúdo que é postado do servidor da página ao navegador nos passa informação no estado bruto. Cabe salientar que ao fazermos isso, não estamos vendo o código-fonte propriamente dito, e sim o resultado que esse código produziu e foi exposto na *Web*;
3. Criação de uma listagem em um banco de dados, hospedado em um servidor devidamente configurado, que armazenou cada link a ser investigado ao longo do

¹² Sequência de caracteres usados para representar uma palavra, frase ou textos de um programa.

¹³ Exportação de todos os dados.

tempo. Esta listagem foi constituída de uma tabela que armazenou o link, o nome do candidato o ID do candidato e o nome da pasta de *download* dos registros. Esta coleta foi feita diariamente de forma manual, com base em observações das postagens dos candidatos nas suas páginas, usando nossa interface *Web*, que foi desenvolvida especificamente para isso;

4. A realização da coleta inicial consistiu no arquivamento imediato do vídeo ou *live* no servidor, armazenando também dados como data da coleta, tamanho e duração do vídeo, se era *live* ou vídeo comum, reações daquele momento, bem como metadados contidos no frame do *Facebook*;
5. Por fim, as coletas posteriores, foram feitas pelo mecanismo automático, programado para pesquisar novamente a base a cada seis horas, limitando-se a coletar novos dados sobre as reações referentes aos vídeos e *lives*.

Com isso, formou-se um banco de dados ricos em informações, que nos permitem fazer diversos tipos de análise, tais como quais os vídeos/*lives* com maior audiência, seus conteúdos associados e quais permaneceram no ar, quais foram removidos etc.

Para viabilizar a análise estatística dos dados, os arquivos coletados foram hospedados em um *Data Center*¹⁴ particular, contratado à *SeedHost*¹⁵, a qual hospeda os dados na Holanda.

O período definido para a realização da coleta foi o fixado para a realização da propaganda eleitoral estipulado pelo TSE, do dia 16 de agosto a 04 de outubro, compreendendo o 1º turno das eleições, e do dia 08 de outubro ao dia 26 de outubro para os dois candidatos do 2º turno.

A **terceira etapa** da pesquisa abarcou a identificação da persistência dos vídeos coletados na fonte de origem, ou seja, nas *fanpages* dos candidatos. Para esta análise foi realizada a inclusão do campo “ativo na fonte” na página de consulta, onde a regra de população na base de dados foi definida como “S” (sim) e quando identificada, manualmente, a indisponibilidade do conteúdo, foi alterada para “N” (não). Esta identificação foi realizada logo após o fim das eleições, nos dias 15, 16 e 17 de novembro de 2018.

Na **quarta etapa** foi realizada uma busca nas plataformas do *Internet Archive*, do *Archive-It* e do *Archive.is* a fim de identificar possíveis conteúdos preservados parcialmente dos *sites* e redes sociais dos candidatos no período de campanha eleitoral a que este trabalho se

¹⁴ Centro de processamento de dados.

¹⁵ Empresa provedora de serviços de hospedagem da Europa.

refere. Para esta busca foram utilizados os endereços disponíveis na *Web* dos *sites* dos candidatos, na ausência destes foram considerados os *sites* do partido, bem como os endereços de mídias sociais, tais como *Facebook*, *Twitter*, *Youtube* e *Instagram*.

A **quinta etapa**, como resultado dos recursos desenvolvidos para esta pesquisa, gerou uma interface de consulta das informações coletadas de forma independente da plataforma *Facebook* que, após a conclusão da pesquisa, será migrado para a base de dados do projeto AWEB (Arquivamento das Eleições Brasileiras), em desenvolvimento pelo Núcleo de Pesquisa em Arquivamento da *Web* e Preservação Digital - NUAWEB (2018).

O NUAWEB é um Grupo de Pesquisa do Programa de Pós-Graduação da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul – UFRGS, e é formado por professores, alunos de pós-graduação e de graduação. Atualmente, o NUAWEB possui duas linhas de pesquisa, Linha 1: Políticas e tecnologias de arquivamento da *Web* e preservação digital e Linha 2: Ética da informação aplicada ao arquivamento da *Web* e preservação digital. (NUAWEB, 2018).

A análise e interpretação dos dados ocorreu em dois macroníveis: o primeiro composto por três subníveis e o segundo por um subnível.

Assim, temos no macronível A:

A1) Identificação das informações constantes nas *fanpages* dos candidatos, como número de seguidores e número de curtidas da *fanpage*, os avatares de apresentação das *fanpages* e se as *fanpages* foram verificadas pelo *Facebook*;

A2) Análise estatística e interpretação dos dados coletados no primeiro e segundo turnos por tipo de registro (vídeo ou *Live*), bem como o comparativo dos dois turnos, quantitativo diário de publicações no primeiro turno, comparativo de número de publicações no primeiro e segundo turnos dos candidatos Fernando Haddad e Jair Bolsonaro, análise do tempo total de exposição dos candidatos em registros audiovisuais no Facebook durante a campanha em contrapartida ao tempo previsto pelo TSE para os blocos em TV aberta;

A3) Análise e interpretação da disponibilidade do formato original dos registros audiovisuais publicados pelos candidatos, apresentação de dados estatísticos, apresentação de exemplos de indisponibilidade.

E no macronível B, analisamos a incidência de preservação dos conteúdos publicados pelos candidatos durante a campanha eleitoral em suas mídias sociais e *sites*. Para isto, consideramos três plataformas que preservam conteúdos brasileiros: *Internet Archive*, *Archive-It* e *Archive.is*.

4 OS PRESIDENCIÁVEIS E AS PUBLICAÇÕES EM MÍDIAS SOCIAIS

Apresentamos aqui a análise e interpretação dos dados, que foi estruturada a partir dos objetivos específicos, delimitando os resultados das cinco etapas da pesquisa.

4.1 Os Candidatos das Eleições Presidenciais de 2018

O cargo de Presidente da República, como explica Blume (2018), pode ser ocupado por brasileiros nascidos no Brasil ou filho de brasileiro, com, pelo menos, 35 anos de idade, filiados, no mínimo, a seis meses antes da data da eleição a um partido político e devem cumprir os critérios de elegibilidade previstos na Ficha Limpa¹⁶ (dentre os quais, não ser condenado em segunda instância na Justiça, por órgão colegiado).

Embora a campanha eleitoral brasileira tenha iniciado no dia 16 de agosto a Justiça Eleitoral julgou até o dia 17 de setembro todos os pedidos de registro de candidatos que concorreram ao pleito de 2018. (TRIBUNAL SUPERIOR ELEITORAL, 2018a). Em decorrência disso, o candidato à Presidência do Partido dos Trabalhadores (PT), até 17 de setembro, era o ex-presidente Luiz Inácio Lula da Silva (Lula) e o candidato a vice-presidente era Fernando Haddad. Todavia, a candidatura de Lula não foi aceita pelo TSE, pelo fato de o ex-presidente ter sido condenado em segunda instância e não cumprir os critérios exigidos na Lei da Ficha Limpa. Com isso, houve a troca de candidatos do PT – passando a ser o candidato à Presidência Fernando Haddad e, a vice-Presidência, Manuela D'Ávila.

No intuito de preservar o contexto peculiar das eleições presidenciais de 2018, para esta pesquisa mantivemos a coleta das informações e dos dados publicados na *fanpage* do ex-presidente Lula durante todo o 1º turno das eleições.

As eleições presidenciais de 2018 contaram com o registro de 13 candidatos, sendo onze homens e duas mulheres. E, segundo Moraes (2018), este número foi o maior desde as eleições de 1989.

Identificamos que os 13 candidatos estavam presentes no *Facebook* através de *fanpages*. Também se observou que o grau de atuação na plataforma variou bastante de candidato para candidato. Mas, no geral, em todas as 13 *fanpages* foram realizadas publicações de conteúdos audiovisuais, de vídeos e *lives*. Durante a coleta, houve a alteração do ID da página do candidato Guilherme Boulos, a qual iniciou como @guilhermeboulos.oficial e, no dia 28 de agosto de

¹⁶ Lei Complementar nº 135, de 4 de junho de 2010.

2018, passou para @guilherme.boulos. Esta alteração ocorreu em função de um acordo realizado com o proprietário da ID @guilherme.boulos.

Para esta pesquisa, não foram identificadas as *fanpages* dos candidatos a vice-presidentes, pois essas informações não compunham nosso objetivo de análise. Assim sendo, listamos na **Tabela 1** os nomes de todos os candidatos à Presidência, a identificação de suas páginas/*fanpages* no *Facebook*, o partido ao qual são filiados, os vice-presidentes da chapa e seus partidos de filiação.

Tabela 1 - Lista dos candidatos à Presidência - informações gerais.



| CANDIDATOS À PRESIDÊNCIA | | | | |
|--------------------------|------------------------|----------|-----------------------------|----------|
| CANDIDATO | ID FACEBOOK | PARTIDO | VICE | PARTIDO |
| Álvaro Dias | @ad.alvarodias | Podemos | Paulo Rabello de Castro | PSC |
| Cabo Daciolo | @deputadocabodaciolo | Patriota | Suelene Balduino Nascimento | Patriota |
| Ciro Gomes | @cirogomesoficial | PDT | Kátia Abreu | PDT |
| Fernando Haddad | @fernandohaddad | PT | Manuela D'Ávila | PCdoB |
| Geraldo Alckmin | @geraldoalckmin | PSDB | Ana Amélia | PP |
| Guilherme Boulos | @guilherme.boulos | PSOL | Sônia Guajajara | PSOL |
| Henrique Meirelles | @hmeirellesoficial | MDB | Germano Rigotto | MDB |
| Jair Bolsonaro | @jairmessias.bolsonaro | PSL | Hamilton Mourão | PRTB |
| João Amoêdo | @JoaoAmoedoNOVO | NOVO | Christian Lohbauer | NOVO |
| João Goulart Filho | @joaogoulart54 | PPL | Léo Alves | PPL |
| José Maria Eymael | @eymaelOficial | DC | Helvio Costa | DC |
| Marina Silva | @marinasilva.oficial | Rede | Eduardo Jorge | PV |
| Vera Lúcia | @verapstu | PSTU | Hertz Dias | PSTU |
| Lula | @Lula | PT | | |

Fonte: Elaborado pela autora, 2019.

Destes dados também observamos a ocorrência de cinco chapas oriundas de coligação partidária, prática usual na política brasileira. Atualmente, segundo o Tribunal Superior Eleitoral (2018d), o Brasil possui 35 partidos políticos e, das 13 chapas, observamos a presença de 18 partidos na composição.

Já na **Tabela 2**, observamos o número de curtidas da *fanpage* de cada candidato, bem como o número de seguidores. Estes dados foram coletados no dia 26 de novembro de 2018., sendo que a maior parte dos candidatos possuía um número um pouco maior de seguidores, com exceção de Geraldo Alckmin e Marina Silva – que obtiveram um decréscimo de seguidores em relação ao número de pessoas que curtiram a página. Isto acontece porque, ao curtir uma página, um usuário-eleitor pode desabilitar a função “seguir” e, assim, não aparecerá em seu *feed* de notícias as postagens daquela página.

Tabela 2 - Número de curtidas das fanpages dos candidatos no *Facebook* e número de seguidores por ordem crescente de curtidas.

| |  pessoas que curtiram |  pessoas que estão seguindo |
|---------------------------|--|--|
| Jair Bolsonaro | 8.826.143 | 9.959.713 |
| João Amoêdo | 3.063.819 | 3.106.461 |
| Marina Silva | 2.393.948 | 2.310.556 |
| Fernando Haddad | 1.789.458 | 1.838.366 |
| Álvaro Dias | 1.220.048 | 1.225.425 |
| Geraldo Alckmin | 1.102.730 | 1.081.427 |
| Ciro Gomes | 770.586 | 799.891 |
| Guilherme Boulos | 688.000 | 732.821 |
| Cabo Daciolo | 484.199 | 533.531 |
| Henrique Meirelles | 249.146 | 252.616 |
| Vera Lúcia | 24.215 | 24.458 |
| José Maria Eymael | 23.729 | 24.289 |
| João Goulart Filho | 17.160 | 17.373 |
| Lula* | 3.908.032 | 4.012.639 |

* candidato não oficial

Fonte: Elaborado pela autora, 2019.

Cabe destacar que, das 13 *fanpages* encontradas, apenas uma não possuía a etiqueta de página Verificada, a qual significa que o *Facebook* confirmou que a página é autêntica para aquela figura pública. Esta verificação não acontece na página do candidato João Goulart Filho, como podemos observar na **Figura 3**.

Figura 3 - Avatar dos candidatos à Presidência em suas *fanpages*.



Fonte: Compilação da autora a partir de *Facebook*, 2019.

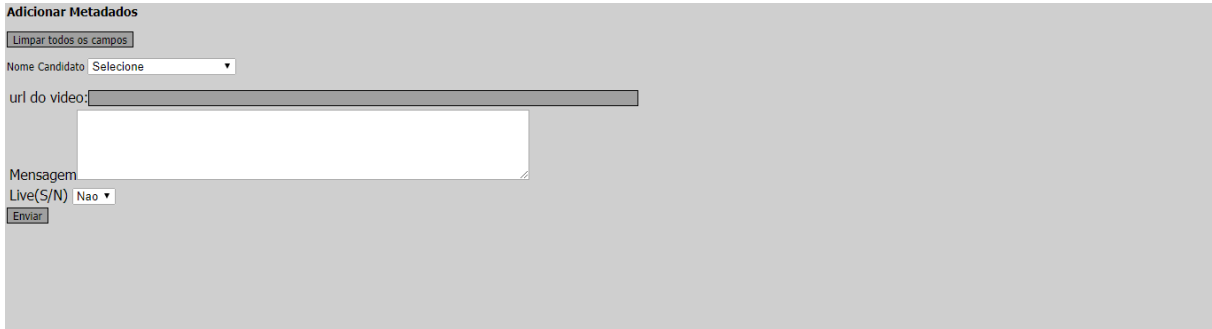
A aplicação do selo de verificação azul é um recurso disponível para marcas, organizações de mídia e figuras públicas. A qualificação para o selo leva em consideração a integridade da conta, conformidade com a política e interesse público. (FACEBOOK, 2018b).

4.2 Conteúdos Audiovisuais Coletados

Os objetos de nossa análise consistiram, primordialmente, na coleta dos conteúdos audiovisuais publicados como vídeos e *lives* pelos candidatos à Presidência. No entanto, a coleta não se limitou a estes conteúdos, reunindo também informações que contextualizam a publicação de forma mais rica, tais como a descrição que resume uma ideia central que será tratada no vídeo ou *live*, o tipo de conteúdo, se consistia em um vídeo convencional ou uma transmissão ao vivo, a data da publicação, o número de curtidas, comentários, compartilhamentos e visualizações que a publicação obteve até a data da coleta. Para realizar o conjunto desta coleta, foi necessário criar interfaces *Web* que possibilitaram a automatização do processo em nível micro, ou seja, um a um.

Como podemos observar na **Figura 4**, a inclusão de cada conteúdo audiovisual ocorreu através de uma interface *Web* onde, durante o 1º turno das eleições, era necessário inserir apenas o *URL* do vídeo disponível na página do candidato no *Facebook*. Já no 2º turno alguns recursos foram bloqueados pela plataforma, sendo preciso acrescentar – além do *URL* do vídeo, a descrição que acompanhava o vídeo e a definição do tipo de vídeo.

Figura 4 - Interface *Web* de inclusão dos registros audiovisuais na base de dados.



The screenshot shows a web form titled "Adicionar Metadados". At the top left, there is a button labeled "Limpar todos os campos". Below it is a dropdown menu for "Nome Candidato" with the text "Selecione". The main part of the form consists of a text input field labeled "url do video:". Below the URL field is a larger text area labeled "Mensagem". At the bottom left, there is a dropdown menu for "Live(S/N)" with the text "Nao" and a small downward arrow. At the bottom center, there is a button labeled "Enviar".

Fonte: Elaborado pela autora, 2019.

Ao clicar no botão “enviar” as informações eram transmitidas como parâmetros para outra página que executava o *script bot-scraping* (**Apêndice E**) para abrir o link internamente em uma versão de não usuário e executar uma “varredura” nos metadados disponíveis. Neste

processo, a primeira ação consistiu em identificar de qual candidato era o link e essa informação foi relacionada a partir do ID do candidato. Na sequência, foram coletados os dados: data de publicação, data de coleta, URL do vídeo para download (baixa resolução - SD), assunto, comprimento do vídeo em bytes, número de curtidas, número de comentários, número de compartilhamentos, número de visualizações, duração do vídeo. Estes dados “popularam” a tabela `tbl_archive` ao mesmo tempo que o comando executou o *download* do vídeo no *server* (servidor) e o direcionou para a pasta do candidato correspondente como mostra a **Figura 5**.

Figura 5 - Pastas dos candidatos com os registros audiovisuais no servidor.

The image shows two side-by-side screenshots of a remote file manager interface. The left screenshot shows a directory listing of candidate folders. The right screenshot shows the contents of the 'ad.alvarodias' folder, which contains a list of MP4 video files.

| Name | Size | Type | Date Modified | Permissions |
|----------------------|--------|-------------|------------------|-------------|
| ad.alvarodias | 20.480 | File folder | 10/10/2018 16:33 | drwxrwxrwx |
| cirogomesoficial | 12.288 | File folder | 06/10/2018 23:18 | drwxrwxrwx |
| deputadocabodaciolo | 4.096 | File folder | 10/10/2018 22:21 | drwxrwxrwx |
| eymaelOficial | 4.096 | File folder | 07/10/2018 10:32 | drwxrwxrwx |
| femandohaddad | 20.480 | File folder | 29/10/2018 14:55 | drwxrwxrwx |
| geraldoalcmin | 12.288 | File folder | 06/10/2018 23:21 | drwxrwxrwx |
| guilherme.boulos | 16.384 | File folder | 10/10/2018 17:00 | drwxrwxrwx |
| hmeirellesoficial | 12.288 | File folder | 10/10/2018 12:37 | drwxrwxrwx |
| jaimessias.bolsonaro | 12.288 | File folder | 02/11/2018 18:40 | drwxrwxrwx |
| JoaoAmoedoNOVO | 12.288 | File folder | 10/10/2018 17:15 | drwxrwxrwx |
| joaogoulart54 | 4.096 | File folder | 09/10/2018 19:19 | drwxrwxrwx |
| Lula | 20.480 | File folder | 10/10/2018 12:13 | drwxrwxrwx |
| marinasilva.oficial | 12.288 | File folder | 10/10/2018 17:04 | drwxrwxrwx |
| verapstu | 4.096 | File folder | 06/10/2018 12:18 | drwxrwxrwx |

| Name | Size | Type | Date Modified | Permission |
|----------------------|--------------|---------------|------------------|------------|
| 100164859999637.mp4 | 1.903.486 | MP4 Video ... | 10/09/2018 14:12 | -rwxr-- |
| 1012802272233920.mp4 | 1.448.691 | MP4 Video ... | 13/09/2018 12:27 | -rwxr-- |
| 1025685447612671.mp4 | 187.125.8... | MP4 Video ... | 11/09/2018 14:54 | -rwxr-- |
| 1029708383867061.mp4 | 51.919.148 | MP4 Video ... | 01/10/2018 22:15 | -rwxr-- |
| 1045769032260888.mp4 | 256.855.8... | MP4 Video ... | 14/09/2018 22:55 | -rwxr-- |
| 1050079635181767.mp4 | 586.647 | MP4 Video ... | 22/09/2018 22:07 | -rwxr-- |
| 1051368798358788.mp4 | 74.516.615 | MP4 Video ... | 18/08/2018 15:29 | -rwxr-- |
| 1051903398329509.mp4 | 604.650 | MP4 Video ... | 27/09/2018 16:07 | -rwxr-- |
| 1054093671430117.mp4 | 13.361.417 | MP4 Video ... | 03/09/2018 19:31 | -rwxr-- |
| 1058352654325417.mp4 | 80.904.989 | MP4 Video ... | 30/09/2018 16:32 | -rwxr-- |
| 1069556629872627.mp4 | 127.983.9... | MP4 Video ... | 27/09/2018 16:04 | -rwxr-- |
| 1069826496512833.mp4 | 1.095.995 | MP4 Video ... | 16/09/2018 15:00 | -rwxr-- |
| 1091793307654550.mp4 | 6.790.687 | MP4 Video ... | 20/09/2018 17:56 | -rwxr-- |
| 1098888573611664.mp4 | 1.059.145 | MP4 Video ... | 11/09/2018 23:35 | -rwxr-- |

Fonte: Elaborado pela autora, 2019.

A página seguinte, **Figura 6**, retornava com algumas informações que utilizamos para aferir a inclusão com sucesso do conjunto de dados, apresentando o nome do candidato o qual pertencia o link do vídeo coletado, a data de criação (da coleta), a mensagem que acompanhava o vídeo, o link do vídeo preservado, o *iframe* de visualização do vídeo na origem/*Facebook*, número de curtidas, número de comentários, número de compartilhamento e número de visualizações. Nesta página também foi feita a verificação do *download* do vídeo, onde clicava-se no link do vídeo para ver se ele abria no servidor de arquivamento.

Figura 6 - Interface Web de inclusão após coleta de vídeo, exemplo.

Nome Criado Mensagem Link_video Frame Atualizar

Ciro Gomes 26-08-2018 18:50

Ainda na década de 90 escrevi um livro no qual propus, entre outras coisas, um IVA (Imposto sobre Valor Agregado), que é um único imposto que substitui vários outros. As lideranças políticas de São Paulo sempre foram contra. Parece que o jogo mudou. ☺ Faça como os candidatos: siga o Ciro!

295385701256745.mp4

IVA - Imposto sobre valor agre...
Publicado por **Ciro Gomes**
10.565 visualizações

curtida: 1399
coment. 92
comp. 433
visual. 10565

Ainda na década de 90 escrevi um livro no qual propus, entre outras coisas, um IVA (Imposto sobre Valor Agregado), que é um único imposto que substitui vários outros. As lideranças políticas de São Paulo sempre foram contra. Parece que o jogo mudou. ☺ Faça como os candidatos: siga o Ciro!

Fonte: Elaborado pela autora, 2019.

Em 13 casos, o *download* do arquivo precisou ser feito manualmente com o auxílio do *plugin*¹⁷ Video DownloadHelper, disponível apenas no navegador do Firefox.

A coleta também foi realizada com o auxílio de uma planilha que serviu como controle manual dos arquivamentos feitos diariamente como podemos ver na **Tabela 3**.

Tabela 3 - Tabela de controle manual da coleta dos registros audiovisuais.

| | A | B | C | D | E | F | G | H | I | J | K | L |
|----|--------------------|--------------|-------------|--------|---------|---------------|-------------|--------------|--------------|-------------|--------|---------|
| 1 | Candidato | 16/ago | 17/ago | 18/ago | 19/ago | 20/ago | 21/ago | 22/ago | 23/ago | 24/ago | 25/ago | 26/ago |
| 2 | Eymael | 1 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| 3 | João Goulart Filho | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| 4 | Lula | 8 | 9 | 5 | 2 | 8 | 7 | 7 | 9 | 9 | 7 | 4 |
| 5 | Vera Lucia | 1 | 1 | 0 | 0 | 2 | 1 | 1 | 1 | 0 | 0 | 1 |
| 6 | Cabo Daciolo | 4 | 1 | 3 | 0 | 2 | 1 | 0 | 2 | 0 | 0 | 1 |
| 7 | Henrique Meirelles | 0 | 1 | 7 | 1 | 1 | 2 | 1 | 3 | 3 | 2 | 2 |
| 8 | Alvaro Dias | 7 | 5 | 5 | 2 | 9 | 5 | 4 | 6 | 12 | 8 | 4 |
| 9 | Ciro Gomes | 5 | 1 | 2 | 1 | 0 | 1 | 1 | 3 | 1 | 1 | 2 |
| 10 | Geraldo Alckmin | 2 | 6 | 3 | 3 | 2 | 0 | 3 | 2 | 3 | 2 | 2 |
| 11 | Guilherme Boulos | 9 | 16 | 7 | 4 | 5 | 8 | 6 | 6 | 6 | 4 | 3 |
| 12 | Jair Bolsonaro | 3 | 4 | 0 | 3 | 2 | 1 | 6 | 5 | 5 | 5 | 2 |
| 13 | Marina Silva | 2 | 1 | 4 | 1 | 0 | 3 | 0 | 4 | 0 | 2 | 2 |
| 14 | João Amoêdo | 7 | 1 | 4 | 4 | 4 | 5 | 2 | 2 | 4 | 2 | 4 |
| 15 | | | | | | | | | | | | |
| 16 | | 50 | 47 | 40 | 24 | 35 | 36 | 31 | 43 | 44 | 34 | 28 |
| 17 | | | | | | | | | | | | |
| 18 | | quinta-feira | sexta-feira | sábado | domingo | segunda-feira | terça-feira | quarta-feira | quinta-feira | sexta-feira | sábado | domingo |
| 19 | | | | | | | | | | | | |

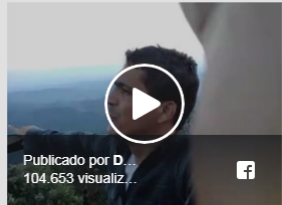
Fonte: Elaborado pela autora, 2019.

¹⁷ Componente de *software* que adiciona um recurso específico a um programa de computador existente.

Após efetuado o arquivamento, o conteúdo já estava disponível para ser pesquisado na página de consulta, **Figura 7**. Esta página, além de mostrar os dados coletados, possibilitava a atualização e acompanhamento das interações ocorridas na postagem durante o 1º turno.

Figura 7 - Página de consulta com exemplo.

Nome Candidato entre e Live?(S/N)

| Nome | Criado | Mensagem | Reacoes | Link Armazenado | Frame Facebook |
|--------------|------------------|----------|-------------------------------|--|---|
| Cabo Daciolo | 24-08-2018 10:54 | SÓ DEUS! | 4760 4021 3724 82917 | 585713815177077.mp4 live(S) |  Publicado por D... 104.653 visualiz... SÓ DEUS! 5,4 mil 4,4 mil 4,2 mil |

Fonte: Elaborado pela autora, 2019.

Nesta página de consulta, abaixo do nome do candidato aparecem dois botões que acionam telas secundárias em forma de *pop-up*, **Figura 8**. O primeiro botão viabiliza a alteração manual dos números das interações e a própria disponibilidade do conteúdo na fonte (recurso utilizado na análise de formato original). Já o segundo botão disponibiliza uma tabela com o histórico das interações coletadas automaticamente pelo *bot-scraping*.

Figura 8 - Telas de atualização e histórico.

Jair Messias Bolsonaro

Atualizacao - Dados

<https://www.facebook.com/ja...>

| | |
|----------------|-------------------------------------|
| curtida: | <input type="text" value="34944"/> |
| coment. | <input type="text" value="25038"/> |
| compart. | <input type="text" value="12745"/> |
| visual. | <input type="text" value="364206"/> |
| ativo na fonte | <input type="text" value="S"/> |

Atualizacao - Dados

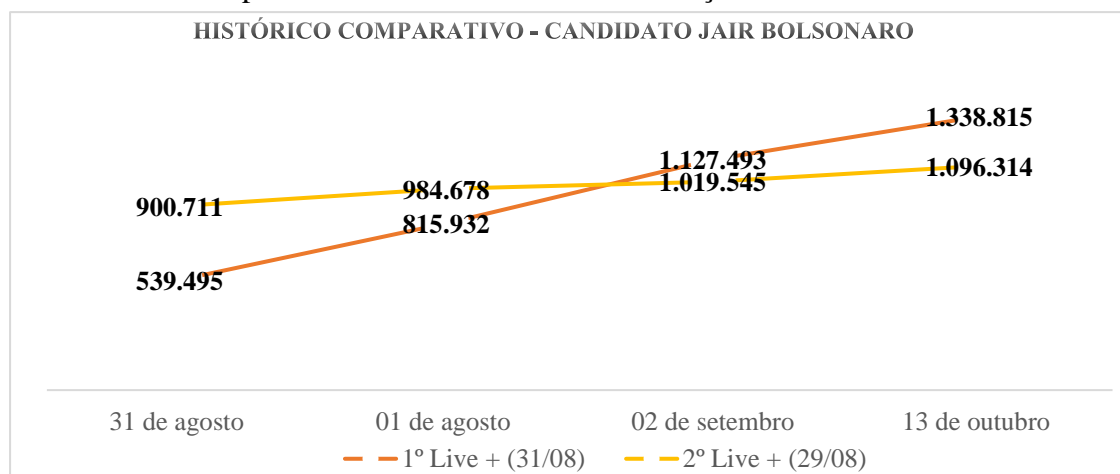
| Data | Curtidas | Comentarios | Compartilhamentos | Visualizacoes |
|------------------|----------|-------------|-------------------|---------------|
| 07-10-2018 18:48 | 46369 | 27659 | 15565 | 724646 |
| 07-10-2018 13:43 | 46372 | 27664 | 15566 | 724633 |
| 06-10-2018 19:05 | 46376 | 27664 | 15565 | 724592 |
| 06-10-2018 13:52 | 46379 | 27664 | 15566 | 724565 |
| 06-10-2018 07:44 | 46380 | 27664 | 15564 | 724544 |

Fonte: Elaborado pela autora, 2019.

O histórico das interações possibilita o acompanhamento do impacto e influência que uma determinada publicação obteve durante o período. A título de demonstração, **Gráfico 1**,

tomamos o exemplo do candidato Jair Bolsonaro onde dois registros de *lives* tiveram o maior número de interações dentre os 13 candidatos.

Gráfico 1 - Comparativo dos *Lives* com maior interação do candidato Jair Bolsonaro.



Fonte: Dados da pesquisa.

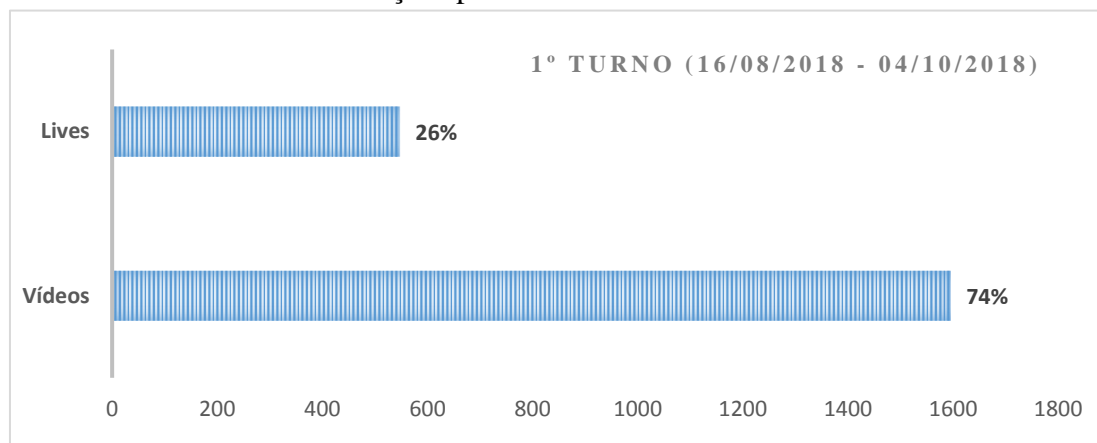
Os dados dos *lives* ID nº 671829229862929, publicado no dia 31 de agosto de 2018 e ID nº 479151242596625, publicado no dia 29 de agosto de 2018, foram analisados considerando quatro pontos de referência, ou seja – dados das mesmas datas. Nota-se que o primeiro *live* iniciou com um pouco mais que 500 mil interações e seguiu crescendo significativamente até atingir 1.338.815. Já o segundo *live*, iniciou com uma interação acima de 900 mil, no entanto o crescimento desacelerou, batendo em 1.096.314. O primeiro registro foi gravado em Porto Velho e possui uma duração mais curta, de 5:54; o segundo foi publicado em Porto Alegre, tendo 11:25 de duração.

4.2.1 Primeiro Turno das Eleições

Foram coletados 3.052 conteúdos audiovisuais no total, sendo que 2.142 registros correspondem ao período do 1º turno das eleições e 275 correspondem ao 2º turno, 231 são registros fora dos períodos definidos para a coleta, mas que foram coletados por conterem informações pertinentes quanto ao apoio aos candidatos ao 2º turno, bem como suas manifestações de agradecimento aos eleitores. Estes critérios também foram considerados para os dois candidatos do 2º turno, buscando preservar a riqueza do contexto das eleições do início ao fim. Além destes registros, acrescentamos os 404 conteúdos audiovisuais provenientes da página do ex-presidente Luiz Inácio Lula da Silva, que não foram analisados em conjunto com

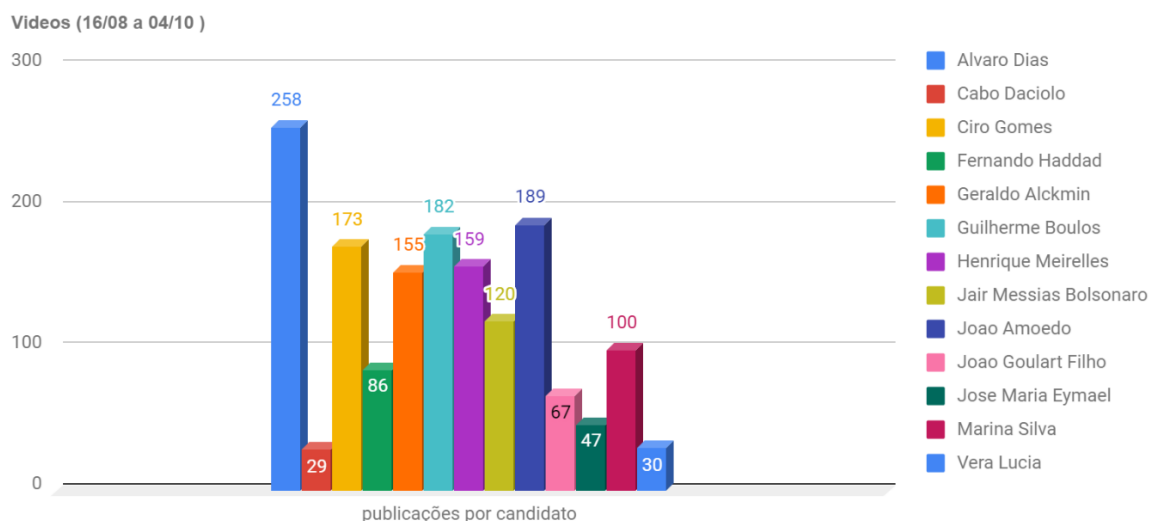
os 13 candidatos oficiais. Assim, a coleta no 1º turno, obteve 1.595 vídeos convencionais e 547 *lives* correspondentes aos 13 candidatos, como demonstrado no **Gráfico 2**.

Gráfico 2 - Quantitativo de registros audiovisuais publicados e coletados no 1º turno das eleições presidenciais de 2018.



Fonte: Dados da pesquisa.

Dos 1.595 vídeos convencionais publicados, observamos que sete candidatos obtiveram um número de publicações acima de 100. Como ilustra o **Gráfico 3**, o candidato Álvaro Dias atingiu o maior número de publicações – com 258 vídeos. Na sequência, apareceu João Amoêdo em 2º lugar com 189 vídeos, Guilherme Boulos em 3º lugar com 182 vídeos, Ciro Gomes em 4º lugar com 173 vídeos, Henrique Meirelles em 5º lugar com 159 vídeos, Geraldo Alckmin em 6º lugar com 155 vídeos e Jair Bolsonaro em 7º lugar com 120 vídeos. Além disso, seis candidatos aparecem com menos publicações de vídeos, Marina Silva em 8º lugar com 100 vídeos, Fernando Haddad em 9º lugar com 86 vídeos, João Goulart Filho em 10º lugar com 67 vídeos, Jose Maria Eymael em 11º lugar com 47 vídeos, Vera Lucia em 12º lugar com 30 vídeos e Cabo Daciolo em 13º lugar com 29 vídeos.

Gráfico 3 - Quantitativo de vídeos publicados no *Facebook* no 1º turno por candidato.

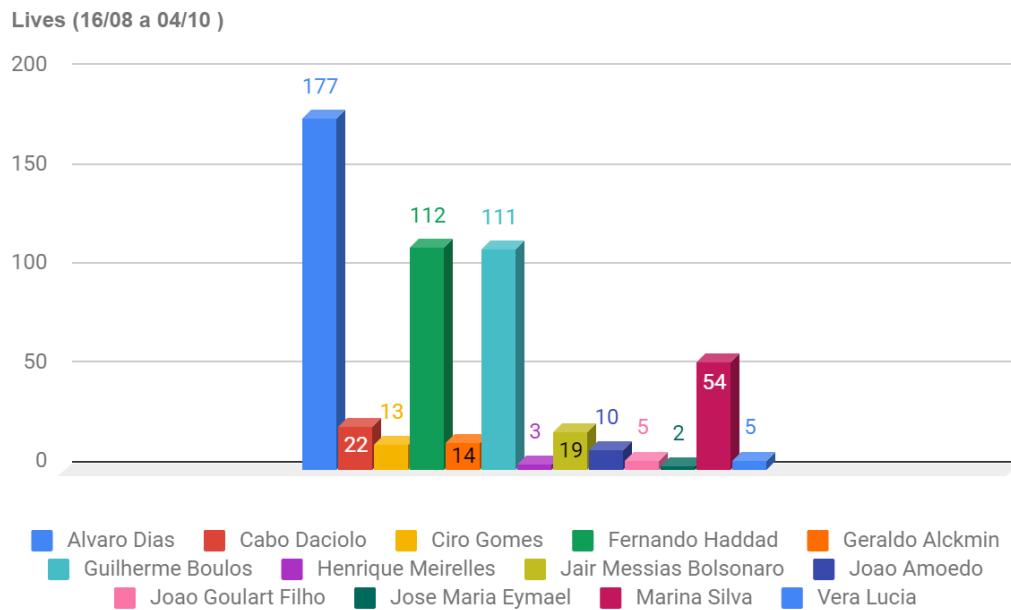
Fonte: Dados da pesquisa.

No que se refere a quantidade de publicações por candidato, não se observou relação de influência nos resultados das eleições considerando que a disputa no 2º turno ocorreu entre dois candidatos com uma produção média em relação a publicações audiovisuais no *Facebook*. Jair Bolsonaro alcançou 46,03% dos votos válidos e Fernando Haddad atingiu 29,28% dos votos válidos (**Anexo A**).

Também é importante considerar que a coleta na página do candidato Fernando Haddad iniciou apenas no dia 11 de setembro de 2018, após a retirada da candidatura do ex-presidente Lula. Outro acontecimento que, provavelmente, influenciou no número de publicações por parte do candidato Jair Bolsonaro, foi o atentado sofrido no dia 06 de setembro de 2018.

Já o número de publicações de *lives*, no geral, foi mais baixo em todos os casos com exceção do candidato Fernando Haddad que teve mais publicações em sua página nesta modalidade. Como mostra o **Gráfico 4**, os candidatos que tiveram mais publicações de *lives* foram Álvaro Dias com 177 registros, Fernando Haddad com 112 registros e Guilherme Boulos com 111 registros. Marina Silva aparece com 54 registros, Cabo Daciolo com 22 registros, Jair Bolsonaro com 19 registros, Geraldo Alckmin com 14 registros, Ciro Gomes com 13 registros, João Amoêdo com 10 registros, João Goulart Filho e Vera Lucia com 5 registros cada, e, por fim, Henrique Meirelles e Jose Maria Eymael com 3 e 2 registros respectivamente.

Gráfico 4 - Quantitativo de *Lives* publicados no *Facebook* no 1º turno por candidato.



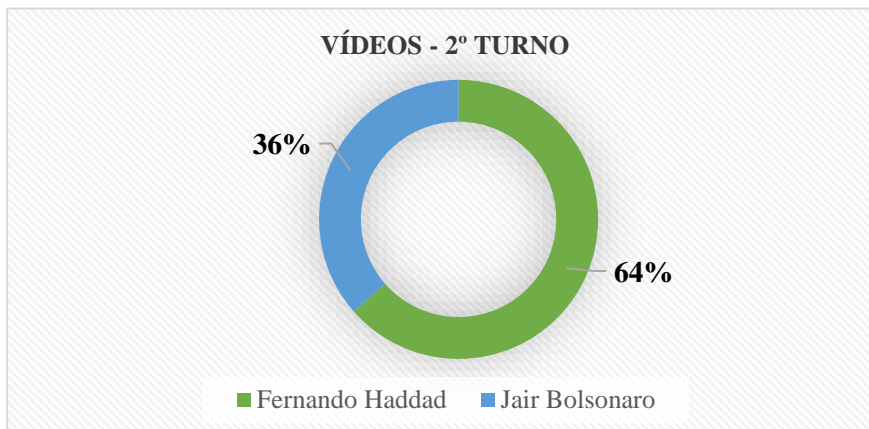
Fonte: Dados da pesquisa.

Por fim, dos 404 registros coletados da página do ex-presidente Lula no primeiro turno, 154 eram vídeos e 250 eram *lives*, representando 38% e 62% respectivamente. No entanto, durante a coleta, observou-se que a maior parte dos *lives* eram produzidos por Haddad, que compartilhou os registros em ambas as páginas. Ou, então, foram gravados vídeos ao vivo de conteúdos audiovisuais antigos onde o ex-presidente aparecia.

4.2.2 Segundo Turno das Eleições

O segundo turno das eleições presidenciais iniciou no dia 08 de outubro de 2018 e terminou no dia 26 de outubro de 2018. Entre os dois candidatos, Fernando Haddad se mostrou mais ativo nas publicações de conteúdos audiovisuais do que o candidato Jair Bolsonaro. Como representado no **Gráfico 5**, foram 101 registros disponibilizados na *fanpage* de Haddad e 58 registros disponibilizados na *fanpage* de Bolsonaro. Assim, gerou uma proporção de 64% e 36% respectivamente.

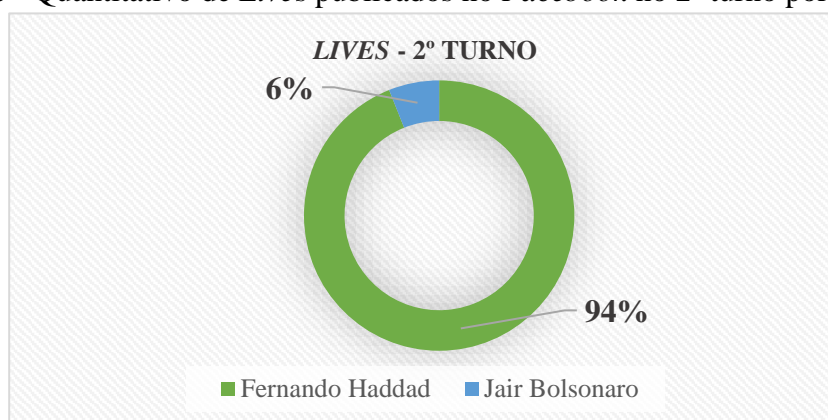
Gráfico 5 - Quantitativo de vídeos publicados no *Facebook* no 2º turno por candidato.



Fonte: Dados da pesquisa.

A coleta dos vídeos ao vivo no 2º turno resultou em 116 registros, sendo a maioria do candidato Fernando Haddad, com 109 *lives* e, apenas, 7 *lives* do candidato Jair Bolsonaro (**Gráfico 6**).

Gráfico 6 - Quantitativo de *Lives* publicados no *Facebook* no 2º turno por candidato.



Fonte: Dados da pesquisa.

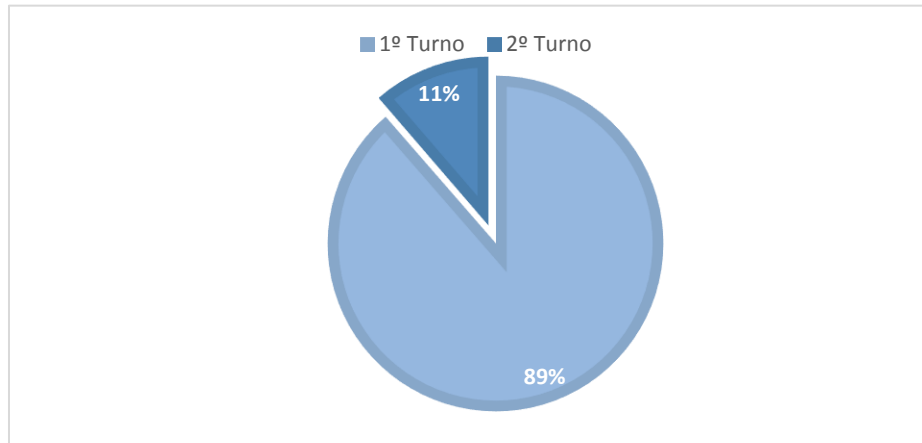
Cabe frisar que a quantidade de *lives* do candidato Jair Bolsonaro pode ter sido drasticamente reduzida pela necessidade de o candidato ter que passar por procedimentos cirúrgicos, o que o afastou da campanha.

4.2.3 Comparativos do 1º e 2º Turnos

Se compararmos o número de registros coletados em ambos os turnos, teremos a proporção de 89% para o 1º turno e 11% para o 2º turno (**Gráfico 7**). Proporcionalmente, no 2º

turno, temos um aumento de publicações, considerando o menor tempo de campanha e apenas dois candidatos.

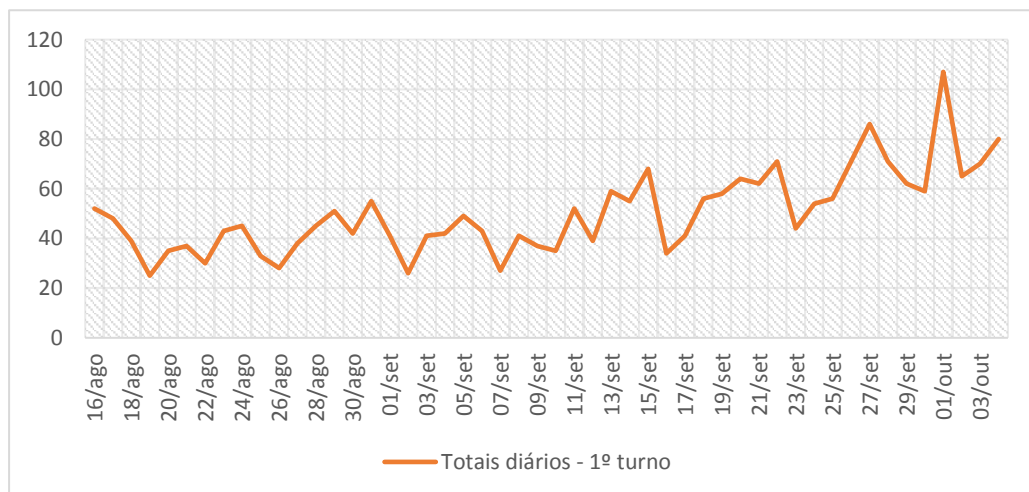
Gráfico 7 - Comparativo de publicações no *Facebook* do 1º e do 2º turno.



Fonte: Dados da pesquisa.

A partir do **Gráfico 8**, observamos que o quantitativo de publicações diárias variou bastante no primeiro turno, situado entre 20 (no início da campanha) e 120 (no final do 1º turno).

Gráfico 8 - Quantitativo diário de publicações no *Facebook* no 1º turno.

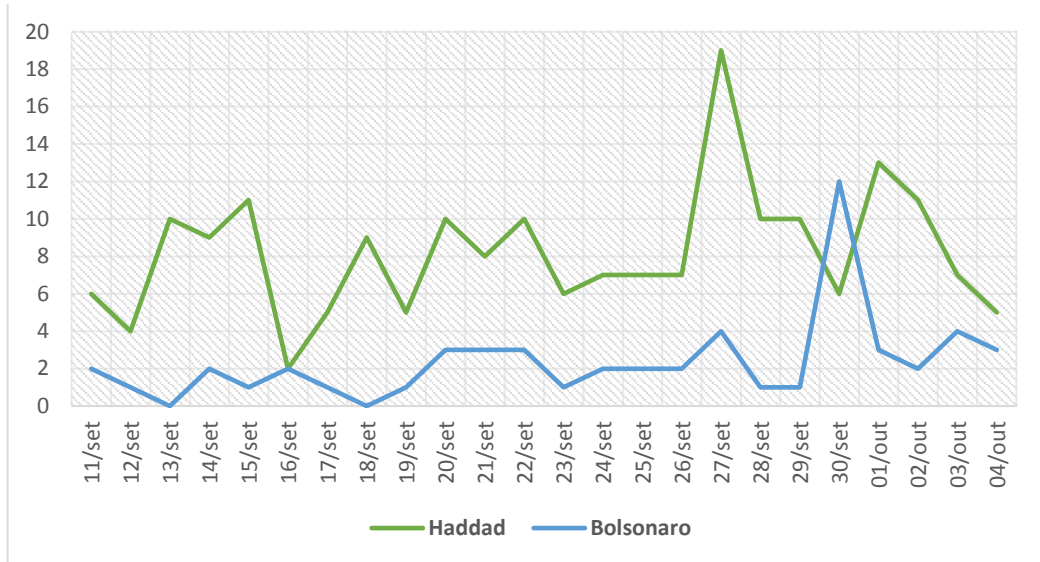


Fonte: Dados da pesquisa.

No comparativo de número de publicações dos dois principais candidatos no primeiro turno das eleições a predominância foi do candidato Fernando Haddad em detrimento de Jair Bolsonaro, que apenas no dia 30 de setembro realizou mais publicações que Haddad. No 1º turno Haddad não ultrapassou de 19 publicações de conteúdos audiovisuais, publicando todos

os dias, pelo menos, 2 registros. Já o candidato Bolsonaro, não efetuou publicações em 2 dias, atingindo o máximo de 12 registros no dia 30 de setembro conforme o gráfico abaixo.

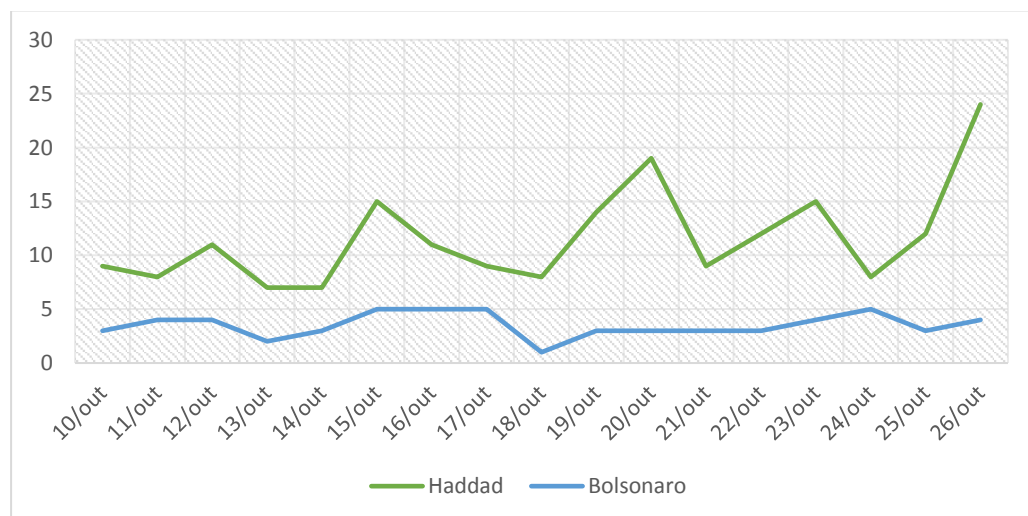
Gráfico 9 - Comparativo de publicações no *Facebook* no 1º turno dos dois principais candidatos.



Fonte: Dados da pesquisa.

No segundo turno, **Gráfico 10**, os quantitativos de publicações dos candidatos ficaram mais heterogêneos ainda, sendo que o candidato Jair Bolsonaro não ultrapassou de cinco publicações diárias, mas manteve a publicação diária de ao menos um registro. Já Haddad publicou até 26 vídeos em um mesmo dia, sem baixar de 7 registros diários.

Gráfico 10 - Comparativo de publicações no *Facebook* no 2º turno dos dois principais candidatos.



Fonte: Dados da pesquisa.

A partir dos metadados coletados, também realizamos uma análise comparativa do tempo de aparição dos candidatos no Facebook e nos blocos da TV. Levamos em consideração a distribuição de tempo ilustrada no infográfico do **Anexo B**, bem como a quantidade total de blocos ocorridos durante a campanha. Todos os valores foram convertidos para segundos, como mostra a **Tabela 4**, com a aplicação dos seguintes cálculos: 1) $TM * 60 + TS * 100$; 2) $TS * 100$; 3) $TM * 60 + TS * 138$ e 4) $TS * 138$. Onde TM corresponde ao tempo em minutos e TS corresponde ao tempo em segundos, o número 100 é a soma de blocos do 1º turno (2 blocos * 50 dias) e o número 138 é a soma de blocos do 1º e 2º (2 blocos * 69 dias).

Tabela 4 - Quantitativo em segundos do tempo de aparição no *Facebook* e TV durante a campanha à Presidência de 2018.

| Candidato | Facebook (vídeos) | TV (blocos) |
|--------------------|----------------------|----------------|
| Álvaro Dias | 190874 | 4000 |
| Cabo Daciolo | 40345 | 800 |
| Ciro Gomes | 33756 | 3800 |
| Fernando Haddad | 243153 | 19734 |
| Geraldo Alckmin | 20614 | 33200 |
| Guilherme Boulos | 137568 | 1300 |
| Henrique Meirelles | 9301 | 11500 |
| Jair Bolsonaro | 54581 | 1104 |
| João Amoêdo | 22944 | 500 |
| João Goulart Filho | 13688 | 500 |
| José Maria Eymael | 6196 | 800 |
| Marina Silva | 66770 | 2100 |
| Vera Lucia | 7847 | 500 |
| Totais | 847637 | 79838 |

Fonte: Dados da pesquisa.

Todos os candidatos receberam do TSE o tempo previsto para a propaganda eleitoral em rádio e TV, assim, trouxemos os dados das aparições previstas nos blocos das 13h e das 20h30. Também observamos que apenas os candidatos Geraldo Alckmin e Henrique Meirelles tiveram o maior tempo de exposição na TV do que no *Facebook* por meio de vídeos. No geral, os demais candidatos usufruíram da liberdade das mídias sociais superando as limitações da TV – o que reforça a tendência migratória da personalização das campanhas para as mídias sociais.

Após concluir a análise dos quantitativos das publicações coletadas, passamos para a análise da disponibilidade desses registros em sua fonte de origem a fim de entendermos os impactos da falta de iniciativas de arquivamento da *Web* e mídias sociais no Brasil.

4.3 Disponibilidade dos Vídeos Publicados pelos Candidatos

O segundo ponto importante de nossa pesquisa consistiu em identificar se os registros coletados ainda estavam em seu formato original após o término das eleições. Esta análise foi realizada considerando a quantidade total dos registros coletados durante o período de campanha eleitoral, ou seja, os 2.821 registros audiovisuais.

A análise ocorreu manualmente, no dia 15 de novembro de 2018 – 18 dias após o término das eleições e teve como resultado os dados da **Tabela 5**, onde temos um total de 32 vídeos e/ou *lives* indisponíveis no link original da *fanpage*. Dos 32 registros indisponíveis, 27 correspondem aos 13 candidatos e 5 são de origem da página do ex-presidente Lula.

Tabela 5 - Tabela de persistência dos conteúdos audiovisuais nas *fanpages* dos candidatos no *Facebook* - novembro.

| Novembro | |
|--------------------|--------------------------|
| Candidatos | Vídeos/ <i>lives</i> off |
| Álvaro Dias | 0 |
| Cabo Daciolo | 2 |
| Ciro Gomes | 0 |
| Fernando Haddad | 2 |
| Geraldo Alckmin | 5 |
| Guilherme Boulos | 2 |
| Henrique Meirelles | 2 |
| Jair Bolsonaro | 3 |
| João Amoêdo | 4 |
| João Goulart Filho | 1 |
| José Maria Eymael | 5 |
| Marina Silva | 0 |
| Vera Lucia | 1 |
| Lula | 5 |
| Total | 32 |

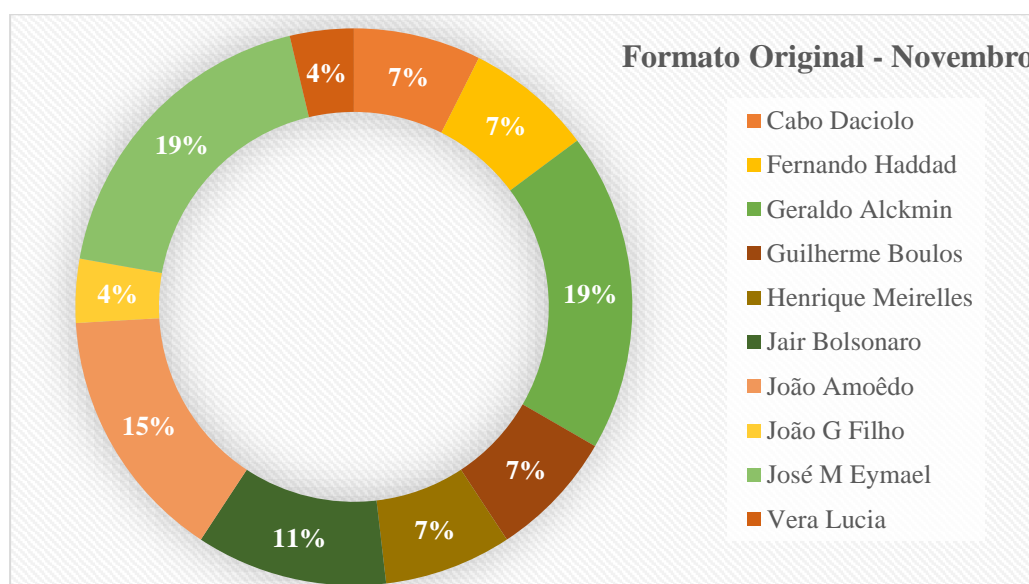
Fonte: Dados da pesquisa.

Apenas três candidatos, em novembro, ainda mantinham todos os conteúdos audiovisuais disponíveis em suas *fanpages*: o candidato Álvaro Dias, o candidato Ciro Gomes e a candidata Marina Silva. Os demais candidatos tiveram, pelo menos, um registro indisponível, chegando a ter até cinco registros indisponíveis de Geraldo Alckmin, José Maria Eymael e do ex-presidente Lula.

Entendemos que esta primeira análise já nos dá subsídios concretos para defender o fomento de pesquisas e desenvolvimentos de projetos sobre o arquivamento da *Web* brasileira em redes sociais como o *Facebook*, principalmente no campo da política, pois – após apenas 18 dias da coleta e num contexto histórico e político muito importante para os brasileiros, 32 registros já não podiam mais ser acessado em seu formato original. Registros estes que, possivelmente, ainda estão sendo estudados sobre o impacto que geraram no resultado das eleições presidenciais de 2018.

O **Gráfico 11** apresenta o percentual correspondente a cada candidato dos 27 conteúdos indisponíveis, sendo os candidatos Geraldo Alckmin e José Maria Eymael os mais atingidos, com 19% cada. Depois aparece o candidato João Amoêdo com 15% e Jair Bolsonaro com 11%. Com 7% cada, aparecem Cabo Daciolo, Fernando Haddad, Guilherme Boulos e Henrique Meirelles, com 4% cada, João Goulart Filho e Vera Lucia. Nesta análise, não constam os dados do ex-presidente Lula pois não se insere diretamente no grupo dos candidatos oficiais.

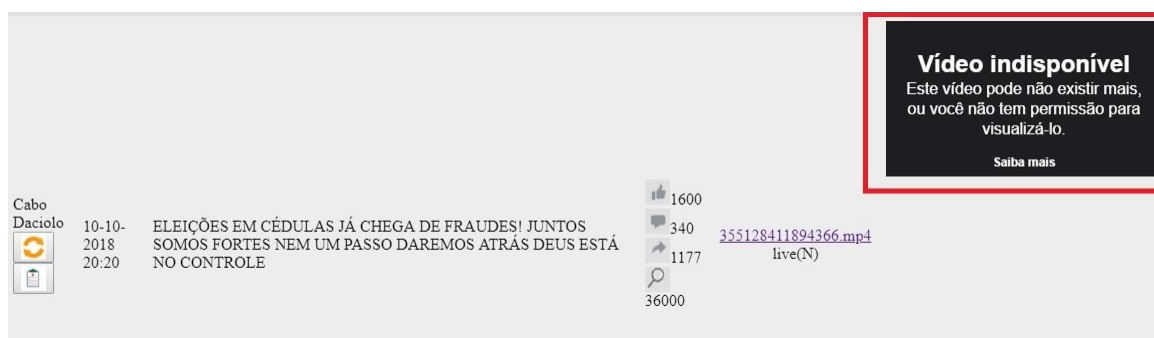
Gráfico 11 - Análise de disponibilidade do formato original das publicações dos candidatos no *Facebook* - novembro.



Fonte: Dados da pesquisa.

Os 32 registros foram verificados um a um de duas formas: a primeira utilizando a página de consulta, que mostra o *iframe* da publicação original enquanto disponível no *Facebook* (**Figura 9**) e a segunda acionando o link de origem da publicação e verificando qual era a mensagem veiculada.

Figura 9 - Exemplo de identificação de registro indisponível a partir do *iframe*.



Fonte: Elaborado pela autora, 2019.

As duas opções de mensagens que aparecem no *Facebook* quando uma publicação não se encontra no seu formato original, como ilustra a **Figura 10**, descrevem que: 1) o conteúdo não está disponível no momento ou 2) o conteúdo está indisponível devido a restrições legais.

Figura 10 - Mensagens padrões de conteúdo indisponível na Página do *Facebook*.



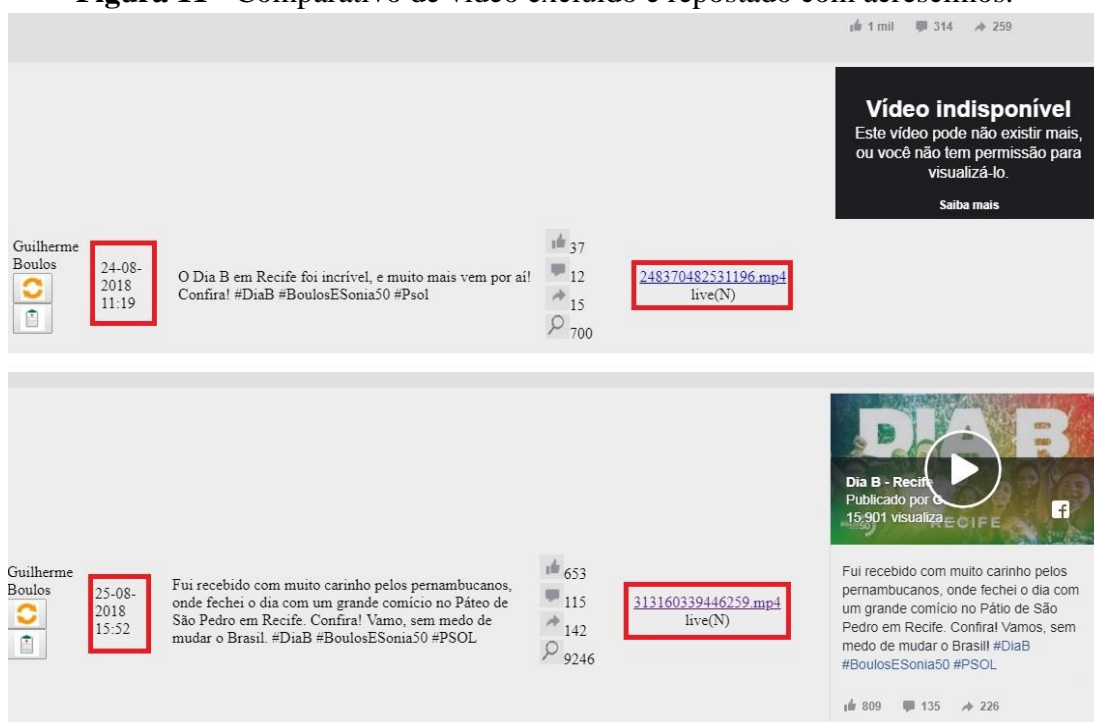
Fonte: Facebook, 2019.

No primeiro caso, o conteúdo pode ter sido removido da página do candidato; ter restringido o acesso a um grupo específico de usuários; ou pode ter sido removido e repostado nas mesmas condições ou com alterações, dentre outras possibilidades que esta pesquisa não se aprofundou.

No segundo caso, uma restrição legal no Brasil significa que uma pessoa ou um grupo de pessoas entrou com um processo na justiça solicitando a retirada do conteúdo da rede por entender ou provar que o registro continha informações prejudiciais.

A título de exemplo, trazemos o caso de um registro publicado na página do candidato Boulos que foi removido, mas que foi percebida a sua “respostagem” no dia seguinte. A **Figura 11** afere que o primeiro vídeo foi publicado no dia 24 de agosto de 2018 com o nº 248370482531196 e com duração de 2 minutos e 16 segundos. E, o segundo vídeo, foi publicado no dia seguinte, 25 de agosto, e – durante a coleta, foi percebido que a imagem inicial do *iframe* era a mesma do vídeo excluído (o que não é comum). Assim, comparou-se ambos os vídeos em busca de esclarecer o ocorrido, percebendo – então, que o segundo registro contava com uma duração de 3 minutos e 28 segundos, ou seja, um minuto e doze segundos a mais que a primeira versão. Também se observou que a descrição do segundo vídeo possuía um texto diferente da constante no primeiro vídeo.

Figura 11 - Comparativo de vídeo excluído e repostado com acréscimos.



Fonte: Elaborado pela autora, 2019.

Todavia, cabe expor que nem todos os registros indisponíveis na data de análise, necessariamente, estarão indisponíveis permanentemente uma vez que podem ter tido seu acesso apenas restringido e, assim – quando de interesse do candidato, voltarem a serem visualizados publicamente de maneira ampla ou ainda terem sido objeto de outro tipo de problema de visibilidade e acesso momentâneo.

4.4 Preservação em Plataformas de Arquivamento da Web

Para entendermos o impacto da presente pesquisa, percebemos que era necessário investigar de forma exploratória se houve preservação dos conteúdos publicados pelos candidatos no *Facebook*. No entanto, constatamos que os presidenciáveis possuíam outras fontes de informação em mídias sociais e, assim, optamos por analisar todas estas fontes disponíveis e os *sites* oficiais das campanhas a fim de tornar a pesquisa um pouco mais completa.

Montamos uma tabela de referência das fontes de informações encontradas de cada candidato (mídias sociais e *sites*), onde notamos, baseado em uma observação ainda superficial, que a *Web* se tornou um ambiente bastante usual pelos candidatos à Presidência, os quais não estão presentes apenas no *Facebook* e *Twitter*, mas também no *Youtube* e *Instagram*. Dos 13 candidatos, apenas José Maria Eymael não possuía uma conta no *Youtube* e Vera Lucia não possuía conta nem no *Youtube* e nem no *Instagram*. Já a conta do *Youtube* onde ocorreu a propaganda e comunicação eleitoral de Fernando Haddad era uma conta da coligação “O Brasil Feliz De Novo” e não do candidato individualmente. Também observamos que os candidatos Cabo Daciolo e Vera Lucia não possuíam um *site* individual, apenas o *site* do partido. Todos os endereços eletrônicos das fontes de referência podem ser consultados no **Apêndice G**.

Para a análise de preservação, escolhemos três plataformas que arquivam conteúdos brasileiros, iniciando pelo *Internet Archive* – o qual coleta automaticamente, mas também sob demandas de usuários. Em seguida, pesquisamos no *Archive-It*, que – apesar de ser um serviço também desenvolvido pelo *Internet Archive*, por ser pago, poderia conter registros não constantes no IA. Por fim, complementando a pesquisa, utilizamos o *Archive.is* ou *Archive.today* por ser uma proposta de preservação totalmente manual e sob a demanda dos usuários.

Para pesquisar nas plataformas, utilizamos os URLs das mídias e *sites* dos candidatos, identificados anteriormente, e consideramos os registros salvos dentro do período das campanhas eleitorais, ou seja, de agosto a outubro de 2018.

A análise ocorreu no dia 18 de novembro de 2018 e, embora ela seja secundária dentro da presente pesquisa, se mostrou necessária para demonstrar o quão pouco os registros dos candidatos das eleições presidenciais foram preservados e o quanto a iniciativa de preservar os conteúdos audiovisuais poderá suprir um hiato histórico, político e da memória digital brasileiros.

Iniciamos nossa análise pelos registros provenientes do *Facebook*, pois se trata do nosso contexto atual de pesquisa. Contudo, apontamos as principais diferenças entre a nossa coleta e a coleta realizada pelas plataformas em questão. A nossa coleta priorizou os registros audiovisuais completos, agregando as informações e metadados que os acompanhavam. Em momento algum, nesta pesquisa, se cogitou preservar a apresentação destes registros de forma idêntica ao conteúdo disponível no *Facebook*. Ao contrário, as plataformas *Internet Archive*, *Archive-It* e *Archive.is* operam coletando um *snapshot* da página, que irá preservar o conteúdo disponível no momento da coleta, isto significa que se a página possuía um vídeo no momento da coleta, foi preservado uma imagem do vídeo e não o vídeo na íntegra. Para preservar o vídeo, seria necessário executar uma nova coleta do vídeo em si o vinculando ao link anterior já preservado.

Um exemplo claro pode ser observado na **Figura 12**, onde aparece a página do candidato Jair Bolsonaro no *Facebook* que foi preservada no dia 17 de setembro de 2018. Percebemos que toda a estrutura visual foi preservada tal qual a da página no *Facebook* no momento da coleta, porém o vídeo que aparece não pode ser reproduzido, pois trata-se de uma imagem, bem como nenhum link pode ser acessado, pois a coleta não foi intensiva.

Figura 12 - *Snapshot* de vídeo do *Facebook* preservado no *Archive-It*.



Fonte: *Archive-It*, 2019.

A preservação de *snapshots* das páginas dos candidatos no *Facebook* ocorreu, predominantemente, através da plataforma *Internet Archive* e do serviço *Archive-It*, mas também se observou que alguns *snapshots* foram realizados através do *Archive.is*.

Embora não pareça significativo os resultados encontrados na plataforma do *Archive.is*, precisamos compreender que as páginas do *Facebook* são dinâmicas e coletas realizadas com segundos de intervalo podem preservar um conteúdo totalmente diferente e único.

Como vemos na **Tabela 6**, apenas a candidata Vera Lucia não teve nenhum registro preservado em nenhuma plataforma.

Tabela 6 - Tabela de plataformas de arquivamento da *Web* - registros do *Facebook*.

| PLATAFORMAS DE PRESERVAÇÃO | | | |
|----------------------------|------------------|------------|------------|
| CANDIDATO | Internet Archive | Archive.It | Archive.is |
| Álvaro Dias | Sim | Sim | Não |
| Cabo Daciolo | Sim | Sim | Não |
| Ciro Gomes | Sim | Sim | Não |
| Fernando Haddad | Sim | Não | Sim |
| Geraldo Alckmin | Sim | Sim | Não |
| Guilherme Boulos | Sim | Sim | Não |
| Henrique Meirelles | Sim | Sim | Não |
| Jair Bolsonaro | Sim | Sim | Sim |
| João Amoêdo | Sim | Sim | Não |
| João G Filho | Sim | Sim | Não |
| José M Eymael | Sim | Sim | Não |
| Marina Silva | Sim | Sim | Não |
| Vera Lucia | Não | Não | Não |
| Lula* | Sim | Sim | Não |

*não foi candidato oficial

Fonte: Dados da pesquisa.

Os candidatos que mais tiveram registros preservados, não em quantidade, mas em diferentes plataformas, foram Geraldo Alckmin e Jair Bolsonaro. O ex-presidente Lula também teve registros preservados nas plataformas analisadas.

Entretanto, observou-se que os registros preservados são limitantes e não cobrem todo o período de publicações realizadas pelos candidatos durante a campanha eleitoral. Bem como todas as publicações audiovisuais não são contempladas, sendo registrados, apenas (e, em alguns casos, em parcialidade) as descrições destes conteúdos audiovisuais. A preservação nestas plataformas ocorre com maior sucesso quando são registros de imagem, no caso de uma imagem por publicação, e textos curtos com 70/75 caracteres aproximadamente, sendo que acima disso é necessário clicar no link “ver mais”, o qual não foi preservado.

Dentre as quatro mídias sociais analisadas como fonte de informações publicadas pelos candidatos, o *Facebook*, embora não de maneira intensiva e satisfatória, foi o que apresentou mais registros coletados pelas plataformas de preservação.

Passamos para a análise dos registros dos candidatos no *Twitter*, **Tabela 7**, onde se verificou uma queda significativa dos índices de preservação em comparação com a tabela anterior.

Tabela 7 - Tabela de plataformas de arquivamento da *Web* - registros do *Twitter*.

| PLATAFORMAS DE PRESERVAÇÃO | | | |
|----------------------------|------------------|------------|------------|
| CANDIDATO | Internet Archive | Archive.It | Archive.is |
| Álvaro Dias | Sim | Não | Não |
| Cabo Daciolo | Não | Não | Não |
| Ciro Gomes | Sim | Não | Não |
| Fernando Haddad | Sim | Não | Sim |
| Geraldo Alckmin | Sim | Não | Sim |
| Guilherme Boulos | Sim | Não | Sim |
| Henrique Meirelles | Sim | Não | Não |
| Jair Bolsonaro | Sim | Não | Sim |
| João Amoêdo | Sim | Não | Não |
| João G Filho | Não | Não | Não |
| José M Eymael | Sim | Não | Não |
| Marina Silva | Sim | Não | Não |
| Vera Lucia | Não | Não | Não |
| Lula* | Sim | Não | Sim |

* não foi candidato oficial

Fonte: Dados da pesquisa.

A preservação de registros publicados no *Twitter* é a que obtém maior sucesso em termos de completude e isto se deve pelo fato dos textos terem uma limitação padrão de 280 caracteres. Sendo assim, a mensagem sempre é capturada na íntegra, bem como a imagem da publicação. Todavia, as publicações de vídeos seguem com a mesma limitação percebida na preservação das páginas do *Facebook*.

Observou-se que a preservação parcial dos registros dos presidencialistas, no *Internet Archive*, ocorreu em dois formatos, predominantemente, no padrão de desktop, mas também houve alguns registros na versão *mobile*. No entanto, para nossa análise não consideramos os registros na versão *mobile* por serem em número inferior.

Para futuras pesquisas que almejem realizar um comparativo, por exemplo, entre as versões arquivadas pelo *Internet Archive* pode ser inserido no link do candidato no *Twitter* a palavra *mobile*, ex.: <https://mobile.Twitter.com/candidatoX>.

Ao passarmos para a análise no *Youtube*, notamos que nem todos os links dos candidatos no *Youtube* contam com um nome de usuário, sendo alguns identificados a partir de um código automático gerado pela plataforma no momento de criação da conta.

Identificamos os mesmos problemas encontrados nos arquivamentos anteriores: poucos registros e nenhuma intensidade. Como apresentamos na **Tabela 8**, apenas cinco candidatos foram contemplados com alguns *snapshots* e todos eles foram realizados pela plataforma *Internet Archive*.

Tabela 8 - Tabela de plataformas de arquivamento da *Web* - registros do *Youtube*.

| PLATAFORMAS DE PRESERVAÇÃO | | | |
|-----------------------------------|-------------------------|-------------------|-------------------|
| CANDIDATO | Internet Archive | Archive.It | Archive.is |
| Álvaro Dias | Não | Não | Não |
| Cabo Daciolo | Não | Não | Não |
| Ciro Gomes | Não | Não | Não |
| Fernando Haddad | Sim | Não | Não |
| Geraldo Alckmin | Sim | Não | Não |
| Guilherme Boulos | Não | Não | Não |
| Henrique Meirelles | Não | Não | Não |
| Jair Bolsonaro | Sim | Não | Não |
| João Amoêdo | Não | Não | Não |
| João G Filho | Não | Não | Não |
| José M Eymael | Não | Não | Não |
| Marina Silva | Sim | Não | Não |
| Vera Lucia | Não | Não | Não |
| Lula* | Sim | Não | Não |

*não foi candidato oficial

Fonte: Dados da pesquisa.

Todos os vídeos em destaque nas páginas dos candidatos não estavam disponíveis para visualização, no entanto, alguns vídeos mais antigos que aparecem nas coleções “populares” e “uploads”, em miniatura, puderam ser assistidos.

Encerrando a análise das mídias sociais, trazemos os dados dos registros preservados do *Instagram* – uma plataforma de compartilhamento de imagens e vídeos – dos candidatos.

Assim como no caso do *Youtube*, apenas cinco candidatos obtiveram seus registros preservados, desta vez: *Ciro Gomes*, *Fernando Haddad*, *Jair Bolsonaro*, *Marina Silva* e, também, o ex-presidente *Lula*. Todos os arquivamentos também foram realizados apenas pelo *Internet Archive* (**Tabela 9**).

Tabela 9 - Tabela de plataformas de arquivamento da *Web* - registros do *Instagram*.

PLATAFORMAS DE PRESERVAÇÃO

| CANDIDATO | Internet Archive | Archive.It | Archive.is |
|--------------------|------------------|------------|------------|
| Álvaro Dias | Não | Não | Não |
| Cabo Daciolo | Não | Não | Não |
| Ciro Gomes | Sim | Não | Não |
| Fernando Haddad | Sim | Não | Não |
| Geraldo Alckmin | Não | Não | Não |
| Guilherme Boulos | Não | Não | Não |
| Henrique Meirelles | Não | Não | Não |
| Jair Bolsonaro | Sim | Não | Não |
| João Amoêdo | Não | Não | Não |
| João G Filho | Não | Não | Não |
| José M Eymael | Não | Não | Não |
| Marina Silva | Sim | Não | Não |
| Vera Lucia | Não | Não | Não |
| Lula* | Sim | Não | Não |

*não foi candidato oficial

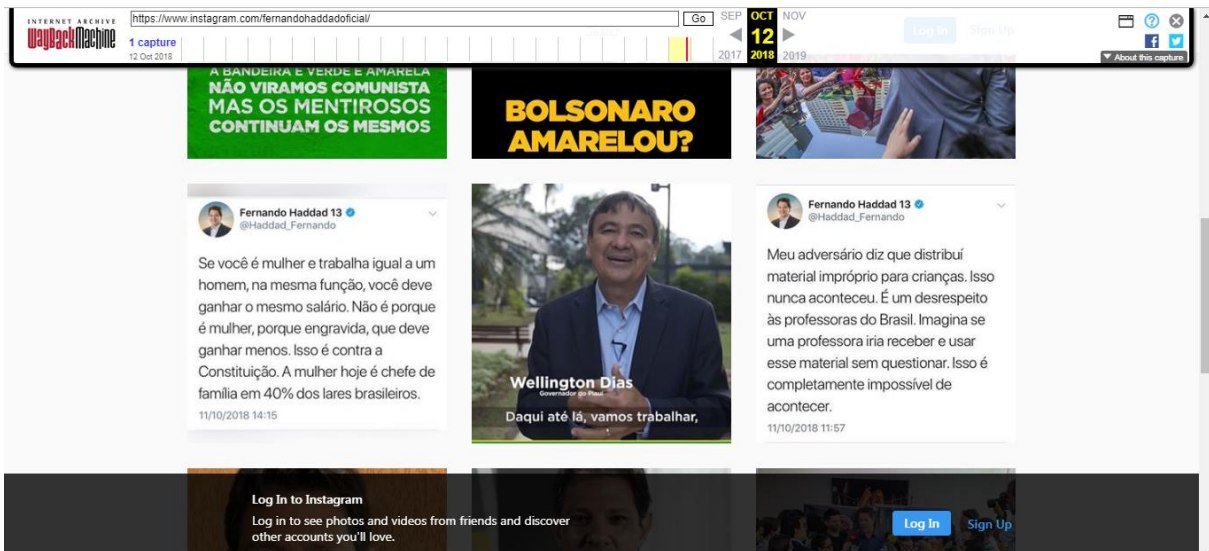
Fonte: Dados da pesquisa.

Da mesma forma que nas demais plataformas de mídias sociais, os vídeos seguem fora das políticas de preservação das plataformas de arquivamento da *Web* analisadas. Isto nos faz refletir sobre a infinidade e variedade de informações que podem se perder nestas mídias caso não haja mudanças significativas na sistemática das coletas.

Não podemos desconsiderar o aumento das publicações de conteúdos audiovisuais, assim como a multiplicidade e abundância de informações que um único registro pode conter.

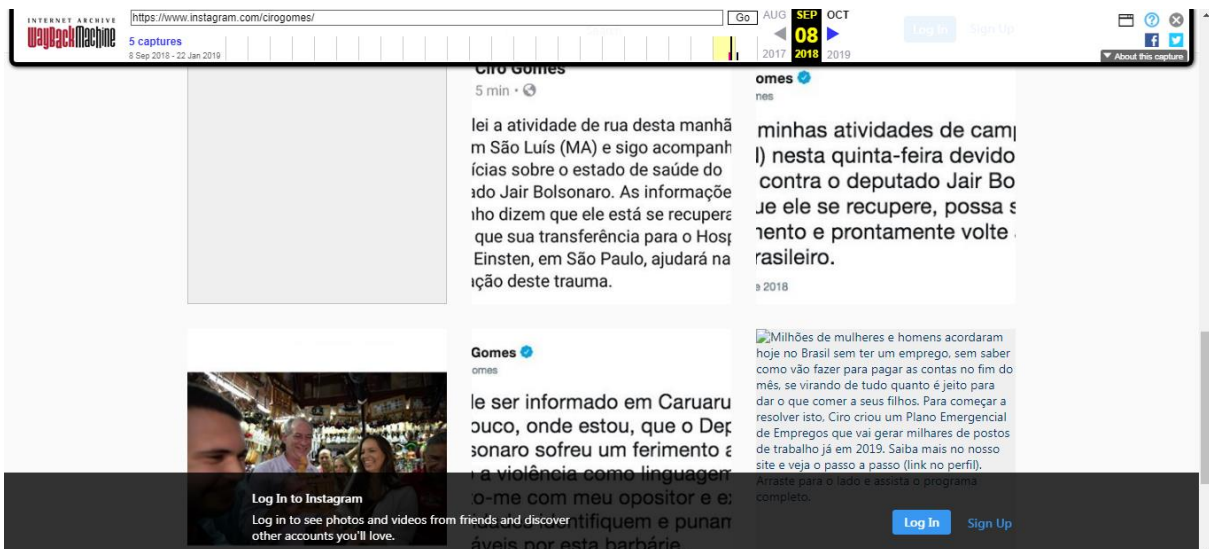
Na plataforma *Instagram*, os registros predominantes tendem a ser imagens, no entanto, para que a preservação desses registros seja completa, é necessário que a mesma tenha sido publicada no formato *square* ou quadrado (**Figura 13**), caso contrário a publicação aparecerá cortada na versão preservada (**Figura 14**).

Figura 13 - Snapshot da página do candidato Fernando Haddad no *Instagram* preservado pelo *Internet Archive*.



Fonte: Internet Archive, 2019.

Figura 14 - Snapshot da página do candidato Ciro Gomes no *Instagram* preservado pelo *Internet Archive*.



Fonte: Internet Archive, 2019.

Diferente do esperado, os *sites* utilizados para a campanha eleitoral dos candidatos não tiveram o maior percentual de preservação, se considerarmos as três plataformas de preservação analisadas. Apenas o *Internet Archive* contemplou todos os candidatos, sendo que o *Archive-It* não preservou nenhum registro e o *Archive.is* realizou apenas a coleta parcial do candidato Jair Bolsonaro (**Tabela 10**).

Tabela 10 - Tabela de plataformas de arquivamento da *Web* - registros dos *sites*.

| PLATAFORMAS DE PRESERVAÇÃO | | | |
|----------------------------|------------------|------------|------------|
| CANDIDATO | Internet Archive | Archive.It | Archive.is |
| Álvaro Dias | Sim | Não | Não |
| Cabo Daciolo | Sim | Não | Não |
| Ciro Gomes | Sim | Não | Não |
| Fernando Haddad | Sim | Não | Não |
| Geraldo Alckmin | Sim | Não | Não |
| Guilherme Boulos | Sim | Não | Não |
| Henrique Meirelles | Sim | Não | Não |
| Jair Bolsonaro | Sim | Não | Sim |
| João Amoêdo | Sim | Não | Não |
| João G Filho | Sim | Não | Não |
| José M Eymael | Sim | Não | Não |
| Marina Silva | Sim | Não | Não |
| Vera Lucia | Sim | Não | Não |
| Lula* | Sim | Não | Não |

*não foi candidato oficial

Fonte: Dados da pesquisa.

As informações preservadas dos *sites* aparecem em diversos formatos, sendo que os registros textuais e fotográficos também apresentam maior êxito no quesito integralidade.

Acreditamos que estes dados reforçam a necessidade de se pensar em preservação da *Web* além das possibilidades mais usuais e difundidas internacionalmente, pois percebemos que, além de escassa, vem deixando de fora muitos conteúdos audiovisuais que contém informações significativas para a nossa história.

Por fim, observamos uma falha, provavelmente no *script* de validação de links, no *Internet Archive*, o qual subtrai uma barra (/) do endereço, após o segundo “HTTPS” do link ou antes do “WWW”, quando acionado os links preservados constantes na tabela do **Apêndice H**. Sendo assim necessária sua inclusão manual na barra de pesquisa do navegador ou no campo de pesquisa da plataforma.

Por ex.: https://web.archive.org/web/*/https://www.instagram.com/candidatox

4.5 Apresentação do Protótipo de Arquivamento

O Protótipo é o resultado da necessidade de desenvolver um recurso autônomo a partir de códigos e plataformas de desenvolvimento *Web open source* voltados, estritamente, para a coleta de registros audiovisuais disponibilizados em *fanpages* públicas do *Facebook*.

O desenvolvimento deste Protótipo mostrou-se relevante por duas razões principais: a primeira consiste no fato da iniciativa ser voltada estritamente para uma produção brasileira,

pública e orgânica – onde, foram dispendidos esforços para coletar e preservar publicações audiovisuais, originais, disponibilizadas pelos candidatos à Presidência de 2018 em suas *fanpages* no *Facebook*, ação esta não identificada em outros estudos até o presente momento. E a segunda razão, foi demonstrada a partir da análise da, praticamente, inexistente preservação de vídeos deste contexto em plataformas globais de arquivamento da *Web*, como o caso do *Internet Archive*.

Contudo, entendemos que o protótipo consiste apenas em um meio de coleta, que proporcionou o êxito desta pesquisa. Assim, não se pretende nominá-lo como um recurso de arquivamento da *Web*, mais especificamente, do *Facebook*.

Todavia, este é um recurso com grande potencial para a pesquisa, no que diz respeito aos registros preservados, e, portanto, consideramos que deve ser garantido o amplo acesso a estes dados.

Para garantir o acesso dos vídeos, *lives* e demais metadados coletados das páginas dos candidatos, disponibilizamos uma interface *Web* simples para a pesquisa dos registros, a qual apresentamos abaixo as principais funcionalidades.

4.5.1 Recursos de Consulta

A principal forma de consulta ocorre através de uma lista de escolha com os nomes dos candidatos às eleições presidenciais de 2018 e do ex-presidente Lula (**Figura 15**).

Figura 15 - Lista de escolha com os nomes dos candidatos às eleições presidenciais de 2018 e do ex-presidente Lula do Protótipo.

| | |
|----------------|---------------------------|
| Nome Candidato | Selecione ▼ |
| | Selecione |
| | Todos |
| | Alvaro Dias |
| | Cabo Daciolo |
| | Ciro Gomes |
| | Fernando Haddad |
| | Geraldo Alckmin |
| | Guilherme Boulos |
| | Henrique Meirelles |
| | Jair Messias Bolsonaro |
| | Joao Amoedo |
| | Joao Goulart Filho |
| | Jose Maria Eymael |
| | Luiz Inacio Lula da Silva |
| | Marina Silva |
| | Vera Lucia |

Fonte: Elaborado pela autora, 2019.

Também é possível pesquisar todos os nomes na mesma consulta através da primeira opção, onde consta “Todos”. Contudo, como estamos pesquisando em uma base de dados com mais de três mil registros, aconselha-se que a pesquisa seja realizada por nome, a fim de garantir uma melhor performance de visualização dos dados.

Outro filtro importante que pode ser utilizado na pesquisa, em conjunto com o nome, é o do período (**Figura 16**), em que o usuário/pesquisador define a data inicial e a data final correspondente aos registros que deseja consultar.

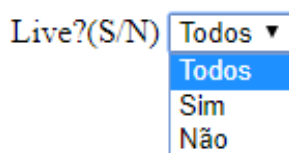
Figura 16 - Campo de período para consulta do Protótipo.



Fonte: Elaborado pela autora, 2019.

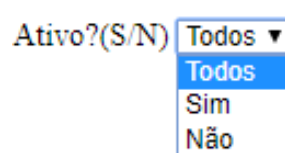
Para tornar a consulta mais específica e reduzida, podem ser usados outros dois filtros disponíveis: de tipo de conteúdo (**Figura 17**), se é *live* ou não; e o de disponibilidade do formato original (**Figura 18**), ativo ou não.

Figura 17 - Seleção de tipo de conteúdo, *Live* ou vídeo.



Fonte: Elaborado pela autora, 2019.

Figura 18 - Seleção de disponível no formato original, ativo ou não.

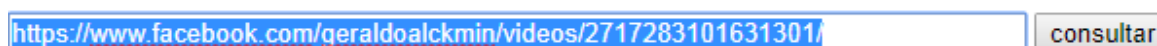


Fonte: Elaborado pela autora, 2019.

Tanto o filtro de período quanto os de seleção de tipo de conteúdo e de disponibilidade no formato original já aparecem preenchidos na tela inicial de pesquisa, sendo necessária a alteração para outras demandas de consulta. Estes filtros atuam obrigatoriamente em conjunto com o campo nome do candidato.

O último recurso de pesquisa consiste em um campo para inserção do *link* da publicação tal qual aparece na página do candidato como mostra o exemplo na **Figura 19**.

Figura 19 - Consulta pelo *link* da publicação conforme página do candidato.



Fonte: Elaborado pela autora, 2019.

Este recurso é independente dos demais pelo fato de o link original já oferecer todos as informações necessárias para a execução da consulta.

4.5.2 Recursos de Visualização

Ao efetuar a consulta pelo link da publicação extraída da página do candidato, obtemos o retorno dos dados referente a este registro conforme mostra a **Figura 20**.

Figura 20 - Visualização do *link* da publicação extraída da página do candidato.

Protótipo - Sistema de Arquivamento da WEB

Nome Candidato | Seleccione | entre 16/08/2018 e 27/03/2019 | Live?(S/N) | Todos | Ativo?(S/N) | Todos | <https://www.facebook.com/fernandohaddad/videos/2132678283716568/> | consultar

Protótipo de Sistema de Arquivamento - Dados de Candidatos Eleicoes 2018

| Nome | Criado | Mensagem | Curtidas | Comentarios | Compartilhamentos | Visualizacoes | Link Armazenado | Frame Facebook | Ativo |
|-----------------|---------------------|---|----------|-------------|-------------------|---------------|--|--|-----------|
| Fernando Haddad | 27-10-2018 12:45 | Há 43 anos, Vladimir Herzog foi assassinado pela ditadura no DOI-CODI. Essa mesma ditadura que Bolsonaro defende com convicção, sem remorso algum. Definitivamente, o Brasil não pode voltar para trás, como quer o deputado. Nós queremos construir um futuro sem violência e com esperança. | 8300 | 3100 | 10000 | 176893 | 2132678283716568.mp4 |  | nov 18: S |

Fonte: Elaborado pela autora, 2019.

Já na consulta por candidato, aparecerá uma lista de todos os registros encontrados para o conjunto de filtros aplicados (**Figura 21**). Os registros são organizados da data mais recente para a mais antiga.

Figura 21 - Apresentação de consulta por candidato.

Nome Candidato | Fernando Haddad | entre 16/08/2018 e 27/03/2019 | Live?(S/N) | Todos | Ativo?(S/N) | Todos | consultar

| | | | | | | | | | |
|-----------------|---------------------|---|-------|-------|-------|--------|--|---|-----------|
| Fernando Haddad | 27-10-2018 12:45 | Há 43 anos, Vladimir Herzog foi assassinado pela ditadura no DOI-CODI. Essa mesma ditadura que Bolsonaro defende com convicção, sem remorso algum. Definitivamente, o Brasil não pode voltar para trás, como quer o deputado. Nós queremos construir um futuro sem violência e com esperança. | 8300 | 3100 | 10000 | 176893 | 2132678283716568.mp4 |  | nov 18: S |
| Fernando Haddad | 27-10-2018 12:43 | Coletiva de imprensa com Fernando Haddad em São Paulo #Haddad13 #ViraVoto | 16000 | 32000 | 8600 | 302581 | 283774408934080.mp4 |  | nov 18: S |
| Fernando Haddad | 27-10-2018 12:41 | Marieta Severo, Osmar Prado, Antônio Pitanga e diversos artistas se somam à nossa campanha por um Brasil sem ódio e intolerância. | 27000 | 6900 | 37000 | 757807 | 1135423039916129.mp4 |  | nov 18: S |
| Fernando Haddad | 27-10-2018 12:40 | Aprendendo uma música nova ao lado de amigos e amigas. É a onda da virada. É importante intensificar a agenda de atividades e conquistar mais votos para o Brasil sair da crise e voltar a ser feliz de novo. #Haddad13 #ViraVoto Video: Ricardo Stuckert | 26000 | 5900 | 25000 | 632709 | 491287991385047.mp4 |  | nov 18: S |

Fonte: Elaborado pela autora, 2019.

Ao acionar o *link* do campo “*Link Armazenado*”, o registro preservado é carregado e já inicia a execução do mp4 (**Figura 22**).

Figura 22 - Execução do *link* armazenado.



Fonte: Elaborado pela autora, 2019.

Além da apresentação dos principais dados e conteúdos coletados, há dois recursos secundários. O primeiro, que mostra a evolução das interações em uma publicação dentro do 1º turno pode ser visualizado pelo botão que aparece abaixo do número de curtidas da publicação. Assim, irá abrir uma nova janela com a tabela de histórico (**Figura 23**).

Figura 23 - Lista de histórico de interações da publicação.

Protótipo - Sistema de Arquivamento da WEB

Nome Candidato entre e Live?(S/N) Ativo?(S/N)

| Data | Curtidas | Comentarios | Compartilhamentos | Visualizacoes |
|------------------|----------|-------------|-------------------|---------------|
| 17-08-2018 11:54 | 35336 | 25157 | 12863 | 370895 |
| 17-08-2018 12:24 | 35970 | 25321 | 13071 | 383651 |
| 17-08-2018 12:54 | 36506 | 25472 | 13263 | 396563 |
| 17-08-2018 13:24 | 36927 | 25637 | 13380 | 406603 |
| 17-08-2018 13:54 | 37350 | 25778 | 13491 | 415101 |
| 17-08-2018 14:24 | 37680 | 25867 | 13580 | 422924 |
| 17-08-2018 14:54 | 38011 | 25956 | 13669 | 430748 |

| Nome Candidato | Data | Conteúdo | Curtidas | Comentarios | Compartilhamentos | Link Armazenado | Visualizacoes |
|------------------------|------------------|--|----------|-------------|-------------------|-----------------|---------------------|
| Jair Messias Bolsonaro | 28-08-2018 11:39 | Continuação Bolsonaro no Ceasa 2 | 12662 | 7134 | | | |
| Jair Messias Bolsonaro | 28-08-2018 11:39 | Bolsonaro no CEASA 1 | 8701 | 5705 | | | |
| Jair Messias Bolsonaro | 22-08-2018 22:42 | Desembarque no aeroporto de Presidente Prudente - SP | 55256 | 35556 | | | |
| Jair Messias Bolsonaro | 17-08-2018 11:36 | Formatura de Sargentos da Polícia Militar do estado de São Paulo. (Live) | 34944 | 25038 | 12745 | 364206 | 428854424188636.mp4 |


Fonte: Elaborado pela autora, 2019.

O segundo recurso adicional consiste na possibilidade de se observar uma prévia do conteúdo tal qual o seu registro original (se ainda disponível) através do *iframe*. Para isto, basta clicar no botão com o logo do *Facebook* que o *iframe* abrirá em outra janela (**Figura 24**).

Figura 24 - Visualização do *iframe* da publicação na origem.

Protótipo - Sistema de Arquivamento da WEB

Nome Candidato: entre e Live?(S/N) Sim Ativo?(S/N)

| | | | | | | | | |
|------------------------|---------------------|--|-------|-------|-------|--------|-------------------------------------|--|
| Jair Messias Bolsonaro | 28-08-2018 11:39 | Continuação Bolsonaro no Ceasa 2 | 12662 | 7134 | | | | |
| Jair Messias Bolsonaro | 28-08-2018 11:39 | Bolsonaro no CEASA 1 | 8701 | 5705 | | | | |
| Jair Messias Bolsonaro | 22-08-2018 22:42 | Desembarque no aeroporto de Presidente Prudente - SP | 55256 | 35556 | | | | |
| Jair Messias Bolsonaro | 17-08-2018 11:36 | Formatura de Sargentos da Policia Militar do estado de São Paulo. (Live) | 34944 | 25038 | 12745 | 364206 | 428854424188636.mp4 |  nov: 18:5 |

The image shows a screenshot of a web application interface. On the left is a table with columns for candidate name, date, event title, and various statistics. On the right is a browser window displaying a Facebook post. The post features a video player with a play button and a caption: 'Formatura de Sargentos da Policia Militar do estado de São Paulo.' Below the video, it shows engagement metrics: 45 mil likes, 27 mil comments, and 15 mil shares. A red box highlights the bottom row of the table and the Facebook post, indicating the connection between the two.

Fonte: Elaborado pela autora, 2019.

A partir desta observação, que ocorre na mesma página, também é possível comparar os números de interações (curtidas, comentários, compartilhamento e visualizações) coletados com os números atuais, por exemplo.

No dia 03 de abril de 2019, foi realizado um registro gratuito de redirecionamento de domínio para o Protótipo no *No-IP Free Dynamic DNS*, que passa a ser nomeado como **Dumpres2018**. A sigla é uma conjunção da expressão inglesa *dump* com a palavra presidenciáveis, acrescida do ano das eleições (2018).

Assim, os dados poderão ser acessados, mesmo que migrando a base de dados pelo endereço eletrônico: **dumpres2018.redirectme.net**

5 CONSIDERAÇÕES FINAS

Compreendemos que, a partir do fenômeno da personalização na política, deu-se origem a inserção dos candidatos às eleições presidenciais de 2018 de maneira ampla e ativa nas principais mídias sociais. Demonstrando, assim, que existe uma busca pela proximidade dos candidatos com seus eleitores através da *Web*.

Contudo, em meio ao crescente domínio de mídias sociais, como o *Facebook*, ainda são escassos os projetos e iniciativas voltados para a preservação dos conteúdos publicados nessas plataformas. E, embora muito tenha se desenvolvido desde a criação do *Internet Archive*, datado de 1996, pouco ainda tem se discutido no contexto brasileiro. A própria preservação de mídias sociais a nível internacional tem se mostrado limitante e pouco abrangente.

Como constatamos durante a pesquisa, é muito difícil preservar conteúdos na íntegra publicados em mídias sociais como o *Facebook*, que em diversas tentativas de automatização da coleta impôs barreiras difíceis para um pesquisador transpor, possibilidades estas abertas apenas para agências publicitárias ou similares. Esta dualidade deixou evidente o real compromisso do *Facebook* com a parte comercial, desconsiderando os interesses social e histórico de seus usuários.

Para realizar a coleta foi necessário um período prévio de, aproximadamente, um mês para compreender como seriam extraídos os dados da página do candidato, bem como, estudar a melhor forma para estruturar as tabelas da base de dados e encontrar *scripts open source* que nos auxiliassem no desenvolvimento dos *bots*. Em suma, chegar nos resultados da pesquisa só foi possível através de uma aplicação prática.

Durante a coleta, pudemos observar alguns aspectos gerais que envolveram as publicações de *lives* dos candidatos. Em alguns casos, as publicações, na verdade, não continham uma comunicação ao vivo o que acabava descaracterizando a finalidade do recurso. Também identificamos a necessidade de aguardar o término da transmissão ao vivo para realizar a coleta com sucesso, algo que seria problemático em uma coleta totalmente automatizada. Ainda registramos a participação de todos os presidenciais de 2018 no Facebook, mesmo que em graus distintos, sendo efetuadas por todas as publicações audiovisuais – onde, ao menos, dois registros foram no formato de *live*, recurso este mais moderno dentro da plataforma e também em outras mídias sociais.

Também percebemos uma limitação da pesquisa, envolvendo a impossibilidade de coletar os comentários das publicações, decorrente da necessidade destes projetos serem executados em conjunto por um grupo de pesquisadores.

Não obstante, os dados vitais da pesquisa provêm da análise de disponibilidade do formato original das publicações coletadas. Este levantamento nos deu subsídios concretos para a defesa de que o arquivamento da *Web* e mídias sociais carece de ser ampliado e estudado nas suas peculiaridades e demandas singularidades. E, ainda, que a responsabilidade dos candidatos deve ser discutida em relação às publicações que ficam indisponíveis sem avisos e/ou explicações a seus eleitores.

Por outro lado, observamos que as três iniciativas de arquivamento da *Web* analisadas, *Internet Archive*, *Archive-It* e *Archive.is*, mesmo com esforços para preservar o conteúdo político brasileiro não deu conta de garantir a homogeneidade – privilegiando alguns candidatos em detrimento de outros. Outro déficit constatado, mas que atinge todos os candidatos, é a quase inexistência de vídeos preservados, não apenas do *Facebook*, mas de todas as mídias sociais verificadas.

Por se tratar de uma pesquisa exploratória, muitas alterações ocorreram para se chegar em objetivos tangíveis e significativos - o que nos mostrou outras possibilidades de pesquisa sob os mesmos dados e aplicações em outras mídias e tipos de conteúdo. A mais expressiva, envolvendo os vídeos coletados diretamente, é a possibilidade de analisar o conteúdo de cada vídeo a fim de extrair os discursos dos candidatos durante os períodos de campanha. Esse tipo de análise contribui para diversos campos de pesquisa, como a Ciência Política, Comunicação e Relações Públicas, História, entre outros, além da própria Ciência da Informação que pode aprofundar-se na epistemologia dessas práticas informacionais.

Os próprios estudos sobre a personalização na política e nas campanhas podem ser beneficiados com os dados coletados e as análises inferidas até o momento para futuros desdobramentos.

Todavia, a partir desta pesquisa, espera-se que a principal influência e interesse surja em forma de estudos mais aprofundados sobre as possibilidades de aplicação do arquivamento da *Web* e mídias sociais aos conteúdos brasileiros. E citamos como exemplo o caso do Orkut, uma mídia social lançada em 2004 e extinta em 2014, muito utilizada pelos brasileiros. Embora tenha sido comunicada a data de extinção e disponibilizado meios de *backup* de seus registros pelos usuários e administradores das comunidades, ainda se perdeu muito conteúdo. Este é um problema que afeta todas as plataformas comerciais que são movidas pela rentabilidade e nem sempre conseguem se manter competitivas frente a novas criações e, como solução, decidem interromper seus serviços e impossibilitar o acesso às informações.

Nossa perspectiva a curto prazo tem como objetivo realizar uma nova análise de disponibilidade do formato original dos 3.052 registros coletados no total, sendo 2.821 dos períodos de campanha eleitoral e 231 fora da campanha, a fim de traçar um comparativo evolutivo com a primeira verificação.

Não obstante, ainda há muitos questionamentos que podem guiar novas pesquisas como inflexões sobre de que modo podemos automatizar os processos de coleta em uma plataforma comercial, como o *Facebook*, que se mostra cada vez mais restritivo. E quais podem ser os impactos dessas restrições para a pesquisa e, ainda, quais são as implicações éticas que estas limitações podem estar sujeitas.

Por outro lado, também se faz necessária a reflexão de quais são as responsabilidades éticas que os candidatos devem ter ao publicar informações e depois removê-las sem aviso prévio a seus eleitores, tendo em vista que os rumos da política e da história são diretamente influenciados por essas interações. Qual é o compromisso que um candidato deve ter com os cidadãos ao publicar afirmações em mídias sociais?

Outros questionamentos que surgem desta pesquisa estão ligados ao nosso objetivo geral e indagam como a preservação dos conteúdos audiovisuais pode ser fomentada e quais são os fatores limitantes? Desta forma, entendemos que o fomento direcionado aos vídeos requer uma abordagem diferente, onde os registros configurem a prioridade e não a estrutura em que foram gerados, bem como a necessidade de compreender quais outros aspectos, além dos recursos, restringem a preservação desses vídeos.

Observou-se que o arquivamento de mídias sociais já ocorre, contudo, a forma como este arquivamento acontece ainda carece de mais reflexões, desenvolvimento e avanços.

As três plataformas de arquivamento analisadas, *Internet Archive*, *Archive-It* e *Archive.is*, mostraram que têm potencial para preservar as principais mídias sociais de maneira extensiva e superficial - isso significa que muitas informações não são contempladas. Na verdade, esse problema decorre por duas razões principais: a tentativa de se preservar a apresentação do formato original dos *links* e a carência de recursos (humanos, de hardware, de software e financeiro), os quais são necessários para realizar uma preservação intensiva.

Estes problemas serão recorrentes enquanto não forem desenvolvidos estudos profundos e aplicáveis sobre avaliação das informações da *Web* com base em sua produção distinta. Alguns aspectos que podemos observar é que estas informações sempre partem de um produtor, mas nem todas as publicações são de origem deste produtor – ou seja, algumas informações provêm de fontes secundárias e que não são, necessariamente, prioritárias na

preservação. Imaginemos então que estes registros, que são compartilhamentos dos produtores, podem ser identificados a partir dos metadados gerados pela publicação, possibilitando criar uma rotina de scripts que incluam este *sublink* em uma tabela de *links* secundários, por exemplo, da base de dados de preservação. E assim, com este filtro, seria possível direcionar os recursos para os *sublinks* primários.

Outras possibilidades envolvem a preservação que privilegia os registros e não obrigatoriamente a estrutura da origem, como o que aconteceu nesta pesquisa, onde preservamos prioritariamente os vídeos e, como adendo, os metadados das publicações. Este tipo de preservação, por ser mais específico, fornece resultados mais contextualizados e completos. No entanto, requer mais espaço de memória para o seu armazenamento que, embora os custos tenham diminuído no decorrer dos anos, ainda são mais proeminentes do que se fossem armazenados *snapshots*. Outro fator importante a ser considerado é a capacidade de processamento necessária ao servidor, que – quando disponibilizado amplamente o acesso aos registros, deve proporcionar um tempo de resposta rápido para não congestionar a plataforma. Isto pode ser atenuado quando a hospedagem dos dados é realizada em servidores comerciais ou institucionais de alta performance.

Mais uma reflexão que pode ser realizada discorre sobre a interrupção que alguns projetos de preservação da *Web* e mídias sociais sofrem por falta de capacidade de armazenamento, como o caso da Biblioteca Nacional dos EUA que deixou de preservar fontes do *Twitter*. Essas limitações tendem a restringir e determinar as políticas de seleção de iniciativas globais, desmotivando futuros projetos e carecendo de fomento dos estudos de seleção e avaliação a fim de desenvolver critérios que auxiliem na escolha de fontes de preservação da memória digital em qualquer mídia e formato.

Em suma, acreditamos que o desafio de desenvolver um protótipo para coletar os vídeos do *Facebook* dos candidatos à Presidência de 2018 durante a campanha eleitoral; a constatação de que 32 registros já estavam indisponíveis no seu formato original; a carência de registros preservados nas plataformas *Internet Archive*, *Archive-It* e *Archive.is* e o fato de desconhecermos projetos que preservem, na íntegra, registros audiovisuais brasileiros no contexto das eleições presidenciais de 2018 demonstram a emergência desta pesquisa. Assim como o potencial que outras pesquisas envolvendo os registros audiovisuais públicos oferecem no contexto da *e-política* (política na *Web*), onde os políticos abrem mão de uma privacidade na *Web* já decadente e assumem a exposição controlada através da personalidade que desejam transmitir aos cidadãos por meio das mídias sociais.

REFERÊNCIAS

ABEJÓN, Paloma; SASTRE, Ana; LINARES, Virginia. *Facebook y Twitter en campañas electorales en España*. **Anuario electrónico de estudios en comunicación social**. Vol. 5, Número 1 / jan-jun, 2012.

ARQUIVO.PT. Serviço de pesquisa sobre o passado português. Disponível em: <http://arquivo.pt/>. Acesso em: abr. de 2018.

BAILEY, Lila. Fair Use & Access to All Human Knowledge. *Internet Archive Blogs*. 24 fev. 2016. Disponível em: <https://blog.archive.org/2016/02/24/fair-use-access-to-all-human-knowledge/>. Acesso em: nov. de 2018. Versão arquivada em: <http://archive.today/NklaQ>.

BELL, Karissa. *Facebook is finally bringing live streaming to everyone*. **Mashable**. 28 jan. 2016. Disponível em: <https://mashable.com/2016/01/28/Facebook-live-video-how-to/#aruPzsvCVgqc>. Acesso em: abr. de 2018. Versão arquivada em: <http://archive.today/Hla2N>.

BLUME, Bruno André. Como é eleito o presidente da república? **Politze**. 20 jun. 2017. Disponível em: <https://www.politize.com.br/presidente-da-republica-como-e-eleito/>. Acesso em: nov. de 2018. Versão arquivada em: <http://archive.today/D9iQG>.

BRASIL. **Lei nº 4.117, de 27 de agosto de 1962**. Institui o Código Brasileiro de Telecomunicações. Brasília, [1962]. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/14117.htm#art71. Acesso em: nov. de 2018.

BRASIL. **Lei nº 12.527, de 18 de novembro de 2011**. Regula o acesso a informações [...]. Brasília, [2011]. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm. Acesso em: nov. de 2018.

BRASIL. **Lei nº 12.965, de 23 de abril de 2014**. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Brasília, [2014]. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm. Acesso em: nov. de 2018.

BRONSTEIN, Jenny. Like me! Analyzing the 2012 presidential candidates' *Facebook* pages. **Online Information Review**, v. 37, n. 2, p. 173-192, 2013.

BROWN, Adrian. **Archiving Websites**: a practical guide for information management professionals. Facet publishing, London, 2006.

BRÜGGER, Niels. *Web archiving: Between past, present, and future*. **The handbook of Internet studies**, p. 24-42, 2011.

CASHMORE, Pete. *Facebook Video Launches: Youtube Beware!* **Mashable**. 24 maio 2007. Disponível em: <https://mashable.com/2007/05/24/Facebook-video-launches/#7wK01FjdHGqg>. Acesso em: nov. de 2018.
Versão arquivada em: <http://archive.today/O54G2>.

CASTELLS, Manuel. **A Sociedade em Rede**. A Era da Informação: Economia, Sociedade e Cultura. São Paulo: Paz e Terra, 6ª ed., vol. 1, 2013.

CASTELLS, MANUEL. **O poder da comunicação**. 2ª ed. Rio de Janeiro/São Paulo: Paz e Terra, 2017.

COSTA, Miguel; GOMES, Daniel; SILVA, Mário J. The evolution of *Web* archiving. **International Journal on Digital Libraries**, v. 18, n. 3, p. 191-205, 2017.

CROOK, Edgar. *Web* archiving in a *Web* 2.0 world. **The Electronic Library**, v. 27, n. 5, p. 831-836, 2009.

DAY, Michael. The long-term preservation of *Web* content. In: **Web archiving**. Springer, Berlin, Heidelberg, 2006. p. 177-199.

Dicionário Eletrônico de Terminologia em Ciência da Informação (DelTCI). Observatório de Ciência da Informação da Universidade do Porto. Disponível em: <https://paginas.fe.up.pt/~lci/novo/index.php/investigar/deltci/itemlist/category/17-deltci>. Acesso em: maio de 2018.

ENLI, Gunn Sara; SKOGERBØ, Eli. Personalized Campaigns In Party-Centred Politics. **Information, Communication & Society**. 2013. DOI:10.1080/1369118X.2013.782330

FACEBOOK. **Políticas de páginas, grupos e eventos**. [s/d]. Disponível em: https://www.facebook.com/policies/pages_groups_events/. Acesso em: nov. de 2018.
Versão arquivada em: <http://archive.fo/nvzR1>.

FACEBOOK. **Como solicitar um selo de verificação azul?** [s/d]. Disponível em: https://www.facebook.com/help/1288173394636262?helpref=faq_content. Acesso em: nov. de 2018.
Versão arquivada em: http://archive.today/bCx_fm.

FERREIRA, Lisiane Braga; MARTINS, Marina Rodrigues; ROCKEMBACH, Moisés. Usos do Arquivamento da *Web* na Comunicação Científica. **Prisma.com** (36) 2018, p. 78-98. DOI 10.21747/16463153/36A5

FERREIRA, Miguel. Introdução à preservação digital: conceitos, estratégias e actuais consensos. Universidade do Minho, Escola de Engenharia, 2006.

FIND YOUR FACEBOOK ID. Disponível em: <https://findmyfbid.com/>. Acesso em: abr. de 2018.
Versão arquivada em: <http://archive.today/OYRhs>.

FRIEDMAN, Thomas L. Thank you for being late: An optimist's guide to thriving in the age of accelerations. Farrar, Straus and Giroux, 2016.

GAUTIER, Thiago. "Memórias" da *Internet* correm o risco de se perderem com o tempo. **Revista Galileu**. 07 mar. 2018. Disponível em:

<https://revistagalileu.globo.com/revista/noticia/2018/03/memorias-da-internet-correm-o-risco-de-se-perderem-com-o-tempo.html>. Acesso em: maio de 2018.

Versão arquivada em:

<https://web.archive.org/web/20180307232953/https://revistagalileu.globo.com/revista/noticia/2018/03/memorias-da-internet-correm-o-risco-de-se-perderem-com-o-tempo.html>.

Gil, Antônio Carlos. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 4. ed., 2002.

GHFLY. Publicação no *Facebook*, impulsionamento de postagens e anúncios de *Facebook Ads*: entenda a diferença entre eles! **GhFly blog**. 27 jun. 2017. Disponível em:

<http://blog.ghfly.com/publicacao-impulsionamento-facebook-ads-entenda-diferencas/>.

Acesso em: nov. de 2018.

Versão arquivada em: <http://archive.today/YlcrN>.

GRAY, Gabriella; MARTIN, Scott. Choosing a sustainable *Web* archiving method: A comparison of capture quality. **D-Lib Magazine**, v. 19, n. 5, p. 2, 2013.

HERITRIX. Wiki. Disponível em: <https://webarchive.jira.com/wiki/spaces/heritrix/>. Acesso em: jun. de 2018.

HOCKX-YU, Helen; CRAWFORD, Lewis; CORAM, Roger; JOHNSON, Stephen. **Capturing and replaying streaming media in a *Web* archive – a british library case study**. The British Library Board, 2010. Disponível em:

<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.232.3180>. Acesso em: mar. de 2018.

HTRACK. *Website copier*. Disponível em: <https://www.httrack.com/>. Acesso em: jun. de 2018.

INTERNATIONAL INTERNET PRESERVATION CONSORTIUM - IIPC. *Website*. Disponível em: <http://netpreserve.org/about-us/>. Acesso em: jun. de 2018.

Versão arquivada em:

<https://web.archive.org/web/20180615122914/http://netpreserve.org/about-us/>.

INTERNET ARCHIVE. Internet Archive Projects. Disponível em:

<https://archive.org/projects/>. Acesso em: jun. de 2018.

INTERNET MEMORY FOUNDATION. Living Web Archives (LiWA). Disponível em: <http://internetmemory.org/en/index.php/projects/liwa>. Acesso em: jun. de 2018.

KIRKPATRICK, David. O efeito *Facebook*. Editora Intrínseca, 2011.

LEV-ON, Azi; HALEVA-AMIR, Sharon. Normalizing or equalizing? Characterizing *Facebook* campaigning. **New media & society**, p. 1461444816669160, 2016.

LIBRARY OF CONGRESS. Brazilian Presidential Election 2010 *Web Archive*. Disponível em: <https://www.loc.gov/collections/brazilian-presidential-election-2010-web-archive/about-this-collection/>. Acesso em: jun. de 2018.

LOMBORG, Stine. Researching communicative practice: *Web archiving* in qualitative social media research. **Journal of Technology in Human Services**, v. 30, n. 3-4, p. 219-231, 2012.

MANIN, Bernard. As Metamorfoses do Governo Representativo, In: **Revista Brasileira de Ciências Sociais**, nº 29, outubro de 1995.

MÁRDERO ARELLANO, Miguel A. **Critérios para a preservação digital da informação científica**. 2008. 354 f. 2008. Tese de Doutorado. Tese (Doutorado em Ciência da Informação) – Universidade de Brasília, Departamento de Ciência da Informação, Brasília.

MASANÈS, Julien. **Web Archiving**. Springer, Berlin, Heidelberg, 2006.

MILLIGAN, Ian; RUEST, Nick; LIN, Jimmy. Content selection and curation for *Web archiving*: The gatekeepers vs. the masses. In: **Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries**. ACM, 2016. p. 107-110.

MORAES, Igor. Veja quem são os candidatos a presidente da República nas eleições 2018. **Estadão**. 27 jul. 2018. Disponível em: <https://politica.estadao.com.br/noticias/eleicoes,veja-quem-quer-ser-presidente-em-2018,70002054149>. Acesso em: nov. de 2018. Versão arquivada em: <http://archive.today/RHcTZ>.

NIU, Jinfang. An overview of *Web archiving*. **D-Lib magazine**, v. 18, n. 3/4, 2012.

Núcleo de Pesquisa em Arquivamento da *Web* e Preservação Digital (NUAWEB). Disponível em: <https://www.ufrgs.br/nuaweb/>. Acesso em: set. de 2018.

OPEN SOURCE INITIATIVE. About. Disponível em: <https://opensource.org/>. Acesso em: jun. de 2018.

OXFORD LIVING DICTIONARIES. *Live streaming*. Disponível em: https://en.oxforddictionaries.com/definition/us/live_stream. Acesso em: jun. de 2018.

PAVÃO, Caterina Marta Groposo; CAREGNATO, Sônia Elisa; ROCHA, Rafael Port da. Implementação da preservação digital em repositórios: conhecimento e práticas. **Revista Digital de Biblioteconomia e Ciência da Informação**. Campinas, SP. Campinas. Vol. 14, n. 3 (set./dez. 2016), p. 407-425, 2016.

POP, Radu; VASILE, Gabriel; MASANÈS, Julien. Archiving *Web* video. In: **International Web Archiving Workshop IWAW 2010**. 2010.

ROSS, Karen; FOUNTAINE, Susan; COMRIE, Margie. Facing up to *Facebook*: politicians, publics and the social media (ted) turn in New Zealand. **Media, Culture & Society**, v. 37, n. 2, p. 251-269, 2015.

RUEST, Nick; MILLIGAN, Ian. An open-source strategy for documenting events: The case study of the 42nd Canadian federal election on Twitter. 2016.

RUTHES, Jefter. Facebook: comunidade de ajuda. **Faceajuda**. 2016. Disponível em: <<https://www.facebook.com/help/community/question/?id=463868263726424>>. Acesso em: nov. de 2018.

Versão arquivada em: <http://archive.today/laGOF>.

SALAHDELDEEN, Hany M.; NELSON, Michael L. Losing my revolution: How many resources shared on social media have been lost?. In: **International Conference on Theory and Practice of Digital Libraries**. Springer, Berlin, Heidelberg, 2012. p. 125-137.

SCHNEIDER, Steven M. et al. Building thematic *Web* collections: challenges and experiences from the September 11 *Web* Archive and the Election 2002 *Web* Archive. **Digital Libraries**, ECDL, p. 77-94, 2003.

SCHOLARS PORTAL DATAVERSE. #elxn42 tweets (42nd Canadian Federal Election). 2015. Disponível em: <https://hdl.handle.net/10864/11310>. Acesso em: jun. de 2018.

SILVA, Armando Malheiro da. A informação. Da compreensão do fenómeno e construção do objecto científico. Porto: Edições Afrontamento, 2006.

SILVA, Armando Malheiro da. et al. **Arquivística: teoria e prática de uma ciência da informação**. Vol. 1, 3ª ed. Porto: Ed.: Afrontamento. 2009. 254p.

SOUZA, Arthur Heleno Lima Rodrigues de et al. O modelo de referência OAIS e a preservação digital distribuída. **Ciência da Informação**, v. 41, n. 1, 2012.

SUMMERS, Ed; PUNZALAN, Ricardo. Bots, seeds and people: Web archives as infrastructure. In: **Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing**. ACM, 2017. p. 821-834.

SUPPORT ARCHIVE-IT. Archiving Facebook. Disponível em: <https://support.archive-it.org/hc/en-us/articles/208333113-Archiving-Facebook>. Acesso em: jun. de 2018.

TECHTARGET. WhatIs.com. **Facebook live**. Disponível em: <https://whatistechtarget.com/definition/Facebook-Live>. Acesso em: abr. de 2018.

Versão arquivada em: <http://archive.today/L5mLW>.

TOYODA, Masashi; KITSUREGAWA, Masaru. The history of *Web* archiving. **Proceedings of the IEEE**, v. 100, n. Special Centennial Issue, p. 1441-1443, 2012.

TRIBUNAL SUPERIOR ELEITORAL. **Confira as principais datas do calendário eleitoral das Eleições Gerais de 2018**. [Brasília: TSE], 2017. Disponível em: <http://www.tse.jus.br/imprensa/noticias-tse/2017/Dezembro/confira-as-principais-datas-do-calendario-eleitoral-das-eleicoes-gerais-de-2018>. Acesso em: abr. de 2018.

TRIBUNAL SUPERIOR ELEITORAL. **Lei das Eleições – Lei nº 9.504, de 30 de setembro de 1997**. [Brasília: TSE]. Disponível em: <http://www.tse.jus.br/legislacao/codigo-eleitoral/lei-das-eleicoes/lei-das-eleicoes-lei-nb0-9.504-de-30-de-setembro-de-1997>. Acesso em: abr. de 2018.

TRIBUNAL SUPERIOR ELEITORAL. **Normas e documentações — Eleições 2018.** [Brasília: TSE]. Disponível em: <http://www.tse.jus.br/eleicoes/eleicoes-2018/normas-e-documentacoes-eleicoes-2018>. Acesso em: abr. de 2018.

TRIBUNAL SUPERIOR ELEITORAL. **Partidos políticos registrados no TSE.** [Brasília: TSE]. Disponível em: <http://www.tse.jus.br/partidos/partidos-politicos/registrados-no-tse>. Acesso em: mar. de 2019.

TRIBUNAL SUPERIOR ELEITORAL. **Resolução nº 23.551, de 18 de dezembro de 2017.** [Brasília: TSE]. Disponível em: <http://www.tse.jus.br/legislacao-tse/res/2017/RES235512017.html>. Acesso em: nov. de 2018.

VALENTE, Jonas. Redes sociais perdem espaço como fonte de notícia, diz relatório global. Agência Brasil. Disponível em: <http://agenciabrasil.ebc.com.br/geral/noticia/2018-06/redes-sociais-perdem-espaco-como-fonte-de-noticiadiz-relatorio-global>. Acesso em: nov. de 2018. Versão arquivada em: <http://archive.today/FXSvp>.

VOERMAN, Gerrit et al. Archiving the *Web*: Political party *Web sites* in the Netherlands. **European Political Science**, v. 2, n. 1, p. 68-75, 2002.

WEARE, Christopher; LIN, Wan-Ying. Content analysis of the world wide *Web*. **Social Science Computer Review**, vol. 18 no. 3, 2000, pp 272-292.

APÊNDICE A – OBJECT INFORMATION DO DB ARCHIVE – TBL_ALVO

```
CREATE TABLE `tbl_alvo` (  
  `seq_alvo` int(11) NOT NULL AUTO_INCREMENT,  
  `id_alvo` varchar(100) DEFAULT NULL,  
  `page_alvo` varchar(300) DEFAULT NULL,  
  `nome_pasta` varchar(200) DEFAULT NULL,  
  `nome_alvo` varchar(200) DEFAULT NULL,  
  `pagina_video` varchar(500) DEFAULT NULL,  
  PRIMARY KEY (`seq_alvo`)  
) ENGINE=InnoDB AUTO_INCREMENT=21 DEFAULT CHARSET=latin1;
```

APÊNDICE B – OBJECT INFORMATION DO DB ARCHIVE – TBL_ARCHIVE

```

CREATE TABLE `tbl_archive` (
  `seq_post` int(11) NOT NULL AUTO_INCREMENT,
  `forcado` varchar(3) DEFAULT 'N',
  `id` varchar(50) DEFAULT NULL,
  `nome_alvo` varchar(100) DEFAULT NULL,
  `live` char(10) DEFAULT 'N',
  `id_post` varchar(100) DEFAULT NULL,
  `ativo_na_fonte` varchar(3) DEFAULT NULL,
  `ativo_na_fonte_090319` varchar(3) DEFAULT NULL,
  `curtidas` varchar(20) DEFAULT NULL,
  `num_comentarios` varchar(20) DEFAULT NULL,
  `num_compartilhamentos` varchar(20) DEFAULT NULL,
  `num_visualizacoes` varchar(20) DEFAULT NULL,
  `created_time` datetime DEFAULT NULL,
  `data_postagem` datetime DEFAULT NULL,
  `data_utime` varchar(20) DEFAULT NULL,
  `abbr_title` varchar(20) DEFAULT NULL,
  `message` varchar(5000) DEFAULT NULL,
  `embeded` varchar(5000) DEFAULT NULL,
  `pasta` varchar(1000) DEFAULT NULL,
  `nome_original` varchar(200) DEFAULT NULL,
  `tempo_duracao` varchar(100) DEFAULT NULL,
  `dt_ultima_atl` datetime DEFAULT NULL,
  `exec` char(1) DEFAULT 'S',
  PRIMARY KEY (`seq_post`),
  UNIQUE KEY `id1` (`id_post`) USING BTREE,
  KEY `id0` (`id`) USING BTREE
) ENGINE=InnoDB AUTO_INCREMENT=25988 DEFAULT CHARSET=latin1
ROW_FORMAT=DYNAMIC;

```

APÊNDICE C – OBJECT INFORMATION DO DB ARCHIVE – TBL_LOG

```
CREATE TABLE `tbl_log` (  
  `seq_log` int(11) NOT NULL AUTO_INCREMENT,  
  `seq_post` int(11) DEFAULT NULL,  
  `id` varchar(50) DEFAULT NULL,  
  `nome_alvo` varchar(100) DEFAULT NULL,  
  `url` varchar(300) DEFAULT NULL,  
  `nova_url` varchar(300) DEFAULT NULL,  
  `curtidas` int(11) DEFAULT NULL,  
  `comentarios` int(11) DEFAULT NULL,  
  `compartilhamentos` int(11) DEFAULT NULL,  
  `visualizacoes` int(11) DEFAULT NULL,  
  `data_log` timestamp NULL DEFAULT NULL ON UPDATE  
CURRENT_TIMESTAMP,  
  PRIMARY KEY (`seq_log`),  
  KEY `id` (`url`)  
) ENGINE=InnoDB AUTO_INCREMENT=289867 DEFAULT CHARSET=latin1;
```

APÊNDICE D – SCRIPT DA PÁGINA DE INCLUSÃO DO PROTÓTIPO

```

<?
error_reporting(E_ERROR);
ini_set('display_errors', 0);
header("Content-type:text/html; charset=iso8859-1");
$connect=odbc_connect('archive','Web','pi');
$sql="SELECT * from tbl_alvo order by nome_alvo";
$result=odbc_exec($connect,$sql);
?>
<html xmlns="http://www.w3.org/1999/xhtml" lang="pt-br"><head>
<link rel="stylesheet" type="text/css" href="estilo_css.css">
<title>Upload Arquivos</title>
<style>
    body {
        font-family: Tahoma;
        font-size: 9pt;
        color: #000000;
        background: #CFCFCF;
        margin: 2px;
    }
    input {
        font-family: Tahoma;
        font-size: 9pt;
        text-align: center;
        border: 1px solid;
        !width: 20px;
        background: #a0a0a0;
    }
</style>
<script>
function nome(a,b){
//alert(b);
document.getElementById(b).innerHTML=a;
}
</script>
</head>
<body>

<h3>Adicionar Metadados</h3>
<form name="dump" method="POST" action="dump.php" enctype="multipart/form-data">

<table>
<input type='reset' value='Limpar todos os campos'/>
<p><label>Nome Candidato</label>
<select class='w3-select w3-border' name='id' id='id' >
<option value="" disabled selected >Selecione</option>
<?
while(odbc_fetch_row($result)){

```

```

$nome_alvo = odbc_result($result,'nome_alvo');
$id = odbc_result($result,'id_alvo');
echo "<option value='$id'>$nome_alvo</option>";
}
?>
</select>
<tr><td>url do video:<input type="text" name="url" id="url" size=100 ></td></tr>
<tr><td>Mensagem<textarea name='mensagem' id='mensagem' rows='5' cols='70'
></textarea></td></tr>
<tr><td>live(S/N)

<select class='w3-select w3-border' name='live' id='live' >
    <option value='N'>Nao</option>
    <option value='S'>Sim</option>
</select></td></tr>
</table>
<input type="submit" value="Enviar">
<tr>
<td>
<span id='sp_resp'></span>
</td>
</tr>
</form>
</body>
</html>

```


APÊNDICE E – SCRIPT DO BOT-SCRAPING DO PROTÓTIPO

```

<?
error_reporting(E_ERROR);
ini_set('display_errors', 0);
$p=explode("/", $url);
$numero=$p[5];
$nome_arquivo = $numero . ".mp4";
//$handle = fopen("https://beamtic.com/", "rb");

$opts = array (
    'http' => array (
        'method' => "GET",
        'header'=>"Accept-language: en\r\n",
        'user_agent' => "Mozilla/5.0 (Windows NT 6.1; WOW64; rv:52.0) Gecko/20100101
Firefox/52.0",
    )
);

$context = stream_context_create($opts);

//ponteiro que aponta para o link do video fornecido como input
$handle = fopen($url, "rb",0,$context);

//variável que sera usada para armazenar o streaming de metadados a ser analisado on-the-fly
$content = "";

while (!feof($handle)) {
    $content = $content . fread($handle, 8192);
}
// Processamento dos metadados
// buscas são feitas no código para identificar onde iniciam determinados pontos de partida e
finalização de dados específicos

//conexao ao db
$conn=odbc_connect('archive','Web','pi');
$canonical='link rel="canonical" href=';
$data_utime='data-utime';
$abbr_title='abbr title';

//ownerid
$search="ownerid:";
$x= strpos($content, $search);
$s=substr($content,$x+8,25);
$e=explode(",",$s);
$ownerid=$e[0];
$ownerid=str_replace("'",",",$ownerid);

//data-utime

```

```

$search="abbr title=";
$x= stripos($contents, $search);
$s=substr($contents,$x+12,17);
//$e=explode(",",$s);
$abbr_title=$s;

//data-utime
$search="data-utime=";
$x= stripos($contents, $search);
$s=substr($contents,$x+12,10);
//$e=explode(",",$s);
$data_utime=$s;

$sql="select nome_pasta,nome_alvo from tbl_alvo where id_alvo=$ownerid";
$result=odbc_exec($connect,$sql);
$nome_pasta=odbc_result($result,'nome_pasta');
$nome_alvo=odbc_result($result,'nome_alvo');
print "nome_pasta=". $nome_pasta ."<br>";
print "nome_arquivo=". $nome_arquivo ."<br>";

//curtidas
$search="likecount:";
$x= stripos($contents, $search);
$s=substr($contents,$x+10,10);
$e=explode(",",$s);
$curtidas=$e[0];

//comentarios
$search="filtered:{range:{offset:";
$search="commentcount:";
$x= stripos($contents, $search);
$s=substr($contents,$x+13,10);
$e=explode(",",$s);
$comentarios=$e[0];

if($comentarios=='ng="en" id'){ $comentarios=0;}

//compartilhamento
$search="sharecount:";
$x= stripos($contents, $search);
$s=substr($contents,$x+11,10);
$e=explode(",",$s);
$share=$e[0];

//visualizacoes
$search="viewCount:";
$x= stripos($contents, $search);
$s=substr($contents,$x+10,10);
$e=explode(",",$s);
$view=$e[0].$e[1];

```

```

$view=str_replace("",",$view);
$view=str_replace('view',$view);

//Tempo de duracao
$search="Duration: ";
$x= stripos($contents, $search);
$dur=substr($contents,$x+10,20);
$e=explode("", $dur);
$duracao=$e[0];

//url - sd
$search="sd_src_no_ratelimit:";
$inicio= stripos($contents, $search);
$x= stripos($contents, $search);
$search=',aspect_ratio:~';
$fim= stripos($contents, $search);
$comp=$fim-$inicio;
$busca=substr($contents,$inicio+20,$comp);

$e=explode(",",$busca);
$link0=$e[0];
$link=str_replace("",",$link0);

///assunto
$search='<meta property="og:description" content="';
$comeco='userContent_3576" data-ft="&#123;&quot;tn&quot;:&quot;K&quot;&#125';

$inicio= stripos($contents, $comeco);
$search='action="/ajax/ufi/modify.php';
$fim= stripos($contents, $search);
$comp=$fim-$inicio;
$comp=1000;
$busca=substr($contents,$inicio+70,$comp);
$e=explode(",",$busca);
$assunto=utf8_decode($e[0]);
$assunto=trim(str_replace('n":"K"}">',",$assunto));
/////

if(strlen($assunto)==0){ $assunto='-';}

print "Comprimento em Bytes:=". strlen($contents) . "<br>";
print "ownerid=". $ownerid . "<br>";
print "Data postagem=". $data_otime . "<br>";
print "abbr_title=". $abbr_title . "<br>";
print "curtidas=". $curtidas . "<br>";
print "comentarios=". $comentarios . "<br>";
print "compartilhamentos=". $share . "<br>";
print "visualizacoes=". $view . "<br>";
print "link=". $link . "<br>";
print "duracao do video=". $duracao . "<br>";

```

```

print "assunto=". $assunto ."<br>";
print "live?=". $live ."<br>";

print "assunto2=". $mensagem ."<br>";
if(strlen($mensagem)==0){ $mensagem=$assunto;}
fclose($handle);

// download do video da url capturada

$handle2 = fopen($link, "rb",0,$context);
$path = '/var/www/htdocs/sistemas/archive/candidatos/'. $nome_pasta .'/'.$nome_arquivo;
$fp = fopen($path,'w');

$content2 = "";
while (!feof($handle2)) {
    $content2 = fread($handle2, 8192);
    $ret = fwrite($fp,$content2,8192);
}
fclose($handle2);
fclose($fp);

$date = date("Y-m-d H:i:s");
$p=explode("/", $url);
$numero=$p[5];

$iframe='<iframe
src="https://www.Facebook.com/plugins/video.php?href=https://www.Facebook.com/'.$nome_alvo.'/videos/'.$numero.' /&show_text=1&width=476" width="476" height="646"
style="border:none;overflow:hidden" scrolling="no" frameborder="0"
allowTransparency="true" allow="encrypted-media" allowFullScreen="true"></iframe>';

$nome_arquivo = $numero . ".mp4";
$pasta="https://server489.seedhost.eu/sistemas/archive/candidatos/$nome_pasta/$nome_arquivo";

$sql = "INSERT INTO tbl_archive (id, id_post,
nome_alvo,created_time,data_utime,abbr_title, pasta,message,embeded,
nome_original,tempo_duracao,curtidas,num_comentarios,num_compartilhamentos,num_visualizacoes,live)
VALUES
('$ownerid','$url','$nome_alvo','$date','$data_utime','$abbr_title','$pasta','$mensagem','$iframe',
'$nome_arquivo','$duracao','$curtidas','$comentarios','$share','$view','$live')";

$sql = "INSERT INTO tbl_archive (id, id_post,
nome_alvo,created_time,pasta,message,embeded, nome_original,live)
VALUES
('$ownerid','$url','$nome_alvo','$date','$pasta','$mensagem','$iframe','$nome_arquivo','$live');

```

```

//$incluir = "INSERT INTO tbl_archive (id, id_post,
nome_alvo,created_time,data_utime,abbr_title, pasta,message,embebed,
nome_original,tempo_duracao,curtidas,num_compartilhamentos,num_visualizacoes)
//      VALUES
('$ownerid','$url','$nome_alvo','$date','$data_utime','$abbr_title','$pasta','$mensagem','$iframe'
,'$nome_arquivo','$duracao','$curtidas','$share','$view')";

$result = odbc_exec($connect, $incluir);
$sup="update tbl_archive set
data_postagem=DATE_ADD(FROM_UNIXTIME(data_utime,'% Y-%m-%d %H:%i:%s'),
INTERVAL '-4' DAY_HOUR) where id_post='$url'";
$rs = odbc_exec($connect, $sup);

//mostra o que foi postado
print '<input type="button" value="Voltar" onClick="javascript: history.back();">';
print "<br>";
print "<table style='width: 100%;' border=0 cellpadding=5 cellspacing=0
bgcolor=#000000>";
print "<tr bgcolor=#a0c0f0>";
print "<input text='hidden' name=id_post id=id_post value='$url' READONLY>";
print "<td width=60px>Nome</td>";
print "<td width=60px>Criado</td>";
print "<td width=320px>Mensagem</td>";
print "<td width=60px>Link_video</td>";
print "<td width=60px>Frame</td>";
print "<td width=60px>Atualizar</td>";
print "</tr>";
$consulta="select * from tbl_archive where id_post='$url'";
$result=odbc_exec($connect,$consulta);

while(odbc_fetch_row($result)) {
$face="";
$nome_alvo = odbc_result($result,'nome_alvo');
$id= odbc_result($result,'id');
$id_post = odbc_result($result,'id_post');
$curtidas = odbc_result($result,'curtidas');
$num_comentarios = odbc_result($result,'num_comentarios');
$num_compartilhamentos = odbc_result($result,'num_compartilhamentos');
$num_visualizacoes = odbc_result($result,'num_visualizacoes');
$created_time = odbc_result($result,'created_time');
$data_criacao= date("d-m-Y H:i",strtotime($created_time));
$mensagem = odbc_result($result,'message');
$nome_arquivo=odbc_result($result,'nome_original');
$link=odbc_result($result,'pasta');
$seq_post= odbc_result($result,'seq_post');
//$frame=odbc_result($result,'embebed');
$atualizar="https://server489.seedhost.eu/sistemas/archive/atualizar.php?id_post=$id_post";

$css1="#e0e0e0";
if($j==0) { $css1="#ededed"; };

```



```

if(xmlHttp2.readyState==4 || xmlHttp2.readyState=="complete") {
document.getElementById(dest).innerHTML+xmlHttp2.responseText
} else { document.getElementById(dest).innerHTML='atualizando ' }

```

```

xmlHttp2=GetXmlHttpObject()
if(xmlHttp2==null) {
    alert("Browser nao suporta requisicao HTTP")
    return
}

```

```

id_post=document.getElementById('id_post').value;
seq_post=document.getElementById('seq_post').value;
data_postagem=document.getElementById('data_postagem').value;
tempo=document.getElementById('tempo').value;
curtidas=document.getElementById('curtidas').value;
comentarios=document.getElementById('comentarios').value;
compartilhamentos=document.getElementById('compartilhamentos').value;
visualizacoes=document.getElementById('visualizacoes').value;
var url="up_reaction.php";
url=url+"?curtidas=" + curtidas;
url=url+"&comentarios=" + comentarios;
url=url+"&compartilhamentos=" + compartilhamentos;
url=url+"&visualizacoes=" + visualizacoes;
url=url+"&id_post="+ id_post;
url=url+"&seq_post="+ seq_post;
url=url+"&data_postagem="+ data_postagem;
url=url+"&tempo="+ tempo;
url=url+"&sid="+Math.random()
//alert(url);
xmlHttp2.onreadystatechange=stateChanged
xmlHttp2.open("GET",url,true)
xmlHttp2.send(null)}
</script>

```

APÊNDICE F – SCRIPT DA PÁGINA DE CONSULTA DO PROTÓTIPO

```

<?
error_reporting(E_ERROR);
ini_set('display_errors', 0);
header("Content-type:text/html; charset=iso8859-1");
$connect=odbc_connect('archive','Web','pi');
$sql="SELECT * from tbl_alvo order by nome_alvo";
$result=odbc_exec($connect,$sql);

if(!isset($protocolo)) {      $proto1 = ";} else {$proto1 = $protocolo;}
?>
<html>
  <head>
    <title></title>
<link href="estilo_css_novo.css" rel="stylesheet" type="text/css"/>
    <style>
      td.t1 {
        font-face: arial;
        font-size: 8pt;
      }
    </style>
  </head>
<body>
<table style='width: 1600px; margin-left: auto; margin-right: auto; height=100% >

<tr>
<td style='text-align: left; width: 25%; vertical-align: top;'></td>
<td style='height: 95px;' align=left ><br>
<h3>Protótipo - Sistema de Arquivamento da WEB</h3></td>
<td colspan=2><span id="resp_consulta2" style=' width: 400px;'></span></td>
</tr>
<tr>
<td height=15 colspan=3 align=left>
<form id=f1 action='consulta_candidato.php' method=post target=if_consulta >
  <p><label>Nome Candidato</label>

<select class='w3-select w3-border' name='id' id='id' ><option value=" disabled selected
>Selecione</option>
<option value="TODOS">Todos</option>
<?
while(odbc_fetch_row($result)){
$nome_alvo = odbc_result($result,'nome_alvo');
$id = odbc_result($result,'id_alvo');
echo "<option value='$id'>$nome_alvo</option>";
} ?>
</select>

<? $data_fim= date("d/m/Y");?>

```



```
entre <input type=text name='data_ini' id='data_ini' value='16/08/2018' size=12 > e <input
type=text name='data_fim' id='data_fim' value=<?=$data_fim?> size=12 >
```

```
<label>live?(S/N)</label>
<select class='w3-select w3-border' name='elive" id='elive" >
<option value='T'>Todos</option>
<option value='S'>Sim</option>
<option value='N'>Não</option>
</select>
<input type=hidden name=fechar value=0>
<input type=submit value=consultar>
</form>
</td>
</tr>
<tr>
<td style='text-align: left;' colspan=3>
<iframe id=if_consulta name=if_consulta style='width: 100%; height: 100%; border: 0px
solid; padding: 1px 1px 1px 1px;' ></iframe></td>
</tr>
</table>
</body>
<script>
</html>
```

APÊNDICE G – LISTA DOS *LINKS* DAS MÍDIAS SOCIAIS E *SITES* DOS CANDIDATOS À PRESIDÊNCIA DE 2018

| FONTES | CANDIDATOS |
|---------------------------|---|
| Álvaro Dias | |
| <i>Facebook</i> | https://www.Facebook.com/ad.alvarodias/ |
| <i>Twitter</i> | https://Twitter.com/alvarodias |
| <i>Youtube</i> | https://www.Youtube.com/user/alvarodias45 |
| <i>Instagram</i> | https://www.Instagram.com/ad.alvarodias/ |
| <i>Site</i> | https://alvarodias.com.br |
| Cabo Daciolo | |
| <i>Facebook</i> | https://www.Facebook.com/depudadocabodaciolo/ |
| <i>Twitter</i> | https://Twitter.com/cabodaciolo |
| <i>Youtube</i> | https://www.Youtube.com/channel/UCdtc9CCVxLBVkN51Lvy_wTA |
| <i>Instagram</i> | https://www.Instagram.com/cabodaciolo/ |
| <i>Site</i> | http://patriota51.org.br/ |
| Ciro Gomes | |
| <i>Facebook</i> | https://www.Facebook.com/cirogomesoficial/ |
| <i>Twitter</i> | https://Twitter.com/cirogomes |
| <i>Youtube</i> | https://www.Youtube.com/cirogomesoficial |
| <i>Instagram</i> | https://www.Instagram.com/cirogomes/ |
| <i>Site</i> | https://www.cirogomes.com.br/ |
| Fernando Haddad | |
| <i>Facebook</i> | https://www.Facebook.com/fernandohaddad/ |
| <i>Twitter</i> | https://Twitter.com/haddad_fernando |
| <i>Youtube</i> | https://www.Youtube.com/channel/UCyL3IelkOB9yI7XvgTYkc0A |
| <i>Instagram</i> | https://www.Instagram.com/fernandohaddadoficial/ |
| <i>Site</i> | https://haddadpresidente.com.br/ |
| Geraldo Alckmin | |
| <i>Facebook</i> | https://www.Facebook.com/geraldoalckmin/ |
| <i>Twitter</i> | https://Twitter.com/geraldoalckmin |
| <i>Youtube</i> | https://www.Youtube.com/channel/UCNxCni0Iv9pr7i_pQZ6ijlg |
| <i>Instagram</i> | https://www.Instagram.com/geraldoalckmin / |
| <i>Site</i> | https://www.geraldoalckmin.com.br/ |
| Guilherme Boulos | |
| <i>Facebook</i> | https://www.Facebook.com/guilhermeboulos/ |
| <i>Twitter</i> | https://Twitter.com/guilhermeboulos |
| <i>Youtube</i> | https://www.Youtube.com/channel/UCFwzvNKpQWKMBiOHJ26IEnQ |
| <i>Instagram</i> | https://www.Instagram.com/guilhermeboulos.oficial |
| <i>Site</i> | https://vamoscomboulosesonia.com.br/ |
| Henrique Meirelles | |
| <i>Facebook</i> | https://www.Facebook.com/hmeirellesoficial/ |
| <i>Twitter</i> | https://Twitter.com/meirelles |
| <i>Youtube</i> | https://www.Youtube.com/henriquemeirellesoficial |
| <i>Instagram</i> | https://www.Instagram.com/henriquemeirelles.real/ |
| <i>Site</i> | http://www.henriquemeirelles.com.br/ |

Jair Bolsonaro

| | |
|------------------|---|
| Facebook | https://www.Facebook.com/jairmessias.bolsonaro |
| Twitter | https://Twitter.com/jairbolsonaro |
| Youtube | https://www.Youtube.com/user/jbolsonaro |
| Instagram | https://www.Instagram.com/jairmessiasbolsonaro/ |
| Site | https://www.bolsonaro.com.br/ |

João Amoêdo

| | |
|------------------|---|
| Facebook | https://www.Facebook.com/JoaoAmoedoNOVO/ |
| Twitter | https://Twitter.com/joaoamoedonovo |
| Youtube | https://www.Youtube.com/channel/UCIYntHsUuqq2GmwvU9T_gCg |
| Instagram | https://www.Instagram.com/joaoamoedonovo/ |
| Site | http://joaoamoedo.com.br/ |

João Goulart Filho

| | |
|------------------|---|
| Facebook | https://www.Facebook.com/joagoulart54/ |
| Twitter | https://Twitter.com/joagoulart54 |
| Youtube | https://www.Youtube.com/channel/UCGoaf93DIN11MjLF-Ymlggg |
| Instagram | https://www.Instagram.com/joagoulartfilho_oficial/ |
| Site | https://joagoulartfilho54.com.br/ |

José Maria Eymael

| | |
|------------------|---|
| Facebook | https://www.Facebook.com/eymaelOficial/ |
| Twitter | https://Twitter.com/eymaeloficial |
| Youtube | não encontrado |
| Instagram | https://www.Instagram.com/eymael_presidente27/ |
| Site | https://www.eymaelpresidente27.com.br/ |

Marina Silva

| | |
|------------------|---|
| Facebook | https://www.Facebook.com/marinasilva.oficial/ |
| Twitter | https://Twitter.com/marinasilva |
| Youtube | https://www.Youtube.com/channel/UC9uefWa6TXIPDRWGZYMCTuA |
| Instagram | https://www.Instagram.com/marinasilva/ |
| Site | https://marinasilva.org.br/ |

Vera Lucia

| | |
|------------------|---|
| Facebook | https://www.Facebook.com/verapstu/ |
| Twitter | https://Twitter.com/verapstu |
| Youtube | não encontrado |
| Instagram | não encontrado |
| Site | https://www.pstu.org.br/ |

Lula

| | |
|------------------|---|
| Facebook | https://www.Facebook.com/Lula/ |
| Twitter | https://Twitter.com/LulaOficial |
| Youtube | https://www.Youtube.com/lulaoficial |
| Instagram | https://www.Instagram.com/lulaoficial/ |
| Site | https://www.lula.com.br/ |

APÊNDICE H – LISTA DOS LINKS DAS MÍDIAS SOCIAIS E SITES DOS CANDIDATOS À PRESIDÊNCIA DE 2018

| FONTES | CANDIDATOS | PRESERVAÇÃO |
|------------------|---|--|
| | Álvaro Dias | |
| Facebook | https://Web.archive.org/Web/20180601000000*/https://www.Facebook.com/ad.alvarodias/ https://wayback.Archive-It.org/all*/https://www.Facebook.com/ad.alvarodias/ | <i>Internet Archive</i> <i>Archive-It</i> |
| Twitter | https://Web.archive.org/Web/20180501000000*/https://Twitter.com/alvarodias | <i>Internet Archive</i> |
| Youtube | nenhum registro de 2018 preservado | |
| Instagram | nenhum registro de 2018 preservado | |
| Site | https://Web.archive.org/Web/20180701000000*/https://alvarodias.com.br | <i>Internet Archive</i> |
| | Cabo Daciolo | |
| Facebook | https://www.Facebook.com/depudadocabodaciolo/ https://Web.archive.org/Web/20180901000000*/https://www.Facebook.com/depudadocabodaciolo/ https://wayback.Archive-It.org/all*/https://www.Facebook.com/depudadocabodaciolo/ | <i>Internet Archive</i> <i>Archive-It</i> |
| Twitter | https://Twitter.com/cabodaciolo nenhum registro de 2018 preservado | |
| Youtube | https://www.Youtube.com/channel/UCdtc9CCVxLBVkn51Lvy_wTA nenhum registro de 2018 preservado | |
| Instagram | https://www.Instagram.com/cabodaciolo/ nenhum registro de 2018 preservado | |
| Site | http://patriota51.org.br/ https://Web.archive.org/Web/20180815000000*/http://patriota51.org.br/ | <i>Internet Archive</i> |
| | Ciro Gomes | |
| Facebook | https://www.Facebook.com/cirogomesoficial/ https://Web.archive.org/Web/*/https://www.Facebook.com/cirogomesoficial/ https://wayback.Archive-It.org/all*/https://www.Facebook.com/cirogomesoficial/ | <i>Internet Archive</i> <i>Archive-It</i> |
| Twitter | https://Twitter.com/cirogomes https://Web.archive.org/Web/20180401000000*/https://Twitter.com/cirogomes | <i>Internet Archive</i> |
| Youtube | https://www.Youtube.com/cirogomesoficial | |

| | | |
|------------------------|---|--|
| | nenhum registro de 2018 preservado | |
| Instagram | https://www.Instagram.com/cirogomes/ https://Web.archive.org/Web/20180401000000*/https://www.Instagram.com/cirogomes/ | <i>Internet Archive</i> |
| Site | https://www.cirogomes.com.br/ https://Web.archive.org/Web/20180401000000*/https://www.cirogomes.com.br/ | <i>Internet Archive</i> |
| Fernando Haddad | | |
| Facebook | https://www.Facebook.com/fernandohaddad/ https://Web.archive.org/Web/*/https://www.Facebook.com/fernandohaddad/ https://Archive.is/https://www.Facebook.com/fernandohaddad/ | <i>Internet Archive</i> <i>Archive.is</i> |
| Twitter | https://Twitter.com/haddad_fernando https://Web.archive.org/Web/20180901000000*/https://Twitter.com/haddad_fernando https://Archive.is/https://Twitter.com/haddad_fernando | <i>Internet Archive</i> <i>Archive.is</i> |
| Youtube | https://www.Youtube.com/channel/UCyL3IelkOB9yl7XvgTYkc0A https://Web.archive.org/Web/20180901000000*/https://www.Youtube.com/channel/UCyL3IelkOB9yl7XvgTYkc0A | <i>Internet Archive</i> |
| Instagram | https://www.Instagram.com/fernandohaddadoficial/ https://Web.archive.org/Web/*/https://www.Instagram.com/fernandohaddadoficial/ | <i>Internet Archive</i> |
| Site | https://haddadpresidente.com.br/ https://Web.archive.org/Web/20180401000000*/https://haddadpresidente.com.br/ | <i>Internet Archive</i> |
| Geraldo Alckmin | | |
| Facebook | https://www.Facebook.com/geraldoalckmin/ https://Web.archive.org/Web/*/https://www.Facebook.com/geraldoalckmin/ https://wayback.Archive-It.org/all*/https://www.Facebook.com/geraldoalckmin/ | <i>Internet Archive</i> <i>Archive-It</i> |
| Twitter | https://Twitter.com/geraldoalckmin https://Web.archive.org/Web/20180701000000*/https://Twitter.com/geraldoalckmin https://Archive.is/https://Twitter.com/geraldoalckmin | <i>Internet Archive</i> <i>Archive.is</i> |
| Youtube | https://www.Youtube.com/channel/UCNxCni0Iv9pr7i_pQZ6ijlg https://Web.archive.org/Web/20180301000000*/https://www.Youtube.com/channel/UCNxCni0Iv9pr7i_pQZ6ijlg | <i>Internet Archive</i> |
| Instagram | https://www.Instagram.com/geraldoalckmin/ nenhum registro de 2018 preservado | |

| | | |
|---------------------------|---|--|
| Site | https://www.geraldoalckmin.com.br/ https://Web.archive.org/Web/20180915000000*/https://www.geraldoalckmin.com.br/ | <i>Internet Archive</i> |
| Guilherme Boulos | | |
| Facebook | https://www.Facebook.com/guilhermeboulos/ https://Web.archive.org/Web/*/https://www.Facebook.com/guilhermeboulos/ https://wayback.Archive-It.org/all*/https://www.Facebook.com/guilhermeboulos/ | <i>Internet Archive</i> <i>Archive-It</i> |
| Twitter | https://Twitter.com/guilhermeboulos https://Web.archive.org/Web/20180815000000*/https://Twitter.com/guilhermeboulos https://Archive.is/https://Twitter.com/guilhermeboulos | <i>Internet Archive</i> <i>Archive.is</i> |
| Youtube | https://www.Youtube.com/channel/UCFwzvNKpQWKMBiOHJ26IEEnQ nenhum registro de 2018 preservado | |
| Instagram | https://www.Instagram.com/guilhermeboulos.official nenhum registro de 2018 preservado | |
| Site | https://vamoscomboulozesonia.com.br/ https://Web.archive.org/Web/20180601000000*/https://vamoscomboulozesonia.com.br/ | <i>Internet Archive</i> |
| Henrique Meirelles | | |
| Facebook | https://www.Facebook.com/hmeirellesoficial/ https://Web.archive.org/Web/*/https://www.Facebook.com/hmeirellesoficial/ https://wayback.Archive-It.org/all*/https://www.Facebook.com/hmeirellesoficial/ | <i>Internet Archive</i> <i>Archive-It</i> |
| Twitter | https://Twitter.com/meirelles https://Web.archive.org/Web/20180515000000*/https://Twitter.com/meirelles | <i>Internet Archive</i> |
| Youtube | https://www.Youtube.com/henriquemeirellesoficial nenhum registro de 2018 preservado | |
| Instagram | https://www.Instagram.com/henriquemeirelles.real/ nenhum registro de 2018 preservado | |
| Site | http://www.henriquemeirelles.com.br/ https://Web.archive.org/Web/20180515000000*/http://www.henriquemeirelles.com.br/ | <i>Internet Archive</i> |
| Jair Bolsonaro | | |
| Facebook | https://www.Facebook.com/jairmessias.bolsonaro | |

| | | |
|---------------------------|---|-------------------------|
| | https://Web.archive.org/Web/*/https://www.Facebook.com/jairmessias.bolsonaro | <i>Internet Archive</i> |
| | https://wayback.Archive-It.org/all/*/https://www.Facebook.com/jairmessias.bolsonaro/ | <i>Archive-It</i> |
| | https://Archive.is/https://www.Facebook.com/jairmessias.bolsonaro | <i>Archive.is</i> |
| Twitter | https://Twitter.com/jairbolsonaro | |
| | https://Web.archive.org/Web/20180415000000*/https://Twitter.com/jairbolsonaro | <i>Internet Archive</i> |
| | https://Archive.is/https://Twitter.com/jairbolsonaro | <i>Archive.is</i> |
| Youtube | https://www.Youtube.com/user/jbolsonaro | |
| | https://Web.archive.org/Web/20180301000000*/https://www.Youtube.com/user/jbolsonaro | <i>Internet Archive</i> |
| Instagram | https://www.Instagram.com/jairmessiasbolsonaro/ | |
| | https://Web.archive.org/Web/20180415000000*/https://www.Instagram.com/jairmessiasbolsonaro/ | <i>Internet Archive</i> |
| Site | https://www.bolsonaro.com.br/ | |
| | https://Web.archive.org/Web/20180501000000*/https://www.bolsonaro.com.br/ | <i>Internet Archive</i> |
| | https://Archive.is/https://www.bolsonaro.com.br/ | <i>Archive.is</i> |
| João Amoêdo | | |
| Facebook | https://www.Facebook.com/JoaoAmoedoNOVO/ | |
| | https://Web.archive.org/Web/*/https://www.Facebook.com/JoaoAmoedoNOVO/ | <i>Internet Archive</i> |
| | https://wayback.Archive-It.org/all/*/https://www.Facebook.com/JoaoAmoedoNOVO/ | <i>Archive-It</i> |
| Twitter | https://Twitter.com/joaoamoedonovo | |
| | https://Web.archive.org/Web/20180701000000*/https://Twitter.com/joaoamoedonovo | <i>Internet Archive</i> |
| Youtube | https://www.Youtube.com/channel/UCIYntHsUuqq2GmwvU9T_gCg | |
| | nenhum registro de 2018 preservado | |
| Instagram | https://www.Instagram.com/joaoamoedonovo/ | |
| | nenhum registro de 2018 preservado | |
| Site | http://joaoamoedo.com.br/ | |
| | https://Web.archive.org/Web/20180501000000*/http://joaoamoedo.com.br/ | <i>Internet Archive</i> |
| João Goulart Filho | | |
| Facebook | https://www.Facebook.com/joagoulart54/ | |
| | https://Web.archive.org/Web/*/https://www.Facebook.com/joagoulart54/ | <i>Internet Archive</i> |
| | https://wayback.Archive-It.org/all/*/https://www.Facebook.com/joagoulart54/ | <i>Archive-It</i> |

| | | |
|--------------------------|---|--|
| Twitter | https://Twitter.com/joagoulart54 nenhum registro de 2018 preservado | |
| Youtube | https://www.Youtube.com/channel/UCGoaf93DIN11MjLF-Ymlggg nenhum registro de 2018 preservado | |
| Instagram | https://www.Instagram.com/joagoulartfilho_oficial/ nenhum registro de 2018 preservado | |
| Site | https://joagoulartfilho54.com.br/ https://Web.archive.org/Web/20180801000000*/https://joagoulartfilho54.com.br/ | <i>Internet Archive</i> |
| José Maria Eymael | | |
| Facebook | https://www.Facebook.com/eymaelOficial/ https://Web.archive.org/Web/*/https://www.Facebook.com/eymaelOficial/ https://wayback.Archive-It.org/all/*/https://www.Facebook.com/eymaelOficial/ | <i>Internet Archive</i> <i>Archive-It</i> |
| Twitter | https://Twitter.com/eymaeloficial https://Web.archive.org/Web/20180501000000*/https://Twitter.com/eymaeloficial | <i>Internet Archive</i> |
| Youtube | não encontrado | |
| Instagram | https://www.Instagram.com/eymael_presidente27/ nenhum registro de 2018 preservado | |
| Site | https://www.eymaelpresidente27.com.br/ https://Web.archive.org/Web/20180501000000*/https://www.eymaelpresidente27.com.br/ | <i>Internet Archive</i> |
| Marina Silva | | |
| Facebook | https://www.Facebook.com/marinasilva.oficial/ https://Web.archive.org/Web/*/https://www.Facebook.com/marinasilva.oficial/ https://wayback.Archive-It.org/all/*/https://www.Facebook.com/marinasilva.oficial/ | <i>Internet Archive</i> <i>Archive-It</i> |
| Twitter | https://Twitter.com/marinasilva https://Web.archive.org/Web/20180501000000*/https://Twitter.com/marinasilva | <i>Internet Archive</i> |
| Youtube | https://www.Youtube.com/channel/UC9uefWa6TXIPDRWGZYMCTuA https://Web.archive.org/Web/20180801000000*/https://www.Youtube.com/channel/UC9uefWa6TXIPDRWGZYMCTuA | <i>Internet Archive</i> |
| Instagram | https://www.Instagram.com/marinasilva/ https://Web.archive.org/Web/20180801000000*/https://www.Instagram.com/marinasilva/ | <i>Internet Archive</i> |

| | | |
|-------------------|---|--|
| Site | https://marinasilva.org.br/ https://Web.archive.org/Web/20180901000000*/https://marinasilva.org.br/ | <i>Internet Archive</i> |
| Vera Lucia | | |
| Facebook | https://www.Facebook.com/verapstu/ nenhum registro de 2018 preservado | |
| Twitter | https://Twitter.com/verapstu nenhum registro de 2018 preservado | |
| Youtube | não encontrado | |
| Instagram | não encontrado | |
| Site | https://www.pstu.org.br/ https://Web.archive.org/Web/20180701000000*/https://www.pstu.org.br/ | <i>Internet Archive</i> |
| Lula | | |
| Facebook | https://www.Facebook.com/Lula/ https://Web.archive.org/Web/20180501000000*/https://www.Facebook.com/Lula/ https://wayback.Archive-It.org/all*/https://www.Facebook.com/Lula/ | <i>Internet Archive</i> <i>Archive-It</i> |
| Twitter | https://Twitter.com/LulaOficial https://Web.archive.org/Web/20180701000000*/https://Twitter.com/LulaOficial https://Archive.is/https://Twitter.com/LulaOficial | <i>Internet Archive</i> <i>Archive.is</i> |
| Youtube | https://www.Youtube.com/lulaoficial https://Web.archive.org/Web/20180601000000*/https://www.Youtube.com/lulaoficial | <i>Internet Archive</i> |
| Instagram | https://www.Instagram.com/lulaoficial/ https://Web.archive.org/Web/20180601000000*/https://www.Instagram.com/lulaoficial/ | <i>Internet Archive</i> |
| Site | https://www.lula.com.br/ https://Web.archive.org/Web/20180701000000*/https://www.lula.com.br/ | <i>Internet Archive</i> |

ANEXO A – RESULTADO DO PRIMEIRO TURNO DAS ELEIÇÕES PRESIDENCIAIS DE 2018

Eleições 2018

Eleição Ordinária Federal - 2018 - 1º Turno

BRASIL

PRESIDENTE

Todos os Candidatos Candidatos com votação Nula ou Anulada

Atualizar

RESULTADO SUJEITO A ALTERAÇÃO

Atualizado em: 22/01/2019
14:12:19

| Seções | 454.490 | Sequencial | Número | Nome | Partido/Coligação | Votos | % Válidos |
|------------------------|----------------------|------------|--------|--------------------|---|------------|-----------|
| Totalizadas | 454.490 (100,00%) | * 1 | 17 | JAIR BOLSONARO | PSL - PSL / PRTB | 49.277.010 | 46,03 % |
| Não Totalizadas | 0 (0,00%) | * 2 | 13 | FERNANDO HADDAD | PT - PT / PC do B / PROS | 31.342.051 | 29,28 % |
| Eleitorado | 147.306.295 | 3 | 12 | CIRO GOMES | PDT - PDT / AVANTE | 13.344.371 | 12,47 % |
| Não Apurado | 470 (0,01%) | 4 | 45 | GERALDO ALCKMIN | PSDB - PSDB / PTB / PP / PR / DEM / SOLIDARIEDADE / PPS / PRB / PSD | 5.096.350 | 4,76 % |
| Apurado | 147.305.825 (99,99%) | 5 | 30 | JOÃO AMOÉDO | NOVO | 2.679.745 | 2,50 % |
| Abstenção | 29.941.171 (20,33%) | 6 | 51 | CABO DACIOLO | PATRI | 1.348.323 | 1,26 % |
| Comparecimento | 117.364.654 (79,67%) | 7 | 15 | HENRIQUE MEIRELLES | MDB - MDB / PHS | 1.288.950 | 1,20 % |
| | | 8 | 18 | MARINA SILVA | REDE - REDE / PV | 1.069.578 | 1,00 % |
| Votos | 117.364.654 | 9 | 19 | ALVARO DIAS | PODE - PODE / PRP / PSC / PTC | 859.601 | 0,80 % |
| Branços | | 10 | 50 | GUILHERME BOULOS | PSOL - PSOL / PCB | 617.122 | 0,58 % |

ANEXO B – INFOGRÁFICO DE TEMPO DE PROPAGANDA NO RÁDIO E NA TV DOS CANDIDATOS À PRESIDÊNCIA DE 2018

Propaganda eleitoral no 1º turno

Veja tempo previsto para cada um dos candidatos à Presidência



Fonte: TSE