

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E INFORMAÇÃO

MARÍLIA GEHRKE

O USO DE FONTES DOCUMENTAIS NO JORNALISMO GUIADO POR DADOS

PORTO ALEGRE

2018

MARÍLIA GEHRKE

O USO DE FONTES DOCUMENTAIS NO JORNALISMO GUIADO POR DADOS

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

Orientadora: Prof.^a Dra. Luciana Mielniczuk
Coorientador: Prof. Dr. Marcelo Träsel

PORTO ALEGRE

2018

CIP - Catalogação na Publicação

Gehrke, Marília

O uso de fontes documentais no jornalismo guiado por dados / Marília Gehrke. -- 2018.

129 f.

Orientadora: Luciana Mielniczuk.

Coorientador: Marcelo Träsel.

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, 2018.

1. fontes jornalísticas. 2. jornalismo guiado por dados. 3. jornalismo em rede. 4. fontes documentais. 5. jornalismo digital. I. Mielniczuk, Luciana, orient. II. Träsel, Marcelo, coorient. III. Título.

MARÍLIA GEHRKE

O USO DE FONTES DOCUMENTAIS NO JORNALISMO GUIADO POR DADOS

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

Aprovado em: 09/01/2018
BANCA EXAMINADORA

Prof.^a Dra. Luciana Mielniczuk – PPGCOM/UFRGS
Orientadora

Prof.^a Dra. Suzana Barbosa – PÓSCOM/UFBA
Examinadora

Prof.^a Dra. Thais Helena Furtado – FABICO/UFRGS
Examinadora

Prof.^a Dra. Marcia Benetti – PPGCOM/UFRGS
Examinadora

Prof.^a Dra. Ilza Maria Tourinho Girardi – PPGCOM/UFRGS
Suplente

RESUMO

Estudar as fontes utilizadas nas notícias de jornalismo guiado por dados (JGD) é a proposta desta dissertação. Para tanto, revisita as classificações de fontes trabalhadas por teóricos da área e situa o contexto atual, derivado de transformações sociais e tecnológicas, sob a perspectiva de sociedade em rede e do jornalismo em rede. O foco do estudo está em descobrir quais fontes são acionadas em notícias do JGD, que emerge neste cenário a partir dos anos 2000. Analisa um *corpus* constituído por 60 notícias veiculadas nos jornais *O Globo*, *The New York Times* e *La Nación*, como veículos tradicionais, e *Nexo*, *FiveThirtyEight* e *Chequeado*, como veículos nativos. A partir do cruzamento entre a teoria e o estudo empírico, propõe a classificação de tipos de fontes nas notícias de JGD. São eles: arquivo documental, estatística e reprodução. Por meio dessa classificação, busca preencher uma lacuna no quadro teórico sobre fontes, superficialmente discutido no jornalismo até então, trazendo o uso de documentos como protagonista neste cenário.

Palavras-chave: fontes jornalísticas; jornalismo guiado por dados; jornalismo em rede; fontes documentais; jornalismo digital.

ABSTRACT

Studying the news sources used in data-driven journalism (DDJ) practices is the proposal of this dissertation. The theoretical approach includes classifications of news sources already discussed in journalism studies. Considering the contemporary context, which is modified by social and technological transformations, this study operates from the networked society and network journalism perspectives. The main point is to detect the use of journalism sources in news developed by DDJ techniques, which emerges in this scenario during the 2000's. It analyzes 60 news records published by *O Globo*, *The New York Times* and *La Nación*, as traditional media, and *Nexo*, *FiveThirtyEight* and *Chequeado*, as the native ones. Combining the theory and the empirical study, it proposes a classification by types of sources of DDJ news: documentary file, statistics and reproduction. Through this classification, it aims to fulfill a gap found in the theoretical sources approach, which is superficially discussed in journalism until now, bringing the use of documents as a protagonist in this scenario.

Keywords: news sources; data-driven journalism; network journalism; documental sources; digital journalism.

AGRADECIMENTOS

Ao Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul (PPGCOM/UFRGS), em nome dos professores e funcionários, pela excelência na formação e a riqueza das discussões em sala de aula. À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), cuja bolsa permitiu dedicação exclusiva aos estudos.

À minha orientadora, Luciana Mielniczuk, pela compreensão, paciência e parceria, pelo incentivo e pelos incontáveis *insights* ao longo das discussões no período do mestrado. Obrigada, Luti, pelas inúmeras conversas sobre a vida dentro e fora da academia, que tanto contribuíram para o meu crescimento profissional e pessoal. Ter uma orientadora presente e disposta a dialogar em todo o processo foi fundamental neste início de trajetória como pesquisadora.

Às professoras da banca de qualificação, Virgínia Fonseca e Thais Furtado, meu agradecimento pelas preciosas contribuições, fundamentais ao desenvolvimento do trabalho. Às professoras Marcia Benetti e Suzana Barbosa por aceitarem o convite para compor a banca final e dialogar sobre o jornalismo. Que orgulho ter uma banca composta por mulheres!

Às colegas do Grupo de Pesquisa em Jornalismo Digital (JorDi) e do PPGCOM agradeço pela amizade e pelo companheirismo no dia a dia e nos congressos. A caminhada ficou mais leve com todas vocês: Fabiana Freitas, Anelise Schütz Dias, Vanessa Kannenberg, Francielly Brites, Máira Evangelista de Sousa, Tali Bezerra, Marlise Brenol, Alciane Baccin, Janaína Kalsing e Bárbara Nickel. Ao meu coorientador, Marcelo Träsel, agradeço pela troca de ideias e sugestões; e ao Francisco Amorim por estar sempre disposto a ajudar, discutir e questionar. Vocês têm a minha admiração.

Por fim, e não menos importante, agradeço aos meus pais, Marlise e Armênio Gehrke, pelo apoio e afeto incondicionais. Impossível chegar até aqui sem o suporte e a estrutura da minha família. Obrigada por acreditarem na importância do meu estudo e da minha formação.

*Se a teoria na prática é outra, então há algo de errado na teoria
(Adelmo Genro Filho)*

LISTA DE FIGURAS

| | |
|---|-----|
| Figura 1 – Mapa da cólera | 41 |
| Figura 2 – Modelo de ficha de leitura..... | 65 |
| Figura 3 – O mapa de reprodução da música <i>Despacito</i> | 68 |
| Figura 4 – Visualização com mapas sobre a preservação da Mata Atlântica..... | 70 |
| Figura 5 – As refeições pelo Brasil e pelo mundo segundo o <i>Twitter</i> | 71 |
| Figura 6 – Profissões de acordo com a similaridade de atividades | 73 |
| Figura 7 – Inserção de profissão em infográfico interativo | 73 |
| Figura 8 – Mapa antigo do Brooklyn..... | 75 |
| Figura 9 – Uso de notas pelo <i>FiveThirtyEight</i> | 77 |
| Figura 10 – Uso de estatística para entender a escolha por chocolate..... | 78 |
| Figura 11 – Documento ligado ao <i>Panama Papers</i> | 80 |
| Figura 12 – Dados qualitativos transformam-se em números | 81 |
| Figura 13 – Nota e publicação no <i>Twitter</i> | 83 |
| Figura 14 – Planilha na organização dos dados..... | 85 |
| Figura 15 – Gráfico da distribuição de fontes documentais por veículo | 96 |
| Figura 16 – Gráfico da distribuição dos tipos por veículo | 107 |

LISTA DE QUADROS E TABELAS

| | |
|--|-----|
| Quadro 1 – Abordagens teóricas de classificação de fontes..... | 28 |
| Quadro 2 – Produção jornalística orientada por <i>software</i> | 52 |
| Quadro 3 – Listagem das notícias estudadas | 61 |
| Quadro 4 – Cabeçalho da ficha de leitura..... | 66 |
| Quadro 5 – Ficha de leitura em formato de relatório | 66 |
| Quadro 6 – Das palavras-chave aos agrupamentos | 89 |
| Quadro 7 – Fontes documentais acionadas em <i>O Globo</i> | 97 |
| Quadro 8 – Fontes documentais acionadas pelo <i>Nexo</i> | 98 |
| Quadro 9 – Fontes documentais acionadas pelo <i>The New York Times</i> | 100 |
| Quadro 10 – Fontes documentais acionadas pelo <i>FiveThirtyEight</i> | 102 |
| Quadro 11 – Fontes documentais acionadas pelo <i>La Nación</i> | 104 |
| Quadro 12 – Fontes documentais acionadas pelo <i>Chequeado</i> | 105 |
| Quadro 13 – Principais características encontradas no <i>corpus</i> estudado | 108 |
| Quadro 14 – As fontes documentais a partir da teoria | 111 |
| | |
| Tabela 1 – Ocorrências de fontes por veículo | 95 |

SUMÁRIO

| | |
|--|----|
| 1 INTRODUÇÃO | 12 |
| 2 FONTES, A ORIGEM DA INFORMAÇÃO NO JORNALISMO | 16 |
| 3 A SOCIEDADE EM REDE, O FENÔMENO BIG DATA E O JORNALISMO | 30 |
| 3.1 Bases para um novo paradigma | 31 |
| 3.2 O jornalismo em rede | 35 |
| 3.3 Jornalismo guiado por dados (JGD): contexto e definições | 38 |
| 3.3.1 Primeiras construções de um cenário em transformação | 39 |
| 3.4 Jornalismo de precisão e Reportagem Assistida por Computador (RAC)..... | 42 |
| 3.5 Jornalismo guiado por dados e jornalismo computacional | 45 |
| 4 ESTUDO DE FONTES | 54 |
| 4.1 Os veículos: tradicionais e nativos no Brasil, Estados Unidos e Argentina | 54 |
| 4.1.1 <i>O Globo e o blog Na Base dos Dados: cobertura diversa</i> | 55 |
| 4.1.2 <i>Nexo Jornal: o contexto a partir dos dados</i> | 56 |
| 4.1.3 <i>The New York Times: explicando as notícias na seção The Upshot</i> | 57 |
| 4.1.4 <i>FiveThirtyEight e a análise estatística</i> | 58 |
| 4.1.5 <i>La Nación e La Nación Data para fiscalizar o governo</i> | 59 |
| 4.1.6 <i>Chequeado e a verificação de discursos públicos</i> | 60 |
| 4.2 A definição das 60 notícias..... | 60 |
| 4.3 Etapas de organização dos dados..... | 64 |
| 4.3.1 Leitura e descrição em fichas | 64 |
| 4.3.1.1 O Globo | 67 |
| 4.3.1.2 Nexo | 69 |
| 4.3.1.3 The New York Times | 72 |
| 4.3.1.4 FiveThirtyEight | 76 |
| 4.3.1.5 La Nación | 79 |
| 4.3.1.6 Chequeado | 82 |
| 4.3.1.7 O que mostram os exemplos | 84 |
| 4.3.2 <i>O uso de planilhas na organização dos dados</i> | 85 |
| 4.4 Proposta de classificação de tipos de fontes para o jornalismo guiado por dados..... | 88 |
| 4.4.1 <i>Classificação inicial incluindo fontes pessoais, documentais e híbridas</i> | 90 |
| 4.4.2 <i>Categorias preliminares</i> | 90 |
| 4.4.3 <i>Categorias refinadas: filtro para fontes documentais</i> | 92 |
| 4.4.3.1 Arquivo documental | 93 |
| 4.4.3.2 Estatística | 93 |
| 4.4.3.3 Reprodução | 94 |

| | |
|---|-----|
| 4.5 Fontes em números: os resultados globais da pesquisa..... | 94 |
| 4.5.1 <i>O Globo</i> | 96 |
| 4.5.2 <i>Nexo</i> | 98 |
| 4.5.3 <i>The New York Times</i> | 99 |
| 4.5.4 <i>FiveThirtyEight</i> | 102 |
| 4.5.5 <i>La Nación</i> | 104 |
| 4.5.6 <i>Chequeado</i> | 105 |
| 4.6 Considerações finais do estudo empírico..... | 108 |
| 5 CONSIDERAÇÕES FINAIS | 110 |
| REFERÊNCIAS | 114 |
| APÊNDICE A – FICHAS DE LEITURA | 126 |
| APÊNDICE B – CATEGORIAS PRELIMINARES | 127 |
| APÊNDICE C – PLANILHA DE FONTES | 129 |

1 INTRODUÇÃO

O jornalismo transforma e é transformado sob o ritmo da reconfiguração da vida em sociedade e de seus modos de organização. Suas bases estão associadas à ideia de verdade, verificação da informação, manutenção da democracia e do interesse público. Os avanços tecnológicos entram como parte do processo à medida que contribuem para o desenvolvimento ou a ascensão de novos conjuntos de práticas. Independente da ferramenta, o jornalismo continua utilizando fontes que dão sustentação às notícias. Dentre o que emerge em um cenário de sociedade em rede e jornalismo em rede está a facilidade de acesso às múltiplas fontes documentais, como bases de dados em portais da transparência, estudos científicos, pesquisas, notas, registros e outros, que são parte do repositório que o jornalismo guiado por dados (JGD) utiliza como fonte.

As fontes são levadas em conta nos estudos de jornalismo desde a primeira tese defendida na área, em 1690, por Tobias Peucer. Da década de 1970 até perto dos anos 2000, os teóricos preocupavam-se especialmente com a relação de poder nas trocas entre jornalistas e fontes (GANS, [1979] 2004; HALL et. al. [1978] 2016). Parte dos pesquisadores que problematizaram essas questões já neste século XXI começaram a perceber mudanças na busca pelas fontes de informação, como o próprio uso de documentos e a disponibilidade de fontes na internet (CANAVILHAS; IVARS-NICOLÁS, 2012; CHAPARRO, 2009; PINTO, 2000). As classificações e tipologias trabalhadas pelos autores são formas de perceber como ocorrem as trocas entre jornalistas e fontes e quais costumam ser acionadas.

As transformações estão inseridas em um cenário de jornalismo em rede (HEINRICH, 2008), cuja estrutura foge à linearidade de processos desenvolvida até meados do século XX, e avança em direção a um espaço de fluxos horizontal, com o apagamento de barreiras geográficas e temporais (CASTELLS, 2016). A gestação da sociedade em rede, que apresenta essas características, começou nos anos de 1970, alterando o modo como as pessoas se relacionam na vida pessoal e profissional. Não significa, porém, que a estrutura anterior, dominada pelos processos industriais, foi completamente extinta. Os modos de organização podem coexistir e se sobressair em maior ou menor escala. O atual cenário é composto pelo fenômeno *big data*, derivado de uma quantidade de informações sem precedentes que se estruturam em uma corrente social, cultural e tecnológica (BOYD; CRAWFORD, 2012; LEWIS; WESTLUND, 2014).

A partir deste panorama, acredita-se na necessidade de compreender de onde vêm as informações empregadas no jornalismo, ou seja, quais fontes ganham espaço nas notícias. O

jornalista é responsável por trazer explicações e articular dados e ideias em uma narrativa, mas as fontes acionadas também determinam a qualidade da informação produzida. No JGD, os dados são as principais fontes da notícia. E essa é pontuada como sua principal diferença em relação às produções tradicionais, que têm como base o contato para a realização de entrevistas, principalmente. Entende-se, então, que é pertinente e importante estudar o emprego de fontes no jornalismo guiado por dados, pois deste contexto podem emergir situações ainda pouco exploradas nas pesquisas da área.

Durante levantamento do estado da arte de pesquisas realizadas, construído a partir da consulta em dois repositórios públicos de teses e dissertações, o da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes)¹ e o do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)², percebeu-se a existência de uma lacuna: não havia estudos específicos sobre o uso de fontes no jornalismo guiado por dados. Havia, sim, dissertações e teses sob outras perspectivas do JGD³.

A discussão sobre fontes documentais ou fontes neste contexto, porém, não foi localizada. Esta dissertação tem a pretensão, portanto, de preencher um espaço em aberto, discutindo as fontes acionadas nesta prática profissional que se utiliza basicamente de informações em bancos de dados públicos. A preocupação com o uso de fontes e de informações quantitativas no jornalismo, a exemplo das práticas de JGD, tornou-se um ponto de inquietação desde o período de atuação como repórter, sendo levada ao início do mestrado, em 2016. A experiência como repórter permitiu vivenciar os conflitos e interesses entre jornalista e fonte, além dos desafios de utilizar repositórios públicos para obter informações noticiáveis.

Pressupondo que órgãos públicos sejam transparentes no fornecimento de informações estruturadas e que o próprio jornalista é capaz de criar seus bancos de dados, as técnicas do JGD permitem fazer consultas a bancos de dados, planilhas e documentos sem precisar passar

¹ A pesquisa no Banco de Teses e Dissertações da Capes foi realizada em 20 de junho de 2017. Acesso pelo endereço <http://bancodeteses.capes.gov.br/banco-teses>.

² A pesquisa na Biblioteca Digital Brasileira de Teses e Dissertações, no site do Ibict, ocorreu em 14 de junho de 2017. Acesso pelo endereço <http://bdtd.ibict.br>.

³ O termo jornalismo guiado por dados, buscado entre aspas, gerou três resultados no portal da Capes: duas dissertações de mestrado – uma de 2014, de autoria de Felipe Rodrigues Grandin (UERJ), sobre a contribuição do JGD na criação de valor para as empresas jornalísticas; e outra de 2015, de Matias Sebastião Peruyera (UTFPR), sobre os usos e apropriações da tecnologia no JGD. Há, ainda, a tese de Marcelo Träsel, defendida em 2014 na Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), que identifica os valores, as crenças e a conduta dos profissionais de JGD tendo o *Estadão Dados*, do jornal *O Estado de S. Paulo*, como espaço de observação participante. Também no portal da Capes ocorreu a consulta pelo termo jornalismo de dados, gerando sete resultados, dos quais quatro apresentam foco na visualização de dados. Entre os demais, a ênfase está no uso do jornalismo de precisão e da Lei de Acesso à Informação como fonte para jornalistas. No Banco de Teses e Dissertações do Ibict, a combinação dos termos documentos e fontes gerou cinco resultados, mas nenhum está relacionado ao jornalismo. Na união dos termos jornalismo de dados ou jornalismo guiado por dados e fonte, com busca avançada em título e assunto, nenhum resultado foi encontrado.

por assessorias de imprensa ou contatar diretamente a pessoa que detém as informações. O emprego de fontes na elaboração do conteúdo jornalístico continuará sendo fundamental, independente de novas transformações que virão. Observar quais tipos de fontes são citadas e com que frequência aparecem nos textos jornalísticos permite se ter uma ideia da origem das informações veiculadas atualmente pelo jornalismo.

A pergunta que esta dissertação se propõe a responder é quais tipos de fontes são acionadas em notícias construídas a partir de técnicas do jornalismo guiado por dados. A partir desta resposta, o objetivo geral é classificar as fontes de natureza documental encontradas no estudo empírico. Os objetivos específicos são a) analisar produções jornalísticas que se utilizam de técnicas do jornalismo guiado por dados, b) observar a ocorrência do uso de documentos como fontes no JGD e c) identificar os tipos de fontes documentais acionados e sua origem.

O *corpus* deste estudo é composto por 60 registros noticiosos publicados entre 1º de janeiro de 2016 e 31 de outubro de 2017 por jornais brasileiros, estadunidenses e argentinos. Os veículos de comunicação eleitos para a composição do recorte apresentam equipes específicas para o trabalho com jornalismo guiado por dados ou então têm em sua redação equipes responsáveis por analisá-los. São eles: *O Globo*, *The New York Times* e *La Nación*, considerados veículos tradicionais, que derivaram de publicações impressas, e *Nexo*, *FiveThirtyEight* e *Chequeado*, chamados de veículos nativos digitais. Os seis escolhidos são importantes para seus países de origem e também para o cenário internacional de jornalismo guiado por dados.

Esta dissertação conta com dois capítulos teóricos e um de abordagem empírica. O primeiro discute os estudos de jornalismo sobre fontes. A preocupação central é trazer perspectivas teóricas que permitam refletir sobre a forma como a classificação de fontes se transformou ao longo dos últimos quase 50 anos. Dessa forma, o foco da abordagem teórica parte de textos publicados principalmente a partir dos anos de 1970 (ERICSON; BARANEK; CHAN, 1987; GANS [1979] 2004; VILLAFANE; BUSTAMANTE; PRADO, 1987), estendendo-se até os dias atuais, com a produção mais recente (LAGE, 2009; PINTO, 2000; SCHMITZ, 2011). É possível perceber, em linhas gerais, um interesse maior sobre questões de relacionamento entre jornalistas e fontes e pouco focado em documentos. Para fins de análise empírica, um conceito operacional de fonte precisou ser criado.

O segundo capítulo teórico procura situar o contexto em que as produções estão inseridas, argumentando que as dimensões da sociedade em rede (CASTELLS, 2015; CASTELLS, 2016; ECHEVERRÍA, 1999) permitem o surgimento de práticas diferenciadas, como o jornalismo em rede (HEINRICH, 2008; 2011) e o jornalismo guiado por dados

(BRADSHAW, 2011; 2017; HOLOVATY, 2006; ROGERS, 2013; TRÄSEL, 2014), caracterizando um cenário em transformação. As bases teóricas do JGD são resgatadas a partir dos conceitos de jornalismo de precisão, Reportagem Assistida por Computador (RAC) e jornalismo computacional.

Por fim, no capítulo de abordagem empírica, busca-se compreender quais são as fontes jornalísticas empregadas no recorte escolhido para analisar a produção de JGD. O eixo principal deste estudo são as fontes de natureza **documental**, como bancos de dados públicos, arquivos, relatórios, projetos, publicações em sites de redes sociais e outros. O estudo empírico permitiu a identificação de um total de 315 fontes consultadas nas 60 notícias, das quais 213 são documentos. As características em comum entre as fontes documentais permitiram agregações em três tipos: **arquivo documental**, **estatística** e **reprodução**, que configuram a proposta de classificação de fontes documentais elaborada nesta dissertação.

2 FONTES, A ORIGEM DA INFORMAÇÃO NO JORNALISMO

O texto jornalístico não nasce de uma epifania. As informações que o compõem são extraídas de fontes que reúnem conhecimento sobre determinado fato ou tema abordado. Fonte remete à ideia de início do percurso de apuração jornalística para construir uma notícia e tornar um fato inteligível ao público. Para compor o processo de identificação e contextualização que permite a compreensão do evento, viabilizado pelo conteúdo noticioso, o jornalista tem de consultar fontes, isto é, pessoas ou documentos munidos de informações.

O emprego de fontes na construção noticiosa é levado em conta na primeira tese de jornalismo (PEUCER [1690], 2004) e desde então ligado à qualidade da produção textual informativa e à veracidade das informações. Já na segunda metade do Século XX, os estudos dessa área problematizam, em sua maioria, a relação de poder entre jornalistas e fontes (GANS [1979], 2004), levam em conta a rotina e os fatos que dela emergem (MOLOTCH; LESTER [1974], 2016; SIGAL, 1973 apud Santos, 1997) e por que os profissionais da imprensa elegeм com mais ou menos frequência os informantes, reproduzindo uma hierarquia de poder já existente (HALL et. al. [1978], 2016). Mesmo estando em uma posição em que conseguem testemunhar eventos, os jornalistas entendem que é conveniente consultar fontes que informem o que ocorreu ou forneçam suas versões (SIGAL, 1987).

Este capítulo dá ênfase aos trabalhos teóricos e técnicos – aqui incluídos manuais de redação – para compreender as tipologias atribuídas às fontes. Após o resgate das ideias de Peucer, são elencados estudos de fontes ao longo do século XX e do século XXI que permitem refletir sobre tais classificações.

As fontes são a quem os jornalistas recorrem para obter informações por meio de observação ou entrevista (GANS, 2004). Os dados obtidos passam por técnicas jornalísticas e por enquadramentos antes de serem publicados. O diálogo entre repórteres e fontes reside no coração da prática profissional (BROERSMA; DER HERDER; SCHOHAUS, 2015), é determinante para a qualidade da informação produzida e veiculada e constitui uma dimensão central para a compreensão do jornalismo (PINTO, 2000). É, sobretudo, uma relação de interdependência. Para Sigal (1987), as notícias formam uma combinação do que os jornalistas pensam e as fontes dizem ser a realidade, sendo que tais aspectos são mediados por organizações jornalísticas. Visão semelhante é compartilhada por Ericson, Baranek e Chan (1987). Para os autores, os jornalistas utilizam as fontes para oferecer ao público relatos da realidade e visões sobre eventos da maneira como consideram mais apropriadas em termos de cobertura. Esses relatos e pontos de vista estão sempre ligados a situações de desvio (o que foge

à rotina) e controle (em nome da estabilidade social), que têm a ver com histórias de interesse humano e com ações conectadas ao papel do jornalista como guardião da democracia.

Estudiosos do tema analisam a relação entre fontes e jornalistas sob diferentes perspectivas. A Manning (2001) interessa saber como a informação pode ser controlada e implantada, com foco no sistema político e no acesso à esfera pública. Em uma linha semelhante, Schudson (2003) aborda a importância de conhecer quem está por trás das informações e sua distribuição – o autor dá ênfase aos representantes de governo, que desempenham papéis poderosos. Chaparro (2007, p. 14), por sua vez, traz a ideia de “[...] revolução das fontes”, em que considera os informantes sujeitos institucionalizados e capacitados para produzir notícias, postura impulsionada pela tecnologia e pela ideia de sociedade da informação.

Machado (2003) discute o ciberespaço como fonte para os jornalistas, de modo que a prática de técnicas de pesquisa e apuração nas redes telemáticas seja complementar ao que se sempre se fez no trabalho de reportagem. Além disso, aponta que atores ligados a movimentos sociais e outras organizações – que não são entendidos como fontes oficiais e por isso nem sempre consultados – têm espaço de manifestação sem necessariamente passar pela mediação de jornalistas. Dessa forma, altera-se “[...] a relação de forças entre os diversos tipos de fontes porque concede a todos os usuários o status de fontes potenciais para os jornalistas” (MACHADO, 2003, p. 27). Muda, portanto, a perspectiva do trabalho de reportagem, que vai além das declarações, mas pode aprofundar-se em questões de interesse público a partir da consulta de múltiplas fontes.

Neste trabalho serão relacionados a seguir os autores que sugerem possibilidades de classificação ou subdivisões para a discussão sobre fontes. O modo de construção deste capítulo é inspirado em Santos (1997), que organiza, cronologicamente, a contribuição teórica nesta área.

A participação das fontes na construção noticiosa recebe atenção nos estudos de jornalismo desde a defesa da primeira tese na área, em 5 de março de 1690, na Universidade de Leipzig (Alemanha). Sob a autoria de Tobias Peucer, o texto escrito em latim levou o título *De relationibus novellis* e foi traduzido para o inglês como *On news reporting* (ATWOOD; BEER, 2001) e *Os relatos jornalísticos* (PEUCER [1690], 2004) em português. Em aproximadamente seis mil palavras distribuídas por 29 parágrafos, Peucer busca definir o que é notícia e como ela surge, levantando questões sobre sua natureza e o formato de apresentação das informações.

Peucer (2004) afirma que o jornalista deve falar a verdade e não acrescentar informações que fujam à realidade. Como o jornalista obtém os dados a partir de fontes, o autor traz

elementos que lembram uma hierarquização de acordo com a importância e de confiabilidade. No décimo parágrafo de sua tese, sugere a existência de três formas para obter os dados necessários à construção noticiosa: o **jornalista como observador dos fatos**, uma terceira pessoa como **testemunha** ou mesmo um **ouvinte**. É categórico ao afirmar que é mais credível o relato do narrador presencial, ou seja, aquele que testemunhou, de forma primária, o evento.

Mais adiante, no décimo quarto parágrafo da tese, aponta a necessidade de confrontar as informações obtidas inicialmente com o que dizem outras fontes, de modo que seja possível garimpar o testemunho de mais pessoas sobre um mesmo tema. Para Tambosi (2004), esse trecho revela a preocupação de Peucer com a adoção da pluralidade de fontes, processo que auxilia na contextualização dos registros noticiosos.

Avançando alguns anos na história e chegando aos anos de 1970 é possível identificar que diferentes autores buscaram classificar as fontes, suscitando reflexões sobre a formação, a construção e a divulgação dos eventos dignos de se tornarem notícia. A ênfase neste capítulo é para os estudos desenvolvidos a partir da década de 1970 porque é nesse período que as transformações sociais e tecnológicas ocorrem de maneira mais acentuada.

O reconhecimento da influência da fonte e da mediação das organizações jornalísticas, suas rotinas e convenções permeiam a percepção de Sigal (1973 apud Santos, 1997) de que a notícia é resultado de um complexo processo. Constrangidos pelas rotinas produtivas, os jornalistas tendem a buscar fontes oficiais, mesmo quando são capazes de testemunhar um evento noticiável.

Para Sigal, as informações chegam aos jornalistas por meio de três canais: **de rotina**, **informal** e **de iniciativa**. O primeiro, de rotina, inclui eventos oficiais, julgamentos, debates parlamentares, campanhas eleitorais, *releases* oficiais, relatórios, discursos e cerimônias. O segundo, chamado de informal, é constituído por encontros reservados entre jornalista e fonte, vazamentos de informação e eventos não relacionados ao poder vigente (como encontros sindicais, convenções e outros). Há, ainda, os canais de iniciativa, em que o ponto de partida são as entrevistas conduzidas pelos jornalistas – nessa subdivisão estão incluídos eventos espontâneos testemunhados pelos repórteres, pesquisa independente envolvendo livros e dados estatísticos ou análises feitas pelos próprios jornalistas. Nos três canais propostos por Sigal transitam as fontes de informação, acionadas quando necessário. Os canais de iniciativa parecem conferir ao repórter certo protagonismo ou autonomia, já que a origem da pauta e os subsídios para o seu desenvolvimento partem do jornalista, principalmente em se tratando de pesquisas.

Em se tratando de fatos públicos, Molotch e Lester ([1974] 2016) compreendem que uma notícia pode ser formada por três agências⁴: **os promotores de notícias** (*news promoters*), categoria que inclui políticos e seus porta-vozes, que identificam e elevam o *status* de uma ocorrência a situação de impacto público; os **news assemblers**, subdivisão que congrega os profissionais envolvidos no trabalho jornalístico e transformam os materiais fornecidos pelos promotores de notícias; e os **consumidores de notícias** (*news consumers*), responsáveis por acompanhar as ocorrências noticiosas e criar uma sensação de tempo público.

A atribuição de papéis que Molotch e Lester (2016) dão às fontes, aos jornalistas e ao público são fundamentais para compreender as tipificações de ocorrências públicas por eles relacionados: há os eventos de rotina, como os que são programados por fontes que desejam obter espaço na mídia; os acidentes, que são ocorrências não previsíveis; os escândalos, que em geral não são esperados e costumam ser divulgados por pessoas externas à situação; e por fim há um quarto tipo, denominado *serendipity*, no qual um agente transforma uma ocorrência não esperada em evento de rotina.

Veículos de imprensa noticiam questões de diferentes naturezas a todo momento. A partir de observações que encontrou em seus estudos nas emissoras *CBS* e *NBC* e nas revistas *Newsweek* e *Time*, Gans ([1979] 2004) procurou responder por que os jornalistas acessam as mesmas fontes. Notou que os jornalistas, especialmente pela falta de tempo no dia a dia das redações, acabam acessando um pequeno grupo de fontes regulares que foram úteis no passado e voltam a ser aproveitadas. Tal oportunidade anterior, juntamente com produtividade, credibilidade, garantia e respeitabilidade, são fatores apontados por Gans como fundamentais para que as fontes voltem a ser acessadas pelos jornalistas. A via é de mão dupla: às fontes interessa obter incentivos, poder, informação credível e proximidade social e geográfica com o jornalista para que sejam concedidas informações.

Uma das principais contribuições de Gans (2004) é fazer emergir as relações de poder entre jornalistas e fontes como ideia de negócio em que ambos representam suas instituições com interesses distintos. O autor trata de **fontes conhecidas** e **desconhecidas**: o primeiro grupo contempla fontes oficiais, como representantes de governos, enquanto o segundo congrega pessoas ligadas a conflitos ou quebra da lei.

As fontes oficiais podem ser classificadas como o que Hall et. al. ([1978] 2016) denominam de **definidores primários**, ou seja, pessoas que possuem lugar de fala reconhecido, reforçando a estrutura de poder existente, e exercem influência sobre os contornos de um texto

⁴ As agências, para Molotch e Lester (2016), são indivíduos ou grupos.

jornalístico. Movidos pela corrida contra o relógio e pelas exigências de imparcialidade e objetividade, o jornalismo dá ênfase aos que já detêm uma posição institucional privilegiada. Às pessoas com poder ou estatuto social elevado é atribuído o caráter de definidor primário na escolha de tópicos que serão notícia.

De acordo com os autores, a interpretação do definidor primário é decisiva para coberturas e debates subsequentes sobre um mesmo tema. Trata-se do enquadramento do problema, da subordinação do jornalismo aos definidores primários por serem fontes acreditadas. Enquadram-se nessa categoria os políticos, os representantes de instituições e os especialistas. Os meios de comunicação detêm apenas um caráter secundário de definição à medida que podem legitimar e reforçar a ação dos primários.

Hall et. al. (2016) exploram, ainda, a ideia de **contradefinidores**, muitos dos quais não possuem acesso ao processo de definição, mas têm como base contraideologias que geram uma definição alternativa aos definidores primários diante de um assunto controverso. No caso das relações industriais, exemplo levantado pelos autores, o papel antagônico caberia aos representantes de um sindicato, por exemplo, o que ocorreria de forma semelhante em outras áreas.

Em Villafañe, Bustamante e Prado (1987), o papel da fonte no contato com o jornalista é relativizado. Os autores trazem aos estudos o conceito de **caráter relativo da fonte** por entender que essa é apenas um intermediário, partindo do pressuposto de que a informação existe antes mesmo de chegar ao seu difusor. O informante, então, é apenas um receptor e mensageiro de dados já consolidados. Dessa forma, entendem que os termos procedência e fontes têm a mesma natureza genérica de servirem como fontes de informação, mas cumprem papéis diferentes no processo de produção de notícias. A procedência está ligada à origem da informação, ou seja, à sua existência antes mesmo de ser comunicada pela fonte jornalística. Com isso, a classificação quanto à natureza física das fontes ultrapassa o espectro humano, podendo incluir documentos, comunicados de imprensa ou instituições.

Como resultado de investigação do processo de produção da informação no rádio e na televisão, os autores apontaram a construção de uma rede de fontes como peça fundamental no fluxo de informações. A partir do trabalho de campo, Villafañe, Bustamante e Prado (1987) identificaram 13 categorias que podem ser acionadas tanto em termos de procedência quanto em termos de fontes propriamente ditas. São elas: **centros regionais**, que dizem respeito à emissora/sucursal ou trabalho em rede nos veículos; **correspondentes e repórteres**, que operam como procedência ou fontes à medida em que propõem a notícia ou selecionam e elaboram as informações; **documentação e agenda do meio**, em geral ligada às fontes oficiais

e institucionais; **fontes governamentais**, que são pessoas ou assessorias; **patronais**, que estão associadas às empresas privadas, pessoas ou assessoria; **sindicatos**, ligados a organizações de trabalhadores; **partidos políticos**; **organizações sociais**; **experts**, que são os especialistas que geram, divulgam ou comentam informações; **depoimentos**, conectados às testemunhas ou protagonistas de fato noticioso; **agências de notícias**, que oficializam e informam sobre eventos; **outros meios de informação**; e **outras fontes**.

A inserção de pesquisadores nas redações gera contribuições importantes ao campo jornalístico. Por conta de um estudo etnográfico que durou mais de 200 dias de trabalho de campo entre 1982 e 1983 e um total de 2,5 mil horas mobilizadas, Ericson, Baranek e Chan (1987) obtiveram informações sobre processos e métodos de rotina desenvolvidos no jornal *Toronto Globe and Mail* e da *CBLT* Televisão, incluindo o diálogo com as fontes jornalísticas. O trabalho de acompanhamento da rotina produtiva ainda foi complementado com entrevistas a jornalistas de Toronto, no Canadá, referentes a outras organizações jornalísticas.

Os autores abordam a existência de um **canal** entre o jornalista e a fonte. É nesse espaço que transitam os meios da fonte, como documentos e releases. Ainda que não esteja expressamente escrito no texto, o canal entre jornalistas e fontes parece ser articulado em termos de acesso do jornalista. Para os autores, existem três tipos de **canais**: **altamente estruturado**, que abrange boletins oficiais, coletivas de imprensa e releases; **menos estruturados**, quando o repórter realiza entrevista pessoalmente ou por telefone, o que pressupõe que a fonte pode resguardar e filtrar alguns aspectos; e canais **não estruturados**, quando o repórter obtém documentos de outra forma, como por meio de exclusividade com a fonte ou a partir de vazamento de informações.

Entre os temas abordados pelos autores está o trabalho de reportagem e a problematização das metodologias utilizadas para a obtenção de informações, além de como os repórteres se apropriam do conteúdo que obtêm. Segundo eles, os repórteres em geral afastavam-se do uso de documentos especialmente se havia outro jeito de conseguir a informação, como por meio de entrevista com uma pessoa (ERICSON; BARANEK; CHAN, 1987), pois havia receio de uma compreensão equivocada do documento. Além disso, havia a tendência de procurar uma fonte humana para explicar ou elaborar alguma fala em cima da publicação. Percepção semelhante parte de Sigal (1987): o autor percebeu que os repórteres preferiam entrevistar as fontes ao invés de investirem tempo em procedimentos mais complexos, como análise estatística ou busca de informações em livros e bibliotecas. O ponto máximo ao qual chegavam era consultar edições antigas do jornal para colaborar com o processo de contextualização da notícia.

Na década de 1990, semelhante aos canais de Sigal (1973 apud Santos, 1997) e à classificação de fonte conhecida e desconhecida (GANS, [1979] 2004), Erbolato (1991) afirma que todas as informações publicadas na imprensa são frutos do contato com **fontes fixas** ou **fora de rotina**. O primeiro grupo envolve pessoas e instituições consultadas com frequência, como as prefeituras, os tribunais, os centros de saúde, as associações e os sindicatos. A subdivisão fora da rotina, por sua vez, diz respeito às consultas excepcionais e esporádicas, que ocorrem quando o esclarecimento dos fatos demanda complementos.

Erbolato ainda classifica os informantes como **diretos**, **indiretos** e **adicionais**. O primeiro tipo compreende pessoas que tiveram envolvimento com alguma ocorrência, comunicados de imprensa ou notas oficiais. Na classificação dos indiretos estão pessoas que, por conta de sua profissão, têm conhecimento circunstancial de algum fato. Conteúdos publicados anteriormente, de caráter documental, também são inseridos pelo autor nessa tipologia. Já as fontes adicionais são aquelas que complementam ou ampliam as dimensões da história que se pretende abordar. Livros, almanaques, enciclopédias e similares estão entre os exemplos mencionados. Uma última tipificação dá conta da identificação das pessoas acionadas na construção de um texto jornalístico: **ostensivas** ou **indeterminadas**, que são, respectivamente, quem teve ou não seu nome divulgado.

Jornalistas e fontes, diz Santos (1997), ocupam um **campo de interação social**, também chamado de **campo da notícia**, em que estão presentes enquadramentos, linguagens e estratégias na relação entre as partes. Esse campo funciona como lugar de interação, de concorrência, de dominação e de intimidação em um jogo permanente de persuasão. O autor chegou a essa proposição depois de percorrer teorias sobre fontes, desenvolvidas por outros estudiosos, e trabalho de campo.

No início dos anos 2000, com o avanço na popularização da internet, as pessoas começaram a ter acesso aos mesmos dados brutos que os jornalistas. Houve uma mudança no cenário, portanto. Pinto (2000) define fontes como pessoas, grupos, instituições e vestígios (incluindo documentos, falas e dados) e sugere diferentes formas de classificação para compreender amplamente o quadro em que estão inseridos jornalistas e fontes. Uma dessas formas diz respeito à identificação de posições e relações sociais, em geral desenvolvidas a partir de estratégias e táticas determinadas.

Outra forma ocorre por meio da tipificação dos diferentes tipos de fontes. O autor enumera modalidades que considera importantes para dar conta de um quadro dinâmico e contextualizado. Para Pinto (2000), a classificação de fontes pode ocorrer segundo sua **natureza** (**fontes pessoais** ou **documentais**), segundo a **origem** (**públicas/oficiais** ou

privadas), segundo a **duração** (**episódicas** ou **permanentes**), segundo o **âmbito geográfico** (**fontes locais, nacionais** ou **internacionais**), segundo o **grau de envolvimento nos fatos** (**oculares/primárias** ou **indiretas/secundárias**), segundo a **atitude diante do jornalista** (**ativas/espontâneas** ou **passivas/resistentes**), segundo a **identificação** (**assumidas** ou **anônimas/confidenciais**) e segundo a **metodologia** ou **estratégia de atuação** (**proativas** ou **reativas**).

Uma terceira abordagem de Pinto (2000) prevê a catalogação e análise de problemas que emergem da atividade dos jornalistas, das fontes ou dessa relação, como é o caso de vazamentos, o uso de fontes anônimas ou de disfarces para a obtenção de dados. Um quarto modo de analisar esta troca tem relação com as intenções de jornalistas e fontes, questionando suas motivações⁵.

Observando que a relação entre jornalistas e fontes é um campo de negociação de interesses, Gomis (2004) classifica o **sistema político** como o principal fornecedor de informações à imprensa e também uma fonte inesgotável de dados. Sua justificativa é que as pessoas que compõem esse meio estão interessadas em ser alvo do olhar de cidadãos, de militantes, de partidos e também de potências estrangeiras. O autor acredita que, em princípio, os próprios políticos são as fontes interessadas e costumam fornecer dados. Quando isso não ocorre, são capazes de liberar informações sobre seus adversários. Coberturas da área da política costumam se valer de informações de bastidores; e vazamentos são comuns. Segundo Gomis (2004), o sistema político não só fornece informações como produz notícias – suas agendas despertam o interesse público e, portanto, da imprensa.

Detalhando as abordagens apresentadas, Lage [2001] (2009) emprega uma classificação mais extensa. Em seu ponto de vista, as fontes podem ser classificadas em **oficiais**, que são mantidas pelo Estado ou instituições a ele relacionadas; **oficiosas**, ligadas a uma entidade ou indivíduo, mas sem autorização para falar em seu nome; e **independentes**, que não possuem vínculo com o poder ou não tem interesse específico em determinado caso. Um segundo grupo de classificação trabalhado pelo autor trata de **fontes primárias** e **secundárias**, que são, respectivamente, quem cede informações essenciais a um texto jornalístico ou subsídios para a preparação de uma pauta ou construção de contexto. Lage ainda traz um terceiro grupo de fontes, composto por **testemunhas** e **experts**. As testemunhas são pessoas que presenciaram

⁵ Com base em resultados de pesquisas empíricas, Pinto (2000) identifica que, ao acessar as fontes, os jornalistas buscam informação inédita, conformação de informações, solução de dúvidas, lançamento de debate, fornecimento de avaliações e atribuição de credibilidade. As fontes, por sua vez, procuram a visibilidade dos meios de comunicação, a marcação da agenda pública, a busca de apoio para ideias e serviços, a prevenção ou reparo de prejuízos, a neutralização de concorrentes e a criação de uma imagem pública positiva.

determinado fato, enquanto os *experts* são fontes detentoras de conhecimento de alguma área específica; em geral são acionadas como fontes secundárias para explicar ou interpretar um evento a ser relatado ou contextualizado na produção jornalística.

A tipologia de Alcântara, Chaparro e Garcia (2005) traz: **fontes de referência**, que precisam conhecer bem o assunto tratado; **fontes de aferição**, que são independentes; **fontes informais**, que auxiliam na humanização do conteúdo; **fontes documentais**, que precisam ter boa procedência; e **fontes bibliográficas**, que ajudam a enriquecer o conteúdo.

A construção de notícias pressupõe a busca de dados, que na concepção de Jorge (2008) pode ocorrer de três formas: por meio da **observação direta do repórter**, indo ao local dos eventos; por meio de **entrevistas** com os envolvidos em um fato; ou acionando outras **fontes de informação**. A autora traz a ideia de **fontes secundárias** como forma de obter dados, pressupondo consultas a relatórios, livros, revistas, recortes, lista telefônica, fitas e filmes e a internet como um todo.

Não significa, necessariamente, que as pessoas precisam ter seu nome divulgado em um texto jornalístico para o fornecimento de dados. Ana Estela Pinto (2009) classifica as pessoas que cedem informações como: **técnicos e especialistas**, que ajudam a explicar determinado assunto; **informantes**, que fornecem observações de bastidores; **personagens da notícia**, que apresentam envolvimento direto nos fatos; e **analistas**, que não possuem envolvimento direto nos eventos, mas ajudam a colocá-lo em contexto.

Complementando as formas de classificação já exploradas, Rodrigo Alsina (2009) traz o conceito de **fontes rotineiras** e **não rotineiras**. O primeiro grupo é consultado com frequência diante de eventos corriqueiros e esperados, enquanto o segundo é acionado quando se pretende ir além das informações de fácil acesso. As fontes não rotineiras, conforme o autor, são consultadas especialmente nos casos do jornalismo de pesquisa, que envolve consultas a outros materiais.

Após a tipologia apresentada com Alcântara e Garcia, Chaparro (2009) amplia a tipologia de fontes e apresenta a seguinte listagem: **fontes organizadas**, que possuem alto nível de profissionalização na produção e divulgação de conteúdos; **fontes informais**, que não possuem atuação institucional; **fontes aliadas**, com as quais o jornalista estabelece relação de confiança para obter informações; **fontes de referência**, formada por pessoas ou instituições que detêm conhecimento reconhecido sobre questões abordadas; **fontes de aferição**, que chancelam informações; **fontes documentais**, que abrangem estudos, diagnósticos e avaliações vinculadas a pessoas e instituições; e **fontes bibliográficas**, que reúnem formas de saber científico, tecnológico, cultural ou popular.

Ao discorrer sobre fontes de informação, Kovach e Rosenstiel (2011) citam algumas possibilidades na prática jornalística: **notícias sem fonte**, que ocorre quando se trata de um evento aberto, como um discurso presidencial, e não há necessidade de atribuição específica; **jornalista como testemunha**, correspondente ao momento em que o repórter descreve em primeira pessoa o que testemunhou; **jornalista como especialista**, ou seja, o profissional que possui conhecimento avançado em determinada área; **testemunha ocular**, que corresponde à categoria composta por pessoas que descrevem com detalhes as situações presenciadas; **participantes, mas não testemunhas**, que dizem respeito às pessoas envolvidas com os eventos, porém sem ter envolvimento pessoal (caso de policiais que investigam uma situação, por exemplo); **experts e analistas**, que são acionados para fornecer contexto ou análise para uma notícia; e, por fim, há ainda as **fontes anônimas**, que participam da notícia sem ter seu nome citado.

Complexificando as classificações, Schmitz (2011) propõe uma matriz diferenciada para compreender a tipificação em busca de uma teoria das fontes. O autor distribui as **fontes** em cinco espaços: quanto à **categoria, grupo, ação, crédito e qualificação**.

Em relação à categoria, as fontes podem ser **primárias** ou **secundárias**, que, respectivamente, fornecem os fatos principais, as versões e os números que serão usados no texto jornalístico ou então ajudam a contextualizar ou desdobrar novos aspectos de uma notícia. O grupo reúne **fonte oficial** (função ou cargo público que representa órgãos mantidos pelo Estado), **empresarial** (de corporação, em geral fontes que possuem interesse comercial), **institucional** (representa organização sem fins lucrativos), **popular** (vítima ou cidadão), **notável** (artista, esportista, profissional liberal, políticos), **testemunhal** (pessoa que viu ou ouviu um fato), **especializada** (especialista, perito, intelectual) e **referência** (documentos, livros, artigos, teses, portais, sites, blogs).

Quanto à ação, Schmitz (2011) traz a **fonte proativa** (oferece notícia pronta), **ativa** (com canais de rotina e material de apoio à produção de notícias), **passiva** (caso das referências à disposição dos jornalistas para consulta) e **reativa** (fonte acredita que o jornalista só busca o lado negativo). Sobre o crédito, há pessoas **identificadas** e **anônimas** no texto jornalístico. Acerca da qualificação, a fonte pode ser **confiável** (a partir do histórico de veracidade das declarações), **fidedigna** (respeitabilidade, notoriedade e credibilidade) ou **duvidosa** (quando há suspeita sobre as informações).

Na criação de um protocolo metodológico para análise de cobertura jornalística⁶, Silva e Maia (2011) discutem a questão das fontes, que são desdobradas a partir de dois grandes grupos: informações de **primeira mão** e de **segunda mão**. O primeiro é composto por pessoas ligadas ao poder público e às instituições, cidadãos, especialistas e assessoria de imprensa, além de fontes não convencionais, como o uso do *off*, e recursos alternativos, quando os repórteres usam disfarce ou infiltração. O segundo inclui agências de notícias, outros veículos jornalísticos, publicações científicas, documentos impressos e eletrônicos, ciberespaço, reedição e republicação.

Aproximando a questão das fontes ao caráter colaborativo da internet, Canavilhas e Ivars-Nicolás (2012) levam em conta o surgimento de informações advindas de blogs e redes sociais. Os autores dividem o campo das fontes de informação em dois grupos: as **fontes da primeira Web**, com caráter informacional, constituído basicamente de informantes oficiais; e as **fontes da Web 2.0**, que tem como pressuposto a colaboração e a contribuição por meio de inteligência coletiva, predominantemente formada por fontes não oficiais. As fontes 2.0 são aquelas que produzem conteúdo por livre e espontânea vontade, em um ambiente de colaboração. Os autores classificam como fontes 2.0: blogs e microblogs, redes sociais, fóruns ou listas de e-mails, *wikis* (enciclopédias colaborativas publicadas na internet), chats e repositórios multimídia, como *Soundcloud* e *YouTube*.

Em um contexto recente, buscando analisar a pluralidade e diversidade de vozes na programação da rádio *BandNews Fluminense FM*, Kischinhevsky e Chagas (2017) fazem uma revisão teórica dos estudos de fontes e apresentam sua própria classificação, que é composta por: **oficiais**, ou seja, pessoas que ocupam cargos eletivos, bem como funcionários dos poderes Executivo, Legislativo e Judiciário e membros de fundações, autarquias e empresas públicas; **empresariais**, que provêm de associações representativas de setores comercial, financeiro, agrícola e outros; **institucionais**, correspondentes aos integrantes de organizações do terceiro setor, organismos multilaterais, movimentos sociais e outros; **testemunhais**, ou seja, personagens que presenciaram eventos; **populares**, que em geral são apresentados como vítima de determinada situação; **especialistas**, que possuem conhecimento em determinado tema; e **notáveis**, categoria correspondente a celebridades, artistas, esportistas e comunicadores.

Revisadas as formas de classificação de fontes na produção acadêmica, direciona-se o olhar, neste momento, para os manuais de redação de jornais nacionais e internacionais para

⁶ O protocolo de análise de cobertura jornalística é edificado a partir de três níveis e permite analisar a cobertura de jornais impressos. O primeiro nível trata de marcas da apuração, e nele estão incluídas as fontes, o segundo aborda as marcas da composição do produto e o terceiro traz os aspectos de contexto.

compreender de que forma são abordadas as tipologias. Ao todo, seis manuais de redação e estilo foram consultados⁷. Em geral, apresentam explicações de como o jornalista deve tratar uma fonte e apontam casos em que há liberdade para sua não identificação. Entre os principais jornais brasileiros, a *Folha de S. Paulo* é o único que constrói sua própria tipologia de fontes.

O jornal enumera quatro tipos que exigem diferentes procedimentos a serem adotados pelos jornalistas durante o processo de pauta e reportagem. A saber: a **fonte tipo zero** tem tradição de exatidão, caso de enciclopédias e documentos emitidos por instituições credíveis. Não necessariamente precisa ser cruzada com outras fontes. A **fonte tipo um** refere-se somente a pessoas e enquadram-se nessa categoria fontes que possuem histórico de confiabilidade. Este tipo de fonte possui conhecimento de causa, está próxima dos fatos e em princípio não tem interesse imediato na sua divulgação. Já a **fonte tipo dois** tem os mesmos atributos da fonte tipo um, porém sem o histórico de confiabilidade. Desse modo, precisa ser cruzada com outras referências para ser utilizada. Por último existe a **fonte tipo três**, com menor grau de confiabilidade. Embora tenha conhecimento sobre os fatos, em geral possui interesses específicos em sua divulgação. Por isso, sua utilização, sozinha, não é recomendada. “Na Folha, há dois caminhos para a informação de fonte tipo três: funcionar como simples ponto de partida para o trabalho jornalístico ou, na impossibilidade de cruzamento com outras fontes, ser publicada em coluna de bastidores [...]” (FOLHA DE SÃO PAULO, 2010, p. 38).

A *Folha* realiza sua classificação conforme a confiabilidade das fontes consultadas, incluindo pessoas e documentos. Mesmo que os quatro tipos sirvam de parâmetro, em última instância é o próprio jornalista que relaciona a origem de suas consultas com cada uma das categorias.

O quadro teórico a seguir resume as abordagens teóricas sobre classificações de fontes. O Manual da *Folha de S. Paulo*, por ter caráter técnico, aparece ao final.

⁷ Além do Manual da Redação, da Folha de S. Paulo, foram consultados os manuais dos jornais *O Estado de S. Paulo*, *O Globo*, *The New York Times*, *La Nación* e *El País*. O material produzido pela *Folha* é o único que apresenta classificação. Os demais, de maneira geral, preocupam-se em orientar o jornalista em seu relacionamento com as fontes (EL PAÍS, 2008), trazem informações de como proceder na coleta e no tratamento dos dados (O ESTADO DE S. PAULO, 2017; O GLOBO, 2001), a maneira como os dados são empregados em gráficos e mapas (LA NACIÓN, 1997), em que circunstâncias é permitido usar fontes anônimas (EL PAÍS, 2008; LA NACIÓN, 1997; O ESTADO DE S. PAULO, 2017; SIEGAL e CONNOLLY, 2015), quando é preciso confrontar informações com mais de uma fonte (EL PAÍS, 2008; LA NACIÓN, 1997) e o cuidado necessário ao resgatar e usar como material de pesquisa os conteúdos antigos publicados pelo jornal (O GLOBO, 2001).

Quadro 1 – Abordagens teóricas de classificação de fontes

| Autor | Classificação |
|--------------------------------------|---|
| Sigal (1973 apud Santos, 1997) | Fontes operam nos canais de rotina, informal e de iniciativa |
| Gans ([1979] 2004) | Pessoas conhecidas e desconhecidas |
| Hall et. al. ([1978] 2016) | Definidor primário, secundário e contradefinidor |
| Ericson, Baranek e Chan (1987) | Canais altamente estruturados, menos estruturados e não estruturados |
| Villafañe, Bustamante e Prado (1987) | Fonte e origem da fonte com as seguintes categorias: centros regionais; correspondentes e repórteres; documentação e agenda do meio; fontes governamentais; patronais; sindicatos; partidos políticos; organizações sociais; experts; depoimentos; agências de notícias; outros meios de informação; e outras fontes. |
| Pinto, M. (2000) | Fontes podem ser classificadas segundo à natureza, origem, duração, âmbito geográfico, grau de envolvimento, atitude diante do jornalista, identificação e metodologia ou estratégia de atuação |
| Alcântara, Chaparro e Garcia (2005) | Fontes organizadas, de referência, de aferição, informais, documentais e bibliográficas |
| Jorge (2008) | Jornalista como observador, testemunhas, outras fontes e fontes secundárias |
| Pinto, A. (2009) | Fontes como técnicos e especialistas, informantes, personagens da notícia e analistas |
| Rodrigo Alsina (2009) | Fontes rotineiras e não rotineiras |
| Lage (2009) | Oficiais, oficiosas e independentes; primárias e secundárias; testemunhas e experts |
| Chaparro (2009) | Fontes organizadas, informais, aliadas, fontes de referência, fontes de aferição, documentais e fontes bibliográficas |
| Kovach e Rosenstiel (2011) | Notícias sem fonte; jornalista como testemunha; jornalista como especialista; testemunha ocular; participantes, mas não testemunhas; experts e analistas; fontes anônimas |
| Schmitz (2011) | Classificação de fontes quanto à categoria, grupo, ação, crédito e qualificação |
| Silva e Maia (2011) | Fontes de primeira mão e segunda mão |
| Canavilhas e Ivars-Nicolás (2012) | Fontes de primeira Web e fontes da Web 2.0, que tem como pressuposto a colaboração |
| Kischinhevsky e Chagas (2017) | Oficiais, empresariais, institucionais, testemunhais, populares, especialistas e notáveis |

| | |
|-------------------------------------|-----------------------------------|
| Manual da Folha de São Paulo (2010) | Fontes tipo zero, um, dois e três |
|-------------------------------------|-----------------------------------|

Fonte: A autora (2018)

Apresentadas as contribuições dos teóricos, percebe-se que o destaque coube à influência das fontes pessoais na relação com os jornalistas. O levantamento bibliográfico das tipologias nos estudos sobre fontes permitiu perceber que o emprego de documentos como livros, artigos, estatísticas, notas e outros são levados em conta nas abordagens desenvolvidas até o momento, mas em menor escala.

Acredita-se, então, na necessidade de refletir sobre a variedade de documentos que compõem o caráter documental das fontes. Entre os estudos mais recentes do qual esta perspectiva está próxima tem-se o trabalho de Pinto (2000), que classifica as fontes de acordo com a natureza pessoal ou documental; Chaparro (2009), que leva em conta a tipologia de fontes documentais, grupo que abrange estudos, diagnósticos e avaliações, e fontes bibliográficas, que são teses, livros, artigos científicos e publicações semelhantes; e Schmitz (2011), que trata de fontes de referência, como documentos, livros, artigos, teses, portais, sites e blogs. Entretanto, tais estudos não contemplam a variedade de fontes empregadas nas práticas do jornalismo guiado por dados.

Para fins de operacionalização no estudo empírico a ser desenvolvido no quarto capítulo, a definição de fonte adotada nesta dissertação é a seguinte: **fonte é pessoa ou documento, incluídas as bases de dados, cuja informação ou conhecimento foi empregado de forma explícita⁸ em uma notícia, seja ao longo do texto, seja por meio de escrita hipertextual ou apenas como referência para uma visualização.**

Este capítulo discutiu como os teóricos classificaram as fontes jornalísticas desde a primeira tese da área e em parte do século XX, especialmente a partir dos anos de 1970, e do século XXI, incluindo estudos recentes. Para entender o contexto contemporâneo, examina-se, a partir de agora, as transformações que levaram à formação da sociedade em rede e das práticas jornalísticas nela inseridas, como o jornalismo em rede e o jornalismo guiado por dados, em que as informações estruturadas em bancos de dados e planilhas também servem de matéria-prima para a elaboração de conteúdo.

⁸ Admite-se que, durante o processo de elaboração e execução da pauta, existe a possibilidade de o jornalista consultar fontes que não necessariamente são citadas no texto. No entanto, como esta é uma análise de produto final, não há como saber se esse procedimento ocorreu. Dessa forma, opta-se por estudar as fontes consultadas e referenciadas de forma clara no conteúdo jornalístico.

3 A SOCIEDADE EM REDE, O FENÔMENO *BIG DATA* E O JORNALISMO

As formas de organização social do século XX e do século XXI nascem distintas e operam em lógicas diferentes. Enquanto a primeira está baseada principalmente na indústria e suas máquinas, em uma situação linear de trabalho, com princípio, meio e fim, a segunda está ancorada sob uma perspectiva sem fronteiras geográficas e temporais, em uma estrutura de rede, na qual questões hierárquicas sofrem apagamento. Tais modos de organização são chamados, respectivamente, de **sociedade industrial** e **sociedade em rede**. As fronteiras não estão tão bem definidas: os dois modos podem coexistir e se sobressair em um ou outro aspecto, dependendo do cenário onde estão inseridos. É fato que as condições tecnológicas contribuem para o avanço de uma sociedade global interconectada, especialmente por conta das práticas vigentes, mas não se trata de uma realidade uniforme.

O termo adotado neste estudo para descrever as práticas sociais da atualidade é **sociedade em rede** porque dá conta de uma estrutura social global construída “[...] em torno de redes ativadas por tecnologias de comunicação e de informação processadas digitalmente e baseadas na microeletrônica” (CASTELLS, 2015, p. 70). Acredita-se, sobretudo, que a ideia de rede está conectada ao jornalismo praticado nos dias de hoje, por isso faz sentido a utilização do termo neste contexto, ao invés de sociedade pós-industrial, por exemplo. Uma rede, conforme o autor, é composta por um conjunto de nós interconectados que interagem. A importância de um nó está em sua capacidade de contribuição para a eficácia da rede, e não em sua especificidade. Na vida social, as redes são estruturas de comunicação e processam fluxos.

As transformações tecnológicas começaram a se desenhar nos anos de 1970 como peça chave para o desenvolvimento da sociedade em rede (CASTELLS, 2016), impactando diferentes setores, desde os mais amplos, como a economia, em direção aos mais específicos, como o trabalho e as relações interpessoais. A sociedade em rede, diz Castells (2015, p. 37), é construída em torno de “[...] redes pessoais e organizacionais movidas por redes digitais e comunicadas através da internet e de outras redes de computadores”.

Um dos conjuntos de práticas oriundos deste modo de organização social é o **jornalismo em rede** (HEINRICH, 2008; 2011), com procedimentos descentralizados e diferentes usos de fontes de informação em um cenário *big data*, ou seja, com grandes quantidades de dados à disposição. Essa condição traz novidades ao jornalismo à medida que pressupõe maior possibilidade de consultas aos dados como fonte para a produção de conteúdo.

Com raízes ainda na era industrial, mas cujo desenvolvimento ganhou força a partir dos anos 2000, por conta das políticas de transparência pública, o jornalismo guiado por dados

(BRADSHAW, 2011; HOLOVATY, 2006; ROGERS, 2013a; TRÄSEL, 2014) é um conjunto de práticas que envolve desde a coleta até a visualização de informações e deriva do jornalismo de precisão (MEYER [1973] 2002) e da Reportagem Assistida por Computador (COX, 2000). Outro termo utilizado para descrever essas técnicas é o jornalismo computacional (COHEN, HAMILTON e TURNER, 2011; STAVELIN, 2013), que dá ênfase à interseção com a programação.

A nomenclatura adotada nesta dissertação é o **jornalismo guiado por dados**⁹ (JGD) – que também pode ser chamado simplesmente de jornalismo de dados –, prática que se atualiza na sociedade em rede a partir da maior disponibilidade de informações em bancos de dados, entre eles os públicos – portanto, com novas possibilidades do uso de fontes – e por conta da facilidade de acesso a ferramentas de visualização e publicação de conteúdo. Para fins de contextualização de cenário, no entanto, é preciso olhar para a teoria e compreender como as mudanças sociais começaram a acontecer.

Este capítulo discute, em um primeiro momento, as transformações ocorridas em um cenário de sociedade em rede. As mudanças aqui relatadas ocorrem em nível internacional, mas em geral têm início nos Estados Unidos e na Europa para então migrarem em direção a outras regiões do mundo. Na sequência são discutidas as transformações que chegaram ao jornalismo, que tem de adaptar sua organização e seus procedimentos operacionais na produção de conteúdo. Neste contexto ganha espaço o JGD e as possibilidades que abre à prática jornalística.

3.1 Bases para um novo paradigma

Ao final do século XX, o mundo viveu o que Castells (2016, p. 87) denomina um dos “[...] raros intervalos da história” com a transformação da cultura material, desenvolvida na sociedade industrial, por mecanismos de um novo paradigma tecnológico que orbita em torno da tecnologia da informação – identificada como o conjunto de tecnologias em microeletrônica, computação, telecomunicações e muito além disso, como a própria engenharia genética e suas aplicações.

⁹ Jornalismo guiado por dados e jornalismo de dados são sinônimos. O termo jornalismo guiado por dados (*data-driven journalism*) é empregado nesta dissertação porque compreende-se que traduz com mais propriedade seu significado. Jornalismo guiado por dados pressupõe a compreensão de que o conteúdo jornalístico teve como ponto de partida de seu desenvolvimento informações de bases de dados, planilhas ou portais da transparência. Em outras palavras, significa dizer que os dados são protagonistas da notícia.

A gestação desse paradigma se deu nos anos de 1970, principalmente nos Estados Unidos, que desde a década anterior associava o progresso tecnológico à cultura da liberdade e da inovação. A partir dos anos de 1980, o processo foi importante na reestruturação do sistema capitalista. A revolução concentrada nas tecnologias da informação passou a remodelar a base da sociedade em ritmo intenso, estabelecendo uma nova relação entre economia, sociedade e Estado, áreas por meio das quais a comunicação é moldada e tem suas relações de poder construídas (CASTELLS, 2015).

A influência dessa revolução sobre o capitalismo diz respeito à flexibilidade de gerenciamento, descentralização das empresas e organização em rede. A reestruturação do modo capitalista de produção no final do século XX é a base do informacionalismo (CASTELLS, 2016), em que a fonte de produtividade está na tecnologia. O capitalismo não é apagado, mas este formato se difere do industrialismo, voltado ao crescimento da economia ao maximizar os processos de produção.

A especificidade do modo informacional de desenvolvimento reside na ação de conhecimento sobre conhecimento como principal fonte de produtividade. Em outras palavras, Castells (2016) explica que se trata da interação entre as fontes de conhecimento tecnológico e a aplicação da tecnologia para aprimorar o processamento de informação. O que caracteriza a revolução tecnológica, esclarece o autor, não é a centralidade desses conhecimentos e da informação por meio deles gerada, mas seu ciclo de retroalimentação cumulativa entre uso e inovação. O ponto principal da teorização de Castells é que a tecnologia não opera sozinha, mas exerce influência sobre a informação, que representa sua matéria-prima.

A difusão do uso de computadores¹⁰ e a chegada da internet privada na segunda metade dos anos de 1990 alteraram a forma de sociedade organizada basicamente de forma industrial. Com a internet, há uma linguagem universal digital, e a lógica das redes do sistema de comunicação geraram condições tecnológicas para a comunicação global e horizontal que se tem hoje. O emprego da rede de computadores, de *softwares* de código aberto, de transmissão digital e da expansão da internet após sua privatização provocou transformações em todos os campos. Uma das mudanças sociais mais aparentes trazidas pela sociedade em rede está justamente na comunicação (CASTELLS, 2016). Até então a tendência das empresas do setor era de utilizar uma organização de trabalho hierárquica (ANDERSON; BELL; SHIRKY),

¹⁰ Castells (2016) diz que a primeira revolução da tecnologia da informação vem da Califórnia, nos Estados Unidos, com a criação do Vale do Silício, onde estão situadas empresas que buscam o desenvolvimento científico e tecnológico a partir de inovação. Naquele local foram desenvolvidos o circuito integrado, o microprocessador e o microcomputador, para ficar em alguns exemplos.

porém tal modelo já não faz sentido com as redes horizontais de comunicação em torno da internet.

A sociedade em rede também parece promover a interação entre poder e contrapoder, que vem, respectivamente, de instituições consolidadas e a partir de demanda da sociedade civil. Um dos exemplos levantados por Castells (2015) é o do *WikiLeaks*¹¹, que segundo o autor configura um caso de interação entre poderes nas redes digitais de comunicação: o sigilo do governo, em meio a questões de segurança nacional sem os processos legais necessários, e as oportunidades para as pessoas monitorarem o poder público e seus representantes e o que eles fazem com essas informações. O próprio jornalismo encontrou o seu lugar em escândalos desse porte à medida que utilizou o vazamento de documentos como fonte de notícia e diálogo com outros veículos de comunicação a partir do trabalho em rede.

Echeverría (1999) e Castells (2016) concordam com relação ao impacto das tecnologias da informação e da telecomunicação e a decorrente criação de um novo espaço social que difere dos entornos naturais e urbanos em que os seres humanos estão acostumados a viver. Echeverría trabalha com a ideia de três entornos: o primeiro é o corpo humano, um ambiente natural – sua fronteira tem relação com a estrutura sensorial, da atuação com o exterior. O segundo entorno é cultural e social, por isso pode ser chamado de urbano. Nele são desenvolvidas formas sociais as mais diversas, como a roupa a ser vestida, o contato com a família, o que é o indivíduo e sua relação com os mercados. O segundo entorno está encoberto por uma produção técnica industrial, como maquiagem, roupas e calçados. O terceiro entorno, por fim, é um espaço social conectado de inter-relações, o equivalente à sociedade em rede de Castells (2016). Trata-se de um espaço desterritorializado dependente de inovações tecnológicas e propriedades que se modificam de acordo com novos avanços científicos. Uma de suas principais manifestações é a Telépolis¹², nome que batiza uma cidade global, eletrônica e digital, que transborda fronteiras geográficas e políticas.

O terceiro entorno de Echeverría (1999) tem em sua constituição um espaço de inter-relação entre seres humanos – hoje, se o conceito fosse atualizado, provavelmente incluiria não

¹¹ O *WikiLeaks*, fundado por Julian Assange, é uma organização internacional que veicula informações não públicas, mas em geral atreladas aos governos, a partir de fontes anônimas protegidas por meio de criptografia (CASTELLS, 2015).

¹² A Telépolis de Echeverría (1999) não é um *habitat*, pois no terceiro entorno não se vive, mas se atua e suporta a interação econômica. Existe uma transformação profunda não só nos setores da comunicação e da informação, mas também na produção, no comércio, no consumo e na própria informação. Os mantenedores da Telépolis, no entanto, são empresas multinacionais ligadas à edificação do terceiro entorno, como companhias telefônicas, audiovisuais, bancárias e informáticas, já que esse entorno demanda recursos tecnológicos, humanos e financeiros com benefícios econômicos que podem surgir. Neste caso, alerta o autor, pode ocorrer uma nova forma de aristocracia no terceiro entorno, a quem Echeverría chama de senhores do ar – que lembram os senhores medievais que tentavam expandir seus territórios e poderes.

humanos. As alterações do terceiro entorno escapam às noções tradicionais de tempo e espaço da sociedade industrial. O tempo, medida vivida de forma diferente pelos seres humanos, de acordo com a estruturação de suas vidas, também sofre alteração na sociedade em rede. O tempo do relógio, que remete à sociedade industrial, é substituído pelo tempo atemporal que, segundo Castells (2016), é desenvolvido quando há uma perturbação sistêmica na ordem sequencial das práticas sociais em contextos como o da sociedade em rede. Além disso, a relação com o tempo tem a ver com o uso que se faz das tecnologias de comunicação e informação.

Se o homem transforma seus entornos de acordo com suas necessidades, faz algum sentido pensar em um terceiro entorno como dependente de inovações tecnológicas. Echeverría (1999) identifica 20 aspectos¹³ presentes nesse espaço e traça paralelos com o primeiro e o segundo entorno combinados. Dessa maneira, o terceiro entorno é informacional digital à medida que as propriedades dos objetos são mais informacionais do que materiais e com maior capacidade de armazenamento; dialoga com a ideia de fluxos eletrônicos com velocidade maior que no primeiro e segundo entornos, apesar da instabilidade nas redes telemáticas.

O mundo é construído a partir da lógica de fluxos e o que interessa é a conexão de rede entre o local e o global. Todas as grandes mudanças, diz Castells (2016, p. 24), têm como base as transformações de tempo e de espaço na experiência humana. Na sociedade em rede, perde força a ideia de espaço como território delimitado e entra em cena a ideia de “espaço de fluxos¹⁴”, que dissolve o tempo e envolve um suporte material de práticas sociais simultâneas comunicadas à distância. Castells (2016) indica que as relações de trabalho, situadas no cerne da estrutura social, representam o principal fator de impacto do paradigma informacional e do processo de globalização sobre a sociedade em geral. Se na concepção do autor o conhecimento e a informação são as principais fontes de produtividade nesta nova economia, é possível afirmar que o jornalismo é certamente uma das áreas mais afetadas, já que trabalha com essas questões diariamente.

¹³ Os 20 aspectos que compõem o terceiro entorno são os seguintes, conforme Echeverría (1999) – entre parênteses, o paralelo com o primeiro e o segundo entorno combinados: distante (próximo); reticulado (*recintual*); informacional (material); representacional (presencial); artificial (natural); multicrônico (sincrônico); compressão (extensão); fluência eletrônica (mobilidade física); circulação rápida (circulação lenta); assentamento no ar (assentamento em terra); instabilidade (estabilidade); globalidade (localidade); bissensorial (multissensorial); memória artificial externa (memória natural interna); digital (analógico); integração semiótica (diversificação semiótica); heterogeneidade (homogeneidade); transnacionalidade (nacionalidade); interdependência (autossuficiência); consumo (produção).

¹⁴ Fluxo, de acordo com Castells (2016), é uma sequência repetitiva, intencional e programável de interação entre posições fisicamente desarticuladas, mantidas por atores sociais nas estruturas econômica, política e simbólica da sociedade. O espaço de fluxos, de acordo com Castells (2015), se refere à possibilidade tecnológica e em termos de organização para praticar a simultaneidade sem contiguidade. Também prevê a interação assíncrona, à distância. O autor exemplifica que mercados financeiros e redes de mídia são alguns grupos e setores que se organizam ao redor do espaço de fluxos.

3.2 O jornalismo em rede

Dentro da perspectiva de sociedade em rede emerge o conceito de jornalismo em rede (*network journalism*)¹⁵, que Heinrich (2008) classifica como uma prática descentralizada e não linear desenvolvida a partir do século XXI e que traz novos formatos importantes para a produção e a disseminação de notícias. A transformação, de acordo com a autora, ocorre a partir de dois processos: da globalização e dos avanços tecnológicos na forma de digitalização. A via de mão única do fluxo de notícias, do emissor em direção ao receptor, principal prática na era industrial, é substituída pela estrutura de rede. Nesta fase mais recente do jornalismo, há flexibilidade dos processos de produção, redução de funções e convergência entre mídias – em oposição ao cenário anterior, que se manifestava na hierarquização presente nas redações e na fragmentação de funções dos jornalistas (FONSECA, 2008).

De acordo com Primo (2011), as práticas jornalísticas atuais envolvem um número maior de produtores e distribuidores de notícias, mas muitos sequer fazem parte de organizações jornalísticas. As empresas situadas nesta malha interconectada do jornalismo em rede sofrem transformações e precisam se adaptar, promovendo um trabalho aberto como uma rede. Deixam de lado as estruturas verticais para a adoção de estruturas horizontais, com organização em torno dos processos e gerenciamento em equipe. Ignorar essas transformações, alerta Heinrich (2008), pode levar ao isolamento na sociedade em rede, em que os usuários misturam fontes de informação e buscam notícias sob demanda. Por isso, características de interação e colaboração demandam incorporação à prática jornalística.

O computador foi um dos equipamentos que transformou os processos jornalísticos. Há registros de seu uso nos anos de 1950 nos Estados Unidos para prever resultados de eleições (BOUNEGRU, 2012), mas a popularização ficou mesmo para os anos de 1980. De acordo com Cox (2000), que fala do cenário nos Estados Unidos, inicialmente os repórteres compravam e utilizavam seus próprios computadores, mas depois as organizações jornalísticas passaram a adquirir os equipamentos. A internet comercial, que se proliferou a partir da segunda metade dos anos de 1990, igualmente impactou a atividade.

Atualmente é possível pensar em um sistema de redes horizontais de comunicação organizado em torno da internet e que abrange os veículos de comunicação tradicionais, mas

¹⁵ De acordo com Heinrich (2008), abordagens anteriores que traziam conceito de jornalismo em rede estavam mais ligadas a conceitos como o de jornalismo online e de jornalismo cidadão. A autora sugere a aplicação do termo jornalismo em rede (*network journalism*) para designar a estrutura dinâmica da produção de notícias que substitui a ideia tradicional de mídia. Descreve uma nova forma da mídia como sistema, impactando na distribuição de notícias impressas, televisão, rádio, televisão, assim como o jornalismo online. Jornalismo em rede torna-se um princípio básico que explica a estrutura do sistema jornalístico na era digital.

também apresenta um espaço de horizontalidade com usuários e atores não humanos que influenciam no processo. Heinrich (2008) observa que as redes transformam o jornalismo profissional em diversas regiões do mundo e de maneira muito similar, incluindo a imprensa de referência, outros veículos menos tradicionais e blogs. A propósito, a estrutura de rede permitiu o surgimento de iniciativas denominadas de nativos digitais, termo que compreende os veículos criados especificamente para a internet, sem derivar de uma publicação impressa ou manter uma versão em papel.

Na visão de Castells (2016), os veículos, individualmente, tornam-se organizações estruturadas internamente em rede e que se conectam globalmente a outras redes na internet – em termos de produção e difusão de conteúdo, portanto. Em relação ao jornalismo, a internet contribuiu para a chegada de um novo ecossistema (ANDERSON; BELL; SHIRKY, 2012) horizontal, com um diálogo próximo entre público e jornalista. Esse público por vezes é quem lança o primeiro alerta que faz o jornalista buscar as informações sobre determinado fato e torna-lo conhecido. O jornalista, por sua vez, tem a possibilidade de considerar o usuário um aliado para obter informações, fotos e vídeos. Ganha espaço um processo de colaboração que já existia, mas se potencializa pela tecnologia, que permite agilidade na troca de informações. Nessa esfera, um crescente fluxo global de notícias conecta permanentemente jornalistas, fontes e audiência.

O jornalismo não está no centro da sociedade, mas integrado a um processo de comunicação em rede que tem seus próprios desafios, como a esfera global com um grande número de competidores, pressões de tempo e espaço e maior quantidade de informação a ser filtrada e contextualizada (HEINRICH, 2011). Considerando que o jornalismo está conectado a uma forma de produção cultural, é preciso levar em conta, como diz a autora, que há alterações no processo de coleta, processamento e apresentação das notícias. Para Briggs (2007), há possibilidades de se fazer um jornalismo qualificado a partir da incorporação de práticas de tecnologia e transparência à rotina jornalística. Percepção semelhante é de Echt (2014), para quem iniciativas como o jornalismo guiado por dados promove um novo contrato entre veículo e leitor a partir da abertura de informações, isto é, a transparência no método, passível de ser empregada de modo que o cidadão possa refazer o caminho adotado na apuração da notícia.

Anderson, Bell e Shirky (2013) indicam que todo o processo de produção das notícias deverá ser repensado na direção de abrir espaços para novos parceiros, aproveitar os dados disponíveis e também construir conteúdo a partir do que há de melhor a ser extraído da inteligência de humanos e máquinas. Na visão de Heinrich (2011), este cenário não representa

uma ameaça à profissão do jornalista, mas é uma oportunidade para rever e reorganizar processos por meio de conexões com audiências e fontes.

Para Bradshaw (2017), são pelo menos quatro os principais papéis do jornalista nesta era em rede: 1) dar voz aos que não têm voz, o que pode ser feito principalmente por meio do uso de *smartphones*, a partir do jornalismo móvel; 2) trazer à tona o que está escondido na *Web* a partir da combinação de fontes e sua utilização em técnicas do jornalismo guiado por dados e da interatividade nesses produtos. Assim, é possível identificar problemas que afetam as comunidades e personalizar a experiência do usuário; 3) conectar comunidades para que seja possível localizar conteúdo fora da bolha de consumo, em que só se tem contato com os tipos de materiais com os quais mais se gera contato; 4) verificar as informações que circulam na *Web* e atuar pelo desmantelamento de informações falsas, um dos maiores desafios atuais do jornalismo.

Isso não quer dizer que os veículos tradicionais não sejam importantes no cenário de jornalismo em rede. Eles são centrais, mas apropriados para a distribuição de informações, trazendo à cena novos atores e proporcionando a reconfiguração do jornalismo (ZAGO, 2014). Os veículos tradicionais, então, cumprem um papel essencial para que outros nós de uma mesma rede possam circular e recircular um fato jornalístico. Em sua tese, Zago (2014) utiliza o termo jornalismo em rede para tratar do momento de configuração dessa prática profissional na contemporaneidade conectada à metáfora de rede e ligada às redes sociais por meio da participação do público.

O jornalismo, para Heinrich (2011), continua sendo um sistema cuja principal atribuição está na distribuição de notícias, mas é verdade que existem mudanças na coleta e na disseminação de informações em um sistema reestruturado de rede. Veículos tradicionais e nativos digitais têm a oportunidade de agregar em suas rotinas produtivas novas ferramentas para a coleta e o trabalho com as informações estruturadas por vezes em grandes quantidades, como nas bases de dados. Isso vai ao encontro do que diz Castells (2016), para quem a transformação tecnológica tem capacidade para ser gerada, armazenada, recuperada, processada e transmitida. É neste cenário que se desenvolve o jornalismo guiado por dados e suas especificidades em um dos aspectos mais básicos e primários da atividade: o uso de fontes.

3.3 Jornalismo guiado por dados (JGD): contexto e definições

A disponibilidade de informações em quantidade sem precedentes, característica da sociedade em rede, é o que determina o fenômeno *big data*¹⁶. O termo é utilizado para definir grandes quantidades de informação produzidas por pessoas, coisas e suas interações (BOYD; CRAWFORD, 2012). De acordo com Lewis e Westlund (2014), trata-se de um fenômeno social, cultural e tecnológico que pressupõe um complexo emaranhado de dados digitais em abundância. Por conta desses fatores, provoca transformações e mudanças de procedimentos em diferentes áreas, como a comunicação, a saúde, a economia, as políticas públicas e outras. Está ligado a diferentes finalidades, pois abrange interações nas redes sociais, registros de saúde, sequências genéticas (usadas na medicina, por exemplo), rastros deixados por usuários e registros governamentais (BOYD; CRAWFORD, 2012; PARASIE, 2014).

Especificamente no jornalismo, a ideia de *big data* evoca atividades ligadas a repositórios de dados, algoritmos, métodos computacionais e processos semelhantes (LEWIS e WESTLUND, 2014). A expressão *big data* não representa apenas uma transição tecnológica, mas constitui um fenômeno sociotécnico que possui origens e implicações culturais, econômicas e políticas (BOYD; CRAWFORD, 2012; LEWIS; WESTLUND, 2014). Para Boyd e Crawford (2012), deriva da interação de três fatores: tecnologia, no que toca à computação e ao uso do rigor algorítmico para coletar, analisar e comparar dados; análise, em relação à identificação de padrões em meio a grandes quantidades de dados; e mitologia, que dá nome à crença de grandes quantidades de dados para oferecer uma forma mais aprimorada de inteligência e conhecimento envoltos na verdade, objetividade e precisão.

O que fazer com os dados no jornalismo é uma questão constantemente debatida entre os estudiosos do tema. Bradshaw (2012, p. 2) admite que os termos “jornalismo” e “dados” são problemáticos e afirma que, duas décadas atrás, o segundo termo estava mais ligado a uma coleção de números dispostos em planilhas. Atualmente, contudo, a disponibilidade de dados é muito maior, motivada por políticas de transparência pública e recursos tecnológicos que

¹⁶ De acordo com Press (2013, s.p.), a primeira referência ao que hoje se denomina *big data* é de 1941. Para quantificar um grande crescimento no volume de dados empregava-se o termo “explosão informacional” (*information explosion*, em inglês). Nos anos de 1940, autores descreviam o crescimento no volume de dados em bibliotecas. Nos anos de 1960, conforme levantamento de Press (2013), os olhares estavam voltados ao aumento na produção científica publicada em periódicos. Em 1981, na Hungria, o Centro Oficial de Estatísticas daquele País começou a medir o volume de informações em bits. Nos anos de 1990 teve continuidade a tendência de observar o agrupamento de dados disponíveis no mundo. Press afirma que o uso do termo *big data* ocorreu pela primeira vez em um artigo publicado no ano de 1999 por Steve Bryson, David Kenwright, Michael Cox, David Ellsworth e Robert Haimes e denominado “Exploração visual de conjunto de dados de gigabyte em tempo real” (Em inglês *Visually exploring gigabyte data sets in real time*). Naquele mesmo ano, parte desses autores participaram de um painel que questionava o que era melhor para o *big data*: interação ou automação.

facilitam a identificação de rastros a partir das atividades na internet, como ocorre no *big data*. Para Lewis e Westlund (2014), tal fenômeno traz a oportunidade de reflexão sobre a epistemologia do jornalismo, ou seja, sua relação com o que agrega/gera de conhecimento, incluindo a potencialidade de pensá-lo mais próximo da ciência – ou da precisão no método jornalístico, como defende Meyer (1991).

As formas de jornalismo na era *big data* são o **jornalismo guiado por dados (JGD)** e o **jornalismo computacional** (LEWIS; WESTLUND, 2014). Segundo Parasie (2014), cresce o número de jornalistas que veem nas bases de dados e nos algoritmos meios apropriados para guiar uma investigação. O termo jornalismo guiado por dados, conforme Bounegru (2012), passou a ser mais utilizado no início do Século XXI e atualiza uma nova fase do jornalismo com volumes de dados disponíveis online – combinados com ferramentas sofisticadas para trabalhar com essas informações. No JGD, as bases de dados e as planilhas utilizadas estão estruturadas de modo que possam ser lidas por computador, o que é fundamental para a execução de cálculos, fórmulas e o cruzamento entre diferentes fontes de informação.

3.3.1 Primeiras construções de um cenário em transformação

O uso de dados é tendência, porém não chega a ser novidade no jornalismo (ROGERS, 2013a). Muito antes da difusão do termo jornalismo guiado por dados, nos anos 2000, já era possível identificar tentativas de criar reportagens baseadas em números e estatística, porém as iniciativas eram isoladas e não configuravam, ainda, a transformação de um cenário.

Hoje existem algumas nomenclaturas para designar o que Coddington (2014) chama de formas quantitativas: além do jornalismo guiado por dados, fala-se de jornalismo de precisão, Reportagem Assistida por Computador (RAC) e jornalismo computacional. Suas especificidades serão exploradas a seguir, em ordem cronológica. Com exceção do jornalismo de precisão, as demais formas pressupõem o uso de computador e todas estão inseridas em um espaço temporal mais ou menos demarcado, mas com possibilidade de sobreposição de conceitos e períodos.

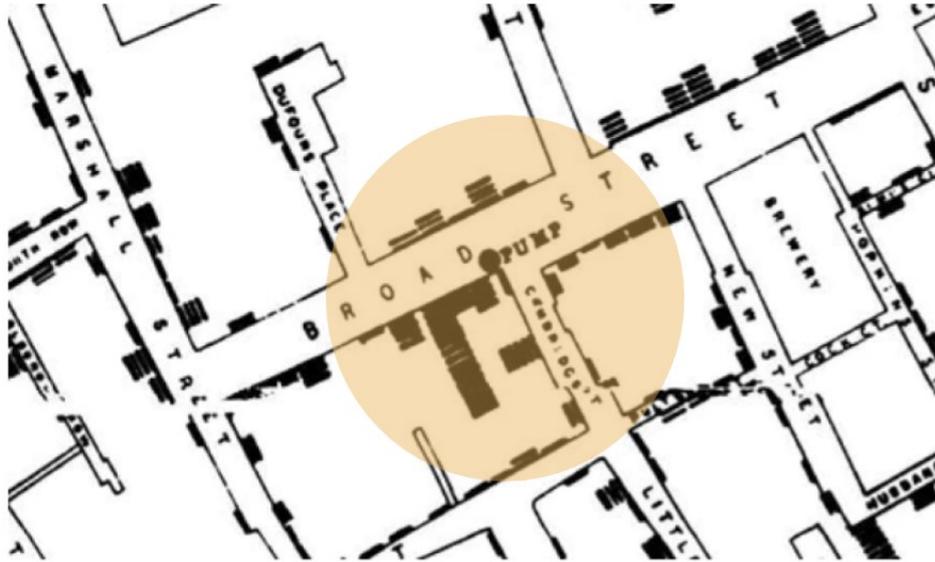
Rogers (2013a) considera uma publicação do jornal britânico *The Guardian*, no Século XIX, o marco inicial do que hoje se conhece como jornalismo guiado por dados, prática que pressupõe o uso de dados estruturados (BRADSHAW, 2017) e conduz toda a narrativa jornalística. No dia 5 de maio de 1821, o jornal, à época chamado de *Manchester Guardian*, publicou um texto acompanhado de uma lista de quantos alunos eram beneficiados pelo ensino

público em escolas de Manchester e Salford. Ir à escola não era obrigatório naquele período, pois as crianças usualmente trabalhavam durante a semana e estudavam aos domingos.

As estatísticas oficiais eram fruto da coleta de dados executada apenas por quatro homens ligados ao clero, o que dificultava a obtenção de resultados precisos. Acreditava-se que cerca de 8 mil crianças eram beneficiadas pelo ensino gratuito nessas duas cidades. A listagem publicada pelo jornal, vazada por uma fonte identificada apenas como *NH*, mostrou que havia distorção: na verdade o número de alunos integrados ao ensino público se aproximava de 25 mil. A publicação do *The Guardian* mostrou, pela primeira vez, uma fotografia verdadeira do cenário educacional daquelas cidades. A justificativa da fonte que vazou as informações era de que a publicação de números mais próximos da realidade permitiria a preparação de políticas para se pensar no futuro da sociedade, incluindo um bom sistema educacional.

O caso das escolas inglesas foi divulgado por meio de texto e de uma tabela com os números, o que auxiliou no processo de visualização. A visualização de dados, aliás, não é obra apenas de *softwares* sofisticados. Pode ser a última etapa do jornalismo guiado por dados, apresentando os resultados obtidos, ou a primeira, quando a busca é pela identificação de padrões e o estabelecimento de conexões. Em 1854, o médico John Snow mapeou os casos de cólera no bairro Soho, em Londres. À época, acreditava-se que a cólera se espalhava pelo ar – o que Snow demonstrou ser uma percepção equivocada. Mapeando os casos (Figura 1), descobriu que as ocorrências de contaminação ocorriam próximas de bombas d'água, o que remeteu à verdadeira origem do problema: uma fralda jogada dentro de um poço, provocando a contaminação por meio do consumo de água (ROGERS, 2013b). A identificação do problema por meio da visualização de dados, além de certa, continuou inspirando trabalhos no jornalismo e na medicina, pois até hoje o caso de John Snow é utilizado como exemplo.

Figura 1 – Mapa da cólera



Fonte: Reprodução The Guardian Data Blog (ROGERS, 2013b) | Captura de tela em 2 dez. 2017

É possível localizar, na área destacada em colorido no mapa, a posição da bomba de água (identificada como *pump*) e o número de mortes no entorno, exibido a partir de pequenos blocos empilhados.

Outra situação envolvendo saúde pública ocorreu também no século XIX, quando a enfermeira Florence Nightingale criou gráficos para representar visualmente as taxas de mortalidade do exército britânico. Seu principal trabalho é de 1858 e foi publicado em um relatório. Além de popularizar o chamado gráfico de pizzas – que mostra as partes de um todo –, originado em 1801, a enfermeira ficou conhecida por desenhar uma visualização semelhante à crista de galo, em espirais. Os gráficos publicados no relatório evidenciaram que a maior parte das mortes de soldados britânicos não ocorria por conta de ferimentos ocasionados por balas de arma de fogo, mas sim em decorrência de doenças que podem ser prevenidas (ROGERS, 2013a).

Os exemplos até aqui mencionados revelam iniciativas pontuais, mas que demonstram o potencial analítico e de descoberta que representa o emprego de dados em notícias. No jornalismo, o uso desse tipo de procedimento passou a ser mais recorrente no Século XX, especialmente após a chegada dos computadores às redações e ao uso de internet comercial. São as primeiras práticas que viriam a dar sustentação ao que se conhece hoje como jornalismo guiado por dados.

3.4 Jornalismo de precisão e Reportagem Assistida por Computador (RAC)

O início dos anos de 1950, nos Estados Unidos, é o ponto inicial do desenvolvimento da carreira do jornalista Philip Meyer, importante profissional para os estudos de jornalismo guiado por dados. Graduado pela Universidade de Kansas, é responsável pela criação do conceito de **jornalismo de precisão**, presente nas origens do jornalismo guiado por dados.

Seu percurso teve início em 1951, no jornal *The Topeka Daily Capital*¹⁷. Na década de 1960, participou da seleção para o programa *Nieman Fellowships*, da Universidade de Harvard, e foi aprovado em 1966. Durante nove meses, estudou métodos quantitativos nas ciências sociais, foi introduzido à linguagem simplificada de computador, o *data-text*, e aprendeu a estabelecer correlações estatísticas (MEYER, 2012). Depois do período em Harvard, em 1967, Meyer trabalhava no jornal *Detroit Free Press* quando começaram a ocorrer protestos¹⁸ de parte da população negra na cidade. O editor delegou a ele e colegas a cobertura dos eventos. Meyer sugeriu, então, que se fizesse uma pesquisa científica¹⁹ sobre as ações e reivindicações dos manifestantes utilizando ferramentas para pesquisa de opinião pública. O trabalho foi executado a partir de uma amostra de 437 pessoas. Identificou-se, também, que os protestos aconteciam em duas áreas não contínuas do mapa de Detroit.

O período de estudos em Harvard e a cobertura dos protestos de Detroit foram aspectos fundamentais para as bases conceituais do jornalismo de precisão, desenvolvido por Meyer a partir da segunda metade dos anos de 1960 e mais conhecido a partir de 1973, quando foi publicada a primeira edição do livro de mesmo nome. Consiste, basicamente, em aplicar métodos de ciências sociais à reportagem a partir de hipótese. Meyer (2002) defende que o jornalismo se aproxime da ciência à medida que opera com uma verdade verificável. Em sua concepção, o próprio profissional deve entrar em ação e fazer pesquisas. O jornalismo de precisão, acredita o autor, amplia o trabalho do repórter sem mudar sua natureza, que é encontrar fatos, compreendê-los e explicá-los.

¹⁷ Topeka é a capital do estado do Kansas e, no início dos anos de 1950, tinha 79 mil habitantes. A tiragem do *The Topeka Daily Capital* era de 60 mil exemplares pagos.

¹⁸ Os episódios tiveram início em 24 de julho de 1967 e ficaram conhecidos como *Detroit riots*, em inglês.

¹⁹ Havia três hipóteses para justificar os protestos: a) os manifestantes estavam frustrados com a economia e não vislumbravam avanços nesse sentido; b) o grupo de pessoas que protestaram era formado por negros que haviam migrado do Sul para o Norte, sendo que o Sul historicamente era uma área de repressão; c) privação relativa entre os manifestantes, pois a ascensão de negros não era uniforme – nem todos tinham as mesmas chances de crescimento em Detroit. A terceira hipótese foi aceita e utilizada para explicar os manifestos. A reportagem foi publicada e, no ano seguinte, de 1968, a equipe do *Detroit Free Press* conquistou prêmio em uma das categorias do Pulitzer por conta do detalhamento da investigação.

Diferente de outras práticas, o uso de computador é dispensável – funciona como um facilitador, mas não é pré-requisito fundamental. Sua contribuição com a criação do jornalismo de precisão, diz Meyer, tem relação com o valor da aplicação de método científico para análise e localização de fatos (GEHRKE; MIELNICZUK, 2017). Para Dader (2002), o jornalismo de precisão parte de uma autocrítica sobre o método utilizado no jornalismo para detectar, selecionar, classificar e hierarquizar os fatos. A prática está atrelada à medição quantitativa e qualitativa da realidade social. Além disso, o autor aposta no rigor técnico e na validade das descobertas do jornalismo de precisão a partir de comprovação meticulosa do método, de modo que o jornalismo de precisão ocorra de fato e em sentido estrito, não apenas como uma transposição de números. Dessa forma, não se trata apenas de uma especialidade temática, mas de uma estratégia de investigação e que implica verificação científica em qualquer situação. Além disso, permite que o jornalista não se prenda às declarações das fontes, que buscam filtrar suas versões dos fatos.

O protagonismo da máquina nos processos jornalísticos fica mais claro com a difusão do termo **Reportagem Assistida por Computador**²⁰, que Coddington (2014) considera continuidade do jornalismo de precisão por conta da proximidade de alguns processos e suas raízes nos métodos de ciências sociais. Nesse aspecto, o autor concorda com Dader (2002), para quem o jornalismo de precisão atingiu seu ponto mais alto, nos Estados Unidos, entre 1987 e 1990, a partir de uma série de prêmios *Pulitzer* vencidos por jornalistas que empregaram técnicas do jornalismo de precisão associado à informática.

Entende-se, contudo, que a RAC também se apropria de técnicas mais simples – uma busca avançada no *Google*, por exemplo, pode ser considerada um procedimento próprio, assim como troca de e-mails e outras pesquisas na internet, procedimentos hoje recorrentes. Em resumo, Stavelin (2013) aponta como ferramentas da RAC o uso de planilhas, a gestão de bancos de dados e recursos disponíveis na internet. Na mesma linha, Crucianelli (2013) entende que técnicas da RAC são úteis, na atualidade, para manusear volumes menores de informação, pois é considerado qualquer processo que utiliza computador em seus procedimentos.

O surgimento do termo RAC está relacionado à cobertura eleitoral da emissora de televisão estadunidense CBS, em 1952, que tentou prever o resultado da eleição presidencial na disputa entre Dwight Eisenhower e Adlai Stevenson. Por meio de resultados parciais, programadores prepararam fórmulas e identificaram que Eisenhower venceria, o que acabou se confirmando. No entanto, com receio de que houvesse um erro de código ou que o computador

²⁰ *Computer-Assisted Reporting (CAR)*, em inglês.

forneceria um resultado inadequado, a emissora de televisão não antecipou o resultado. A situação vai ao encontro do que afirma Briggs (2007) sobre a resistência dos jornalistas na incorporação de novos métodos às suas práticas. O autor relata que os profissionais resistiram inclusive ao uso do telefone para entrevistar fontes e, já nos anos 1990, ao utilizar e-mail como forma de comunicação. De qualquer maneira, o caso da CBS serviu para que, dali em diante, o uso de computadores passasse a ocorrer com frequência para a previsão de resultados eleitorais (COX, 2000).

Ainda que o resultado das eleições tenha sido o previsto pelo computador, o episódio de 1952 foi um fato isolado. Os computadores só começariam a se popularizar a partir dos anos de 1980, inicialmente nos Estados Unidos – neste período, sim, o termo RAC passou a ser mais conhecido e incorporado às práticas de apuração jornalística. Uma década antes, contudo, repórteres já buscavam trabalhar com bases de dados no dia a dia (COHEN; HAMILTON; TURNER, 2011).

Mesmo que em desuso, essa sigla²¹ ainda faz sentido em alguns países onde nomes de programas ou instituições não foram modificados por conta das alterações nas práticas com números. Nos Estados Unidos existe, por exemplo, o programa *National Institute for Computer-Assisted Reporting* (NICAR)²², fundado em 1989 juntamente com a Escola de Jornalismo da Universidade de Missouri e que ainda realiza conferências anuais ligadas a *Investigative Reporters and Editors* (IRE), referência do jornalismo investigativo nos Estados Unidos. Dader (2002) avalia que a criação de centros especializados como esse foram fundamentais para a popularização do acesso online a documentos e também para a consolidação da formação em estatística e informática nas faculdades de jornalismo e no mercado de trabalho.

Cox (2000, p. 20) ressalta, sobretudo, que a propagação da RAC só ocorreu porque demandava habilidades já existentes entre os jornalistas, que desta forma puderam explorar a tecnologia. “O desenvolvimento da Reportagem Assistida por Computador foi um evento conduzido pelo repórter. Por mais que a tecnologia ajude, só é boa se tiver um repórter tão bom quanto por detrás dela”²³. Para Lage (2009), o principal mérito da RAC é permitir a combinação

²¹ Meyer (1999) publicou um artigo reivindicando a aposentadoria do termo RAC por acreditar que já não faz sentido manter a palavra computador na sigla, considerando que vários processos são informatizados.

²² No final da década de 1980, Elliot Jaspin, do Providence Journal, mostrou que era possível relacionar a tecnologia das bases de dados e a linguagem estruturada Query em referências cruzadas para encontrar notícias. Ele viajou os Estados Unidos para ensinar tais ferramentas e fundou uma organização que depois se chamaria NICAR (COHEN, HAMILTON e TURNER, 2011).

²³ No original “The development of computer-assisted reporting has been a reporter-driven event. As helpful as technology is, it is only as good as the reporter who is behind it”.

do uso da internet com métodos de pesquisa, o que se repete, em alguma medida e níveis diferentes de complexidade, no jornalismo guiado por dados.

3.5 Jornalismo guiado por dados e jornalismo computacional

No JGD, os dados são a fonte principal do conteúdo jornalístico. Segundo Bradshaw (2017), há algumas possibilidades de fontes a serem monitoradas por jornalistas de dados: órgãos de governo, organizações, instituições acadêmicas e científicas, grupos de pressão e a própria mídia. Outra maneira é construir a própria base de dados, coletando informações a partir de métodos de pesquisa e observação. Independente da fonte, o autor indica que tudo começa com uma pergunta, que pode ser feita a partir do olhar crítico para as bases de dados, boletins ou releases, ou a partir de uma percepção de curiosidade sobre o mundo e cujas questões podem ser contextualizadas nas notícias.

As nomenclaturas até aqui trabalhadas se sobrepõem em alguns aspectos e se diferenciam em outros. A partir da metade dos anos 2000, o jornalismo guiado por dados passou a ganhar relevância nas redações e nas pesquisas. Um dos motivos foi o texto escrito em 2006 por Adrian Holovaty, jornalista que trabalhou com sites de notícias e descreveu, de forma simples, o que se entende por jornalismo de dados. Sua principal crítica está no fato de que os jornais não haviam mudado a operacionalização da construção de notícias, centralizada em coletar informações e escrever uma história. O autor acredita na necessidade de rever essa conduta, pensando em formas de agregar informações e organizá-las de modo que possam ser lidas por máquinas.

Holovaty defende que os jornalistas apostem na criação de dados estruturados para posterior recuperação no conteúdo jornalístico. Exemplo citado pelo autor diz respeito a incêndios. O que o jornalismo faz, na maioria das vezes, é noticiar o fato em si, sem explorar outras possibilidades. Com as informações organizadas, seria possível criar uma reportagem abordando incêndios em determinado período, sendo possível detalhar dia, horário, causa da ocorrência, tempo de demora para o atendimento, entre outros fatores. Semelhante ao incêndio estão casamentos e obituários, que possuem informações coletáveis e estruturáveis. O fato é que os dados são o ponto inicial para esse tipo de prática²⁴.

²⁴ Mancini e Vasconcellos (2016) diferem as práticas do jornalismo guiado por dados em relação ao que identificam como jornalismo com dados. Enquanto no JGD os dados são a razão de ser do trabalho de reportagem, tal como um núcleo de desenvolvimento e conduzido por todo o texto, o jornalismo com dados, menos profundo, significa apenas o uso de informações quantitativas sem desenvolvê-las ou problematizá-las e promover cruzamentos.

Autores como Machado (2006) e Barbosa (2007) definem as bases de dados como aspecto essencial no jornalismo digital. Para Machado (2006), as bases de dados estruturam e organizam todas as etapas do processo de produção jornalística – atuam desde o processo de apuração e elaboração do conteúdo até a sua edição e circulação. Em resumo, as funções das bases de dados são, basicamente: “[...] 1) de formato para a estruturação da informação; 2) de suporte para modelos de narrativa multimídia e 3) de memória dos conteúdos publicados” (MACHADO, 2006, p. 16). Nesse sentido, as bases de dados funcionam como suporte para contar histórias no jornalismo, incluindo o processo de observação e exploração das informações até a contribuição para um ambiente interativo de navegação. Além disso, estão atreladas aos sistemas automatizados de armazenamento e recuperação de dados, de modo que as informações possam ser reaproveitadas.

No modelo teórico Jornalismo Digital em Bases de Dados (JDBD), cuja proposição é justificada pela complexidade dos processos que envolvem os produtos jornalísticos digitais, as bases de dados são

“[...] definidoras da estrutura e da organização, bem como da apresentação dos conteúdos de natureza jornalística, de acordo com funcionalidades e categorias específicas, que vão permitir a criação, a manutenção, a atualização, a disponibilização e a circulação de produtos jornalísticos digitais dinâmicos” (BARBOSA, 2007, p. 218).

As bases de dados, afirma a autora, configuram um elemento estruturante e com potencial de reinvenção. São identificadas rupturas qualitativas para os produtos jornalísticos digitais, gerando inovação. Dessa forma, suas funções são desempenhadas tanto em nível de gestão interna dos produtos quanto em termos de apuração e contextualização, além da estruturação e recuperação das informações, bem como a composição do conteúdo. A configuração do paradigma JDBD ocorre em um cenário com características como base tecnológica ampliada, uso de conexão de banda larga, expansão das bases de dados, algoritmos, diferentes linguagens de programação, produtos automatizados e emprego de conhecimentos específicos para extração de informações. As bases de dados, conforme Barbosa (2007), não desempenham apenas uma função documental – são uma nova forma cultural ligada à produção em que o computador é utilizado para criação e disseminação de conteúdo jornalístico.

A Web, no JDBD, é tratada como uma grande base de dados. Dessa forma, o jornalismo guiado por dados é abrangido conceitualmente pelo JDBD e considerado uma de suas vertentes ou extensões porque amplia a possibilidade de uso das bases de dados desde a esfera de produção do conteúdo jornalístico até a circulação (BARBOSA; TORRES, 2013).

Bounegru (2012) diz que o jornalismo guiado por dados é resposta de uma tradição longa, que responde a novas circunstâncias, como o crescimento no volume de dados oferecidos, assim como ferramentas de publicação. A entrada dos microcomputadores e do acesso à internet no final do Século XX geraram um aumento na oferta de bases de dados públicos e *softwares* de análise (TRÄSEL, 2014), estimulando a adoção dessas técnicas entre os jornalistas. Bradshaw (2017) concorda: atribui a ascensão do jornalismo guiado por dados ao aumento do poder da computação, da capacidade de armazenamento de dados e do acesso às bases como fontes de notícias para os jornalistas.

A partir de uma análise a 48 sites ligados a produtos e formatos trabalhados no JGD, Barbosa e Torres (2013) perceberam possibilidades de expansão do modelo teórico JDBD por meio do JGD no que se refere aos modos de narração, criação e apresentação de conteúdo jornalístico. Identificaram uma tipologia de quatro categorias para distinguir as iniciativas. São elas: *Mashups*, que abrange a combinação de fontes variadas e apresentadas em formatos, por exemplo, de infográficos interativos e *newsgames*; Data/Dados, como termo utilizado para denominar equipes e seções de veículos que trabalham com dados; Sistemas de Recomendação de Notícias (SRN), com listas das mais lidas e comentadas, por exemplo; e Fontes, que dizem respeito às informações divulgadas pelo poder público e também pelas empresas jornalísticas.

No caso da categoria Fontes, que interessa a este estudo, identificam sites governamentais que disponibilizam dados públicos, como os portais da transparência, e também de instituições jornalísticas que desenvolvem projetos e oferecem os dados estruturados em plataformas abertas para que as informações sejam propagadas e utilizadas por quem tiver interesse. Para Barbosa e Torres (2013), a tendência dos governos em aumentar a transparência a partir da disponibilidade de dados públicos combinada aos ideais de iniciativas não governamentais em oferecer informações *open data* favorecem a expansão do jornalismo guiado por dados.

A transparência pública ativa, por meio de repositórios de governos, é realidade há alguns anos. Em 2009, nos Estados Unidos, o governo de Barack Obama lançou um portal²⁵ com dados do governo federal, que reúne planilhas de diferentes setores, como agricultura, saúde, violência e comércio. Austrália, Nova Zelândia, Reino Unido e outras nações seguiram o exemplo e também providenciaram seus sites (ROGERS, 2013a). No Brasil, apesar da existência anterior de leis de transparência e da própria Constituição Federal de 1988²⁶, o acesso

²⁵ <http://www.data.gov/>

²⁶ No inciso XXXIII do artigo 5º da Constituição Federal consta que “todos têm direito a receber dos órgãos públicos informações de seu interesse particular, ou de interesse coletivo ou geral, que serão prestadas no prazo

público passou a ser obrigatório com a Lei Nº 12.527, conhecida como Lei de Acesso à Informação (LAI), em vigor desde maio de 2012, que prevê a divulgação de informações de interesse público dos poderes Legislativo, Executivo e Judiciário em qualquer esfera. Quando os dados procurados são de interesse público, mas não estão disponíveis, podem ser obtidos via Serviço de Informação ao Cidadão (SIC). O Decreto 8.777, assinado pela presidente Dilma Rousseff quatro anos mais tarde, indica que os arquivos do governo federal têm de ser disponibilizados em formato aberto.

Para que cumpra com os requisitos da acessibilidade, as informações precisam ser compartilhadas em formatos reutilizáveis. Também é necessário que se apresente levantamentos completos, que permitam análise dos dados. Souza (2016, p. 60) afirma que, a partir da LAI, a divulgação de informações públicas é regra e o sigilo a exceção. “O acesso a documentos, arquivos e estatísticas, constitui-se em um dos fundamentos para a consolidação da democracia que possibilita a transparência dos atos públicos e busca pelas informações constituindo-se em uma forma de participação do povo”.

Não apenas a disponibilidade de dados oficiais em bases de dados públicas modificou o cenário e a necessidade de os jornalistas trabalharem com dados. Outro ponto levantado por Bradshaw (2017) está na necessidade de conhecimentos do jornalismo guiado por dados com relação ao que fazer com documentos vazados, tal como o *WikiLeaks*, em 2010, inicialmente com documentos sobre a guerra no Afeganistão e no Iraque. De acordo com Rogers (2012), que à época era jornalista de dados do *The Guardian*, os documentos de guerra foram compartilhados também com o *The New York Times* e o *Der Spiegel*. Havia 92.201 linhas de dados na planilha, cada uma com detalhes sobre eventos militares no Afeganistão. A quantidade massiva de documentos vazados fez com que os jornalistas passassem a enxergar a necessidade de trabalhar com bases de dados extensas, contribuindo, dessa forma, para o desenvolvimento do jornalismo guiado por dados como tendência (BAACK, 2011)²⁷.

Em 2016 foi a vez do *Panama Papers*, que permitiu o trabalho de investigação em rede entre redações jornalísticas ao redor do mundo. Nesse episódio, mais de 11 milhões de documentos foram vazados do Mossack Fonseca, um escritório de advocacia no Panamá. A redação do alemão *Süddeutsche Zeitung* obteve os registros a partir de uma fonte anônima e dividiu os documentos com outros veículos associados ao Consórcio Internacional de

da lei, sob pena de responsabilidade, ressalvadas aquelas cujo sigilo seja imprescindível à segurança da sociedade e do Estado”. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm

²⁷ Também é nesse contexto e em casos como do próprio *WikiLeaks* que surgem iniciativas como o *crowdsourcing* (HOWE, 2006), espécie de apuração coletiva delegada a humanos para tarefas cognitivas que o computador não é capaz de cumprir.

Jornalistas Investigativos (ICIJ, na sigla em inglês). Os registros foram analisados por mais de 370 jornalistas de 76 países e o resultado da investigação foi publicado em 2016, mostrando o uso de companhias *offshore* para facilitar o pagamento de propinas, negócios envolvendo armas, evasão de divisas e fraudes (ICIJ, 2016). As fraudes envolveram políticos, bilionários e celebridades.

O tratamento de informações obtidas por meio de documentos vazados tem relação com o conceito de Träsel (2014, p. 119), para quem o jornalismo guiado por dados se refere “a aplicação da computação e dos saberes das ciências sociais na coleta, processamento, interpretação e apresentação de dados, com o objetivo de ampliar a função da imprensa como defensora do interesse público” e como uma das respostas da comunidade profissional jornalística à crise econômica e de identidade pela qual passa o jornalismo. Na perspectiva de Crucianelli (2013)²⁸, trata-se de da combinação de técnicas do jornalismo investigativo, do jornalismo em profundidade, do jornalismo de precisão, da Reportagem Assistida por Computador e do jornalismo analítico, esse último ligado à transposição de dados para sistemas.

O modelo de pirâmide invertida voltada ao jornalismo guiado por dados é apresentado por Bradshaw (2011) com as seguintes etapas: a) *compilação dos dados*, que deriva da pergunta que o jornalista de dados pretende responder a partir de bases estruturadas; b) *limpeza*, que envolve a remoção de erro humano nas planilhas e a padronização de números com os quais se trabalha; c) *contextualização*, que compreende a clareza sobre onde os dados foram extraídos, quem coletou e com que propósito; d) *combinação*, etapa que prevê o cruzamento de bases de dados diferentes e complementares; e, por fim, a e) *comunicação*, que gera uma segunda série de procedimentos por conta do detalhamento previsto pelo autor: *visualização*, *narração*, *socialização*, *humanização*, *personalização* e *utilização*, todas ligadas ao modo como o leitor vai se conectar com o conteúdo produzido.

Para fins de aplicação nesta dissertação, trabalha-se com um conceito operacional de jornalismo guiado por dados, baseado principalmente em Bradshaw (2017). Assim, serão considerados, no estudo empírico de fontes do próximo capítulo, conteúdos cujas marcas do produto dão pistas de que foram compostos a partir de **técnicas do JGD, o que inclui consultas a fontes como base de dados, criação dos próprios repositórios de dados e outros documentos, conectado a um jornalismo de pesquisa e que por vezes culmina com a visualização. O produto pode ser só texto, mas também pode ter/ser uma visualização. Embora esteja diretamente associado ao uso de planilhas e/ou estatística no processo de**

²⁸ A autora utiliza o termo jornalismo de dados e jornalismo de banco de dados como sinônimos, afirmando que o primeiro prevalece por ser mais curto.

apuração, o JGD não precisa, necessariamente, apresentar números na forma final do produto.

As práticas do JGD também estão ligadas aos saberes de outras áreas. Segundo Heravi (2017), é uma disciplina emergente que congrega saberes de ciências sociais, ciência da informação, ciência da computação, análise de dados, design e *storytelling*. A partir de uma pesquisa online com 206 participantes de 43 países, intitulada *Global Data Journalism Survey*, a autora constatou que houve crescimento no JGD nos últimos anos e que a comunidade que se dedica a essa prática apresenta altos níveis de educação com formação principalmente em jornalismo e comunicação e em menor escala em disciplinas ligadas à informática. Pesquisa de Fink e Anderson (2014) igualmente se desenvolve nessa direção. De acordo com os autores, jornalistas de dados não necessariamente começam a carreira como repórteres, mas por vezes passam, antes, por mestrado e doutorado em outras áreas como ciência política e ciência da informação. Ainda que o uso de fontes documentais não seja novo no jornalismo, passou a ter seu valor redescoberto por conta da necessidade de novos caminhos para o jornalismo (MIELNICZUK; TRÄSEL, 2017).

No que toca ao uso de linguagem de programação e a proximidade com a área da informática, há o termo **jornalismo computacional**, que prevê o uso de automação para que os jornalistas consigam trabalhar com grandes bases de dados sem precisar classificá-los manualmente. Pode ser considerado um híbrido entre a ciência da computação e o jornalismo, operando entre os dois campos (STAVELIN, 2013) e oferecendo ferramentas tecnológicas para uso na prática. Essa afirmação vai ao encontro da definição de Cohen, Hamilton e Turner (2011), que destacam a integração entre jornalistas e programadores para o uso de algoritmos, o desenvolvimento de novas técnicas, métodos e interfaces para publicação de histórias de interesse público. Nas universidades, Stavelin diz que o termo vem sendo ligado ao ensino de práticas como programação, mineração de dados, jornalismo estruturado, visualização, análise de rede, jornalismo investigativo e práticas similares.

Na visão do autor, o jornalismo computacional se difere de outras formas desenvolvidas a partir do computador pelos seguintes motivos: a) o centro está na plataforma (no sentido de espaço para publicação) e não na história a ser contada, pois dá ênfase não a facilitar uma narrativa, mas sim na possibilidade de que se possa explorar essas histórias sob aspectos diferentes; b) modelos computacionais podem ser adicionados – lembra a personalização de dados, fazendo com que as pessoas possam trabalhar com seus dados a partir de grandes bases; c) o pensamento computacional pode ser aplicado para solucionar questões importantes para a sociedade e os usuários.

Das nomenclaturas apresentadas até aqui, Stavelin (2013) considera todos os termos dentro de um escopo maior, chamado de produção guiada por *software* ou produção guiada por computador, demandando saber especializado ou conhecimento aprimorado da técnica em comparação a outras formas de jornalismo. Por conta de suas especificidades, profissionais especializados na área possuem nomenclatura específica: são os jornalistas de dados (*data journalists*).

Em essência, o jornalismo de precisão enfatiza o uso de método científico, RAC as ferramentas digitais, jornalismo em bases de dados enfatiza o armazenamento e a recuperação de informações, jornalismo de dados ou guiado por dados busca encontrar histórias em bases de dados, enquanto jornalismo computacional destaca a fusão de valores jornalísticos e da computação para a criação de ferramentas e aplicação de métodos (STAVELIN, 2013, p. 46).²⁹

Em todas essas formas, o autor enxerga fases de entrada e saída (*input e output*), afirmando que, na fase de entrada, todas essas formas são tratadas sob a perspectiva de gerar um produto jornalístico (fase de saída), como notícias produzidas para jornal, rádio, televisão e internet, ambiente em que há interatividade. A diferença em relação ao jornalismo computacional reside principalmente na fase de saída, em que Stavelin aponta como escopo dessa prática a criação de *softwares* para as notícias ou para as redações.

A compreensão de Crucianelli (2013) sobre o JGD igualmente se aproxima do jornalismo computacional. Além de ter como premissas o emprego de grandes volumes de dados e visualização interativa, a autora vê como característica desse conjunto de práticas a incorporação do programador à equipe de jornalismo, à medida que passa a ser necessária a presença de alguém que saiba extrair dados e desenhar suas aplicações. Coddington (2014) afirma que, assim como o JGD, o jornalismo computacional também é considerado descendente da RAC. Para ele, o jornalismo computacional está mais preocupado com a coleta, a busca de sentido e a apresentação de informações por meio de técnicas de computação do que uma preocupação jornalística propriamente dita.

Dos conceitos examinados até aqui é possível resumir o seguinte:

²⁹ Do original “In essence, precision journalism emphasizes the use of scientific method, CAR emphasizes digital tool use, database journalism emphasizes structure of information storage and retrieval, data and data-driven journalism emphasizes finding stories in data sets, while computational journalism emphasizes the merging of computing and journalistic values in tool creation and method application.”

Quadro 2 – Produção jornalística orientada por *software*

| Nome | Período inicial | Prática |
|---|------------------------|---|
| Jornalismo de precisão | Década de 1960 | Aplicação de métodos de pesquisa social e estatística na prática da reportagem |
| Reportagem Assistida por Computador (RAC) | Década de 1980 | Uso de computador na apuração jornalística e, a partir da metade dos anos de 1990, consultas feitas pela internet e entrevista via e-mail |
| Jornalismo computacional | A partir dos anos 2000 | Automação de processos para coleta e organização de informações com forte vínculo de programação |
| Jornalismo guiado por dados | A partir dos anos 2000 | Busca de histórias em produção jornalística orientada por dados desde a coleta até a visualização |

Fonte: A autora (2018)

Em comum, todas as práticas possuem a orientação via *software* e o uso do computador, com exceção do jornalismo de precisão, em que essa característica não é fundamental. Ainda que todos os termos tenham suas especificações, em geral demandam um conhecimento de bases de dados por parte do jornalista. No caso do jornalismo computacional, precisa estar relacionado à ciência da computação ou a uma equipe multidisciplinar, capaz de dar conta dos processos.

O trabalho de um arquivista, descrito por Beltrão (1969), não é tão diferente do que faz um jornalista hoje, na era *big data*. Apesar do abismo temporal, o ponto principal diz respeito à organização das informações de modo que sua recuperação seja rápida³⁰ quando o repórter precisar acionar algum material de arquivo. A necessidade de busca ágil e recorrente, atualmente, faz com que a disponibilização de informações em bancos de dados estruturados torne a prática mais acessível. A migração para a plataforma digital, contudo, também gera mais exigências quanto à precisão na busca de informações (BARBOSA, 2007).

Este capítulo discutiu o contexto de desenvolvimento do jornalismo guiado por dados, a partir do fenômeno *big data*, que surge em um espaço de sociedade em rede e de jornalismo em rede por conta da facilidade de acesso à tecnologia e da difusão de bancos de dados públicos. Os precursores do JGD foram estabelecidos ainda na etapa industrial de organização da

³⁰ Beltrão defendia que o responsável pelo arquivo do veículo fosse um jornalista especializado com quatro funções principais: recolher e catalogar fotos e outros materiais publicados pelo jornal; dirigir e organizar a coleção de recortes; receber e catalogar materiais externos, que são utilizados pelos jornalistas como “base de informações” (BELTRÃO, 1969, p. 66) e redigir notas biográficas, sínteses histórias, estatística, fichas de referência e outros.

sociedade, época em que o uso de computador na apuração jornalística ainda era incipiente e pontual, e o conjunto de práticas se acentuou em uma época de sociedade em rede.

O uso de dados como fonte sofreu transformações à medida que passou a integrar um cenário de sociedade e jornalismo em rede, a partir de políticas de transparência e da disponibilidade de bases de dados públicas. Além disso, o contexto de geração de dados a partir do uso de sites de redes sociais, por exemplo, também se revelou matéria-prima para o JGD. Soma-se a isso a possibilidade ampla de consulta a estudos científicos, artigos e outras fontes complementares, mais facilmente explorada com o auxílio da tecnologia em comparação à sociedade industrial.

A partir da apresentação de estudo empírico, o próximo capítulo identifica fontes acionadas em 60 notícias elaboradas a partir de técnicas de jornalismo guiado por dados. A análise envolve produções dos jornais *O Globo* e *Nexo*, do Brasil, *The New York Times* e *FiveThirtyEight*, dos Estados Unidos, e *La Nación* e *Chequeado*, da Argentina. Partindo de semelhanças encontradas na identificação de fontes propõe-se uma tipologia de classificação aplicável às fontes acionadas nas notícias de jornalismo guiado por dados.

4 ESTUDO DE FONTES

O percurso teórico até aqui desenvolvido sustenta as bases para uma leitura crítica do material empírico que será analisado neste capítulo. Entende-se que, a união de esforços teóricos e práticos, materializada pela combinação da revisão bibliográfica e de um estudo de identificação de fontes acionadas em notícias, permite gerar uma investigação mais completa, à medida que as duas frentes exploradas são complementares e geram contribuições mútuas.

A delimitação do *corpus* demandou uma série de escolhas: foi preciso eleger os veículos de comunicação, as notícias publicadas e os procedimentos metodológicos que resultaram na criação de categorias para o estudo das fontes acionadas em publicações que utilizaram técnicas de jornalismo guiado por dados – e cujas marcas são perceptíveis no produto final.

As etapas deste estudo empírico podem ser assim divididas: 1) seleção de veículos; 2) seleção das notícias a serem estudadas; 3) leitura e organização dos dados em fichas; 4) organização dos dados em planilhas; 5) refinamento das categorias de fontes e definição de tipos; 6) análise dos materiais reunidos. A seguir os procedimentos serão detalhados.

4.1 Os veículos: tradicionais e nativos no Brasil, Estados Unidos e Argentina

A primeira decisão foi quanto à escolha dos veículos de comunicação analisados: *O Globo* e o *Nexo Jornal*, do Brasil, *The New York Times* e *FiveThirtyEight*, dos Estados Unidos, e *La Nación* e *Chequeado*, da Argentina. A navegação exploratória realizada durante o período de estudos no mestrado, de março de 2016 a outubro de 2017, permitiu essa definição por conta do trabalho realizado pelos veículos a partir do uso de técnicas de jornalismo guiado por dados e pela regularidade de suas produções. Também foi levada em conta a importância do veículo para o seu país de origem e no cenário internacional.

A seleção de veículos de países diferentes tem como propósito a diversidade no *corpus*. Dessa forma, optou-se pela análise de notícias publicadas por veículos de comunicação tradicionais – *O Globo*, *The New York Times* e *La Nación* – consagrados em seus países e conhecidos como mídia de referência; e também veículos nativos digitais com menor tempo de existência e que executam trabalhos semelhantes com o uso de dados – *Nexo*, *FiveThirtyEight* e *Chequeado*.

Observando o contexto dos três países, todos apresentam momentos diferentes em relação às políticas de transparência e acesso à informação pública. De acordo com Souza (2016), as leis de acesso à informação no mundo se expandiram principalmente nos anos de

1990. O Brasil conta com instrumentos legais desde a Constituição de 1988 e mais especificamente com a aprovação da Lei de Acesso à Informação, em vigor desde 2012. Nos Estados Unidos, o caminho em direção ao acesso à informação pública é mais antigo. A Lei de Liberdade de Informação, chamada de *Freedom of Information Act* (FOIA) em inglês, data de 1966 (SOUZA, 2016), embora não seja aplicável a todos os departamentos do poder público. Ao final de 2009, essa lei ficou muito próxima da ideia de governo aberto com a disponibilidade de bancos de dados pelo governo Barack Obama. Na Argentina, por sua vez, existe o *Decreto de Acesso à Informação*, de dezembro de 2003, válido para o governo federal. Segundo a autora, tal decreto está inserido na Constituição daquele país, não havendo uma lei específica de acesso à informação. De acordo com Crucianelli (2013), não existe uma cultura de dados abertos no país, dificultando o acesso dos cidadãos às informações de interesse da sociedade.

A seguir, serão apresentados os veículos cujas notícias foram estudadas, começando pelos tradicionais e na sequência os nativos digitais, de acordo com seus países de origem. O objetivo não é fazer uma reconstituição histórica de cada um, mas trazer brevemente alguns elementos que justifiquem a escolha pelas seis iniciativas e sua importância no contexto do jornalismo guiado por dados.

4.1.1 O Globo e o blog *Na Base dos Dados*: cobertura diversa

O jornal *O Globo*³¹, sediado no Rio de Janeiro, foi lançado em 29 de julho de 1925, sob a fundação de Irineu Marinho, e circula em versão impressa até os dias de hoje. É um dos principais jornais brasileiros de circulação nacional, ao lado de nomes como *Folha de S. Paulo*³² e *O Estado de S. Paulo*³³. A estreia de *O Globo* na *Web* ocorreu em 29 de julho de 1996, em uma das comemorações de aniversário do veículo. *O Globo On*, como era chamado o site,

³¹ <http://www.oglobo.globo.com>

³² O jornal *Folha de S. Paulo* criou o blog *FolhaSPDados*, que aparentemente surgiu e parou de ser atualizado em 2012. A página era alimentada pelo jornalista Gustavo Faleiros em uma parceria da Folha com o programa *Knight International Journalism Fellowships*, do Centro Internacional para Jornalistas (ICFJ, na sigla inglês). A iniciativa trabalhava com dados disponibilizados por órgãos governamentais, universidades e institutos de pesquisa. A Folha ainda conta com o blog *Afinal de Contas*, assinado pelo jornalista Marcelo Soares (2016), atualizado pela última vez em outubro de 2016. Em 2017, a *Folha* contratou o cientista de dados Daniel Mariani, ex *Nexo*, e vem retomando as publicações com o emprego de dados.

³³ O primeiro jornal brasileiro da mídia de referência a formar uma equipe de profissionais para trabalhar com técnicas do jornalismo guiado por dados foi *O Estado de São Paulo*, que em 2012 lançou o *Estadão Dados*, inicialmente formado por dois jornalistas, um engenheiro de *software* e um *designer*. A ideia era unir estatística, programação e edição, gerando informação a partir de dados. O primeiro produto especial do grupo foi o *Basômetro*, ferramenta de monitoramento das votações de parlamentares brasileiros (GANDOUR, 2014). A equipe do *Estadão Dados* foi objeto de observação participante da tese de Marcelo Träsel (2014). Embora seja referência e caracterize um momento histórico, o blog do *Estadão Dados* diminuiu a regularidade de produção no período de desenvolvimento desta dissertação.

buscava, desde então, ser mais do que uma transposição do papel para o ambiente digital, apostando na agilidade das publicações e no aprimoramento da prestação de serviços (O GLOBO MEMÓRIA, 2017).

O blog *Na Base dos Dados*³⁴, espaço destinado aos profissionais do Núcleo de Jornalismo de Dados de *O Globo*, foi lançado em 22 de agosto de 2014. As produções da equipe são voltadas para veiculação na internet e também para a versão impressa. O lançamento do blog prometia, desde o início, análises a partir de bases de dados e visualizações por meio de infografia. Sua razão de ser está, portanto, na disponibilidade de dados públicos: justifica-se, na primeira postagem do blog, o fato de nunca ter havido tantas bases disponíveis desde a disseminação do conceito de *Open Government* (ALLEGRO, 2014).

A proposta da seção é abordar temas e assuntos relacionados ao jornalismo guiado por dados, disponibilizando repositórios, análises, gráficos e infográficos produzidos pela equipe (NA BASE DOS DADOS, 2017). Ainda que o jornal *O Globo* não tenha sido o pioneiro na formação de um time específico e especializado em jornalismo guiado por dados no País, é o que até outubro de 2017 realizou atualizações todas as semanas – na maioria dos casos, três vezes por semana. Nesse sentido, a escolha por *O Globo* se justifica pela frequência das publicações e pela variedade da área de cobertura, que inclui pautas sobre política, economia, polícia, cultura, educação e outras áreas.

4.1.2 Nexo Jornal: o contexto a partir dos dados

O *Nexo Jornal*³⁵ representa uma das iniciativas independentes listada pela Agência Pública em um mapa³⁶ brasileiro sobre o tema no ano de 2016, logo depois de ser criado, em novembro de 2015. É um jornal que nasceu digital, com foco em “explicações precisas e interpretações equilibradas” (NEXO, 2017, documento eletrônico) a partir de informações contextualizadas sobre eventos do Brasil e do mundo, incitando o debate público. O nome *Nexo* representa algo que busca fazer sentido, oferecendo informações claras e contextualizadas.

A equipe é composta por 29 profissionais³⁷ das áreas de jornalismo, arte, tecnologia, pesquisa, negócios e estratégia, divididos em seis núcleos que operam de forma integrada. A sede está localizada em São Paulo. Tem como objetivo a busca pelo equilíbrio, o uso de

³⁴ <http://blogs.oglobo.globo.com/na-base-dos-dados/>

³⁵ <https://www.nexojournal.com.br/>

³⁶ O mapa do jornalismo independente é uma iniciativa da Agência Pública e tem como objetivo reunir iniciativas que surgiram de projetos coletivos e não vinculados a grupos de mídia e outras organizações de influência (AGÊNCIA PÚBLICA, 2017). Está disponível no endereço <https://apublica.org/mapa-do-jornalismo/>

³⁷ De acordo com documento remetido pelo Nexo em julho de 2017.

evidências e dados e a clareza do conteúdo produzido. O *Nexo* é atualizado todos os dias e de maneira frequente, mas não tem o compromisso de operar no estilo *breaking news*, com atualizações minuto a minuto, justamente por ter uma proposta diferente e conectada à explicação (NEXO, 2017). Seus fundadores são Paula Miraglia, que atua como diretora geral, Renata Rizzi, diretora de estratégia e negócios, e Conrado Corsalette, editor-chefe.

Segundo seus criadores, o *Nexo* teve como ponto de partida a visão de que o jornalismo tem potencial transformador ao oferecer conteúdo claro e explicativo, fatores que contribuem para um debate qualificado. O jornal procura oferecer mapas e infográficos junto aos conteúdos, assim como ferramentas interativas, com o intuito de cativar não apenas o que consideram o leitor de mídia tradicional, mas também um público jovem e interessado pelo debate público. O veículo foi premiado em 2017 no *Online Journalism Awards*³⁸, que reúne iniciativas do mundo inteiro, com o prêmio de excelência em redações pequenas.

4.1.3 *The New York Times*: explicando as notícias na seção *The Upshot*

Sediado em Nova York, nos Estados Unidos, o *The New York Times*³⁹ foi fundado em 1835, à época chamado de *The New York Herald*, por James Gordon Bennett. A inclusão deste jornal no estudo se deve à importância a ele atribuída, pois é referência em qualidade de jornalismo e iniciativas inovadoras para veículos de comunicação não apenas nos Estados Unidos, mas em todo o mundo. Ao longo de sua história conquistou prêmios de jornalismo, incluindo o *Pulitzer*⁴⁰, o mais importante deles.

Após a formação em 1995 do *The New York Times Electronic Media Company*, visando expandir as publicações digitais, o site foi lançado em 22 de janeiro de 1996, disponibilizando aos leitores acesso ao conteúdo do jornal (THE NEW YORK TIMES, 2017). A ideia também era oferecer material não disponível no jornal impresso e atividades interativas, como as palavras-cruzadas (LEWIS, 1996).

A seção *The Upshot*⁴¹ foi criada em 22 de abril de 2014 no site do *The New York Times* e reúne conteúdo produzido a partir de dados públicos e pesquisas, com foco em polícia, política e economia (MCDULING, 2014) e um de seus destaques está na visualização dessas

³⁸ A premiação ocorre desde os anos 2000 e valoriza os produtos de excelência em jornalismo digital no mundo. Mais informações estão disponíveis no site: <https://awards.journalists.org/about/>

³⁹ <https://www.nytimes.com/>

⁴⁰ O primeiro *Pulitzer Prize* recebido pelo *The New York Times* foi em 6 de março de 1918, por conta da publicação de textos sobre dezenas de boletins oficiais, documentos e discursos sobre a Primeira Guerra Mundial. Entre os mais recentes, o *The New York Times* recebeu Pulitzer 2017 em três categorias: *International Reporting*, *Feature Writing* e *Breaking News Photography* (2017 PULITZER PRIZES, 2017).

⁴¹ <https://www.nytimes.com/section/upshot>

informações. A proposta é explicar as notícias, permitindo que os leitores façam suas próprias conexões entre diferentes histórias. Outro ponto é inserir dados entre as notícias divulgadas no dia a dia (LEONHARDT, 2014), de modo que auxilie no processo de contextualização.

4.1.4 *FiveThirtyEight* e a análise estatística

O site *FiveThirtyEight*⁴² é um projeto do economista e estatístico Nate Silver e começou a se desenhar com este nome em 2012 (BRADSHAW, 2017), embora Silver tenha trabalhado com atividades semelhantes antes mesmo desse período. O espaço reúne notícias desenvolvidas a partir de análise estatística e oferece conteúdo nas áreas de política, esportes, ciência e saúde, economia e cultura. Leva esse nome porque 538, ou *FiveThirtyEight*, é o número de eleitores que compõe o colégio eleitoral dos Estados Unidos, utilizado para eleições indiretas a presidente e vice.

A iniciativa tornou-se tão popular que foi comprada pelo jornal *The New York Times* de 2010 a 2013. Após esse período, a empresa *ESPN*, de esportes, estabeleceu uma parceria para que a equipe do *FiveThirtyEight* pudesse ampliar o leque de cobertura de áreas, incluindo esportes. Em 2016, a variedade de tópicos de cobertura de fato foi ampliada para áreas como ciência e cultura. Para Lewis e Waters (2017), ainda que não seja jornalista de formação, Silver e suas análises contribuíram para o desenvolvimento do jornalismo guiado por dados nos Estados Unidos. Silver ficou conhecido principalmente pelas previsões corretas sobre o vencedor das eleições presidenciais nos Estados Unidos em 2008 e 2012.

O site *FiveThirtyEight* reúne profissionais de jornalismo, tecnologia e análise de dados, incluindo técnicas de visualização e programação. Assim, busca manter os princípios tradicionais do jornalismo à medida que oferece informações de interesse público, porém com mais atenção para o uso de dados, a partir da coleta dessas informações e da construção de notícias a partir delas (CARLSON; USHER, 2015). Para os autores, o formato do *FiveThirtyEight* une tradição e inovação no jornalismo.

⁴² <http://fivethirtyeight.com/>

4.1.5 *La Nación* e *La Nación Data* para fiscalizar o governo

O jornal argentino *La Nación*⁴³ teve sua primeira edição publicada no dia 4 de janeiro de 1870. Com sede em Buenos Aires, é um dos principais jornais da Argentina e considerado parte da mídia de referência daquele país. Seu fundador é o general Bartolomé Mitre, que foi historiador, militar e presidente da República em 1862 (ARES, 1985). Disponibiliza conteúdo na *Web* desde 1995.

O *La Nación Data*, de onde foram extraídas as notícias analisadas nesta pesquisa, é uma seção do site em que estão concentrados os conteúdos produzidos a partir de técnicas de jornalismo guiado por dados. Dentro dessa seção ainda há o blog *Nación Data*, que divulga iniciativas *open data* e os bastidores das notícias veiculadas. Na descrição do blog, os jornalistas escrevem que o objetivo é compartilhar informações em um momento tecnológico de disponibilidade maior de dados para desenvolver conteúdo, investigações, visualização e gerar novo conhecimento sobre as bases de dados (NACIÓN DATA, 2017). A equipe do *La Nación Data*⁴⁴ é composta por pessoas de diferentes áreas: além de jornalistas, conta com profissionais que trabalham com operação de sistemas e bases de dados, acesso à informação pública e design.

Em entrevista a Abad (2015), uma das líderes do *La Nación Data*, Florencia Coelho, disse que a ideia de criar uma equipe de dados surgiu em setembro de 2010, em diálogo com Angélica Peralta, então gerente de desenvolvimento multimídia do *La Nación*, após participação em uma oficina de jornalismo de dados para principiantes. Na mesma época, jornalistas precisavam de auxílio da área tecnológica do jornal para acessar bases de dados extensas do *WikiLeaks*. Naquele momento, jornais de referência de outros países, como o *The Guardian*, na Inglaterra, já haviam criado iniciativas específicas para o trabalho com dados.

Um dos principais propósitos do *La Nación Data* é investigar as contas do poder público. A equipe do *La Nación Data* foi premiada três vezes no *Data Journalism Awards*⁴⁵, premiação voltada para publicações em jornalismo guiado por dados. Entre os prêmios, na edição de 2015 como site do ano em jornalismo guiado por dados⁴⁶ e em 2016 na categoria

⁴³ <http://www.lanacion.com.ar/>

⁴⁴ <http://www.lanacion.com.ar/data>

⁴⁵ O *Data Journalism Awards* premia, desde 2012, as melhores produções de jornalismo guiado por dados no mundo. É organizado pela Rede Global de Editores (Global Editors Network) com o apoio do Google News Lab e das fundações John S. e James L. Knight Foundation. As informações completas estão disponíveis no site da premiação: <https://www.datajournalismawards.org/>

⁴⁶ O nome original da categoria é *Data Journalism Website of the Year*.

Dados Abertos. Dessa forma, corresponde a uma das principais iniciativas de jornalismo guiado por dados na América Latina.

4.1.6 *Chequeado* e a verificação de discursos públicos

O *Chequeado*⁴⁷ surgiu no ano de 2010 em um momento de proliferação de informações duvidosas principalmente nos sites de redes sociais, tornando-se um desafio fornecer informações de qualidade no jornalismo (ECHT, 2014). Para tanto, combina técnicas do *fact-checking*⁴⁸, que Echt (2014) considera a comparação dos discursos com fatos e dados, e do jornalismo guiado por dados, ou seja, a extração de histórias a partir de grandes bases de dados e o processo de contextualização.

Trata-se do principal projeto da *Fundação A Voz Pública*⁴⁹, produto jornalístico digital que visa filtrar as declarações de políticos, economistas, empresários, pessoas públicas e outros, de modo que contemple os formadores de opinião pública. Dessa forma, o *Chequeado* busca ser uma ferramenta pública para que os cidadãos tenham consciência da veracidade (ou não) das informações propagadas em discursos. Com isso, o objetivo do *Chequeado* é reduzir a margem de impunidade intelectual que acaba nos discursos públicos (ECHT, 2014).

De acordo com Echt (2014), o *Chequeado* não apenas publica dados, mas faz com que esses números sejam acompanhados de um contexto que permita a compreensão do evento público, visando o aumento da transparência e do debate para aprimorar a democracia na Argentina. Como forma de manter a transparência, o *Chequeado* costuma disponibilizar os dados brutos ou já trabalhados sobre as notícias que cobre.

4.2 A definição das 60 notícias

Definidos os veículos de comunicação, a etapa subsequente desta pesquisa consistiu na escolha das notícias que compõem o *corpus* de análise. Optou-se por eleger dez produções de

⁴⁷ <http://chequeado.com/>

⁴⁸ O *fact-checking* é uma prática de verificação em cima de declarações verificáveis (com números e informações que possam ser verificadas) de políticos e personalidades públicas cujas informações e discursos podem gerar impacto sobre a vida em sociedade. O uso de *fact-checking* ainda visa combater afirmações rasas, confrontando falas com dados verificáveis, valendo-se da missão do jornalismo como mantenedor da verdade e vigia do poder político (GRAVES, NYHAN e REIFLER, 2015).

⁴⁹ A fundação *A Voz Pública* tem como principal projeto o *Chequeado*, que identifica como espaço de checagem do discurso público e de jornalismo guiado por dados. Tem como principais pilares a mídia, por meio do *Chequeado*, e de iniciativas nas redes sociais, a inovação e a educação. Mais sobre a fundação: <http://chequeado.com/acerca-de-la-fundacion/>

cada um dos veículos, totalizando 60 unidades, por entender que o número permite verificar alguma recorrência entre os casos, ainda que seja uma amostra não probabilística.

O *corpus* desta pesquisa é composto por notícias publicadas entre 1º de janeiro de 2016 e 31 de outubro de 2017 e esse é o único recorte temporal adotado. As publicações de *O Globo* e do *Nexo* são as mesmas utilizadas no estudo piloto da qualificação, mas passaram por nova leitura e análise⁵⁰.

Para manter a homogeneidade das publicações, foram selecionadas notícias com profundidade intermediária: não tão rasas e breves quanto uma nota, nem tão aprofundadas quanto uma grande reportagem. Uma notícia geralmente dá conta de temas factuais, do dia a dia, responde às questões do lide (o que, quem, quando, onde, como e por quê) e explora alguns aspectos. Neste trabalho, a prioridade era manter a variedade de temas noticiados. Por isso, a listagem final dos registros noticiosos selecionados abrange publicações de diversas áreas, como política, economia, cultura, polícia, saúde, meio ambiente e outras.

Quadro 3 – Listagem das notícias estudadas

| Veículo | Tema/Editoria | Código/Unidade de análise | Título |
|-------------|-------------------|---------------------------|---|
| <i>Nexo</i> | Educação | UN1 | 7 dados que mostram como está a educação brasileira hoje |
| <i>Nexo</i> | Cultura | UN2 | O horário das refeições no Brasil e no mundo, segundo o Twitter |
| <i>Nexo</i> | Esporte | UN3 | Campeonato Brasileiro: história conturbada, disputas por dinheiro e paixão nacional |
| <i>Nexo</i> | Meio ambiente | UN4 | O que resta da mata atlântica no Brasil |
| <i>Nexo</i> | Política/economia | UN5 | 10 índices econômicos e sociais nos 13 anos de governo PT no Brasil |
| <i>Nexo</i> | Política/economia | UN6 | Públicas e privadas: a divisão de terras no território brasileiro |
| <i>Nexo</i> | Política | UN7 | Qual é o grau de discordância e concordância entre os ministros do Supremo |
| <i>Nexo</i> | Saúde | UA8 | Por que o Japão tem as menores taxas de obesidade do mundo |
| <i>Nexo</i> | Violência | UA9 | Por que o Brasil não dá uma resposta efetiva ao aumento persistente dos homicídios? |
| <i>Nexo</i> | Política/economia | UN10 | Qual o quadro do investimento estrangeiro direto no Brasil |

⁵⁰ O estudo piloto apresentado na qualificação desta dissertação contava, inicialmente, com a criação de categorias como passo inicial para a posterior análise. A partir de sugestão da banca, outro processo foi desenvolvido: primeiramente todas as notícias foram selecionadas e descritas para então ocorrer a definição das categorias com base em semelhanças entre as fontes.

| | | | |
|---------------------------|---------------------|------|---|
| <i>O Globo</i> | Polícia | UN11 | Mapeamos as apreensões de armas de fogo no Rio desde 2007 |
| <i>O Globo</i> | Cultura | UN12 | O mapa de países onde 'Despacito' faz mais sucesso |
| <i>O Globo</i> | Política | UN13 | Total de greves caiu após FHC e depois retomou aumento no fim do 2º governo Lula |
| <i>O Globo</i> | Polícia | UN14 | Denúncias de violência doméstica e familiar contra a mulher crescem 133% |
| <i>O Globo</i> | Política | UN15 | Despesa com vices ultrapassa R\$ 64 milhões em 21 estados |
| <i>O Globo</i> | Segurança pública | UN16 | O mapa interativo do sistema prisional brasileiro |
| <i>O Globo</i> | Brasil | UN17 | Pará tem 30% das mortes em conflitos por terra no Brasil |
| <i>O Globo</i> | Demografia | UN18 | Trabalhadores negros com escolaridade igual e de brancos recebem até 30% menos |
| <i>O Globo</i> | Política | UN19 | Mudança de faixa populacional salva receita do FPM de 13 municípios |
| <i>O Globo</i> | Política e cultura | UN20 | O volume de financiamento de projetos culturais via Lei Rouanet em 4 gráficos |
| <i>The New York Times</i> | Demografia | UN21 | <i>A Secret of Many Urban 20-Somethings: Their Parents Help With the Rent</i> |
| <i>The New York Times</i> | Demografia/economia | UN22 | <i>Actually, Income in Rural America Is Growing, Too</i> |
| <i>The New York Times</i> | Carreira | UN23 | <i>Switching Careers Doesn't Have to Be Hard: Charting Jobs That Are Similar to Yours</i> |
| <i>The New York Times</i> | Gênero/economia | UN24 | <i>The Gender Pay Gap Is Largely Because of Motherhood</i> |
| <i>The New York Times</i> | Segurança pública | UN25 | <i>Is Terrorism Getting Worse? In the West, Yes. In the World, No.</i> |
| <i>The New York Times</i> | Saúde | UN26 | <i>Which Metrics on Hospital Quality Should Patients Pay Attention To?</i> |
| <i>The New York Times</i> | Alimentação | UN27 | <i>How Restricting Food Stamp Choices Can Fight Obesity</i> |
| <i>The New York Times</i> | Demografia | UN28 | <i>How Redlining's Racist Effects Lasted for Decades</i> |
| <i>The New York Times</i> | Violência | UN29 | <i>Inside the Algorithm That Tries to Predict Gun Violence in Chicago</i> |
| <i>The New York Times</i> | Demografia | UN30 | <i>Most Detailed Map of Gay Marriage in America</i> |
| <i>FiveThirtyEight</i> | Saúde | UN31 | <i>Data On Drug Use Is Disappearing Just When We Need It Most</i> |
| <i>FiveThirtyEight</i> | Economia | UN32 | <i>In One Of Latin America's Poorest Countries, Boosting The Economy Is Women's Work</i> |

| | | | |
|------------------------|----------------------------|------|--|
| <i>FiveThirtyEight</i> | Política | UN33 | <i>What The World Thinks Of Trump</i> |
| <i>FiveThirtyEight</i> | Demografía | UN34 | <i>'Normal America' Is Not A Small Town Of White People</i> |
| <i>FiveThirtyEight</i> | Gênero | UN35 | <i>How Much Progress Have We Made Since Women Went On Strike In 1970?</i> |
| <i>FiveThirtyEight</i> | Imigração | UN36 | <i>Democrats Weren't Always Super Liberal On Immigration</i> |
| <i>FiveThirtyEight</i> | Cultura | UN37 | <i>The Ultimate Halloween Candy Power Ranking</i> |
| <i>FiveThirtyEight</i> | Política | UN38 | <i>Sexual Harassment Isn't Just A Silicon Valley Problem</i> |
| <i>FiveThirtyEight</i> | Política | UN39 | <i>Marching Scientists Will Have A Lot In Common With Angry '70s Farmers</i> |
| <i>FiveThirtyEight</i> | Esporte | UN40 | <i>Who Hates The Patriots The Most?</i> |
| <i>Chequeado</i> | Economía | UN41 | <i>Impuesto a las Ganancias: cómo varió la cantidad de personas que lo pagaron en los últimos cinco años</i> |
| <i>Chequeado</i> | Economía | UN42 | <i>Venezuela: cuatro datos para contextualizar la crisis</i> |
| <i>Chequeado</i> | Economía | UN43 | <i>¿Qué muestran los datos sobre los juicios laborales?</i> |
| <i>Chequeado</i> | Educação | UN44 | <i>Aprender 2016: ¿qué relación hay entre el nivel socioeconómico del alumno y que vaya a una escuela pública o privada?</i> |
| <i>Chequeado</i> | Justiça e direitos humanos | UN45 | <i>Baja de imputabilidad: ¿cuál es el perfil de los adolescentes en conflicto con la ley en el país?</i> |
| <i>Chequeado</i> | Demografía | UN46 | <i>Los hogares con padres que no terminaron la primaria cuadruplican los niveles de pobreza de aquellos que la tienen</i> |
| <i>Chequeado</i> | Economía | UN47 | <i>Crisis del sector lechero: datos para entender los problemas de la industria</i> |
| <i>Chequeado</i> | Transparência | UN48 | <i>Índice de Datos Abiertos: en qué mejoró y en qué no la Argentina para subir en el ranking</i> |
| <i>Chequeado</i> | Política | UN49 | <i>La Argentina es el cuarto país de la región con más mujeres en el Congreso</i> |
| <i>Chequeado</i> | Meio ambiente | UN50 | <i>A nueve años de su sanción, ¿se cumple la llamada "ley de bosques"?</i> |
| <i>La Nación</i> | Violência | UN51 | <i>Radiografía de los secuestros exprés: cuándo y dónde se producen y cómo hay que actuar</i> |
| <i>La Nación</i> | Demografía | UN52 | <i>Originales: unos 29 mil argentinos tienen un nombre único que no se repite en todo el país</i> |
| <i>La Nación</i> | Política | UN53 | <i>Los argentinos que figuran con sociedades offshore en paraísos fiscales</i> |
| <i>La Nación</i> | Violência | UN54 | <i>Las cifras sobre violencia de género, antes y después de Ni Una Menos</i> |
| <i>La Nación</i> | Política | UN55 | <i>Mirá el comparativo entre los discursos de Obama y Trump</i> |

| | | | |
|------------------|---------------|------|---|
| <i>La Nación</i> | Telefonia | UN56 | <i>La telefonía móvil y los bancos se mantienen al tope de los reclamos desde hace cinco años</i> |
| <i>La Nación</i> | Saúde | UN57 | <i>El 93% de las personas que esperan un trasplante necesita un riñón o hígado</i> |
| <i>La Nación</i> | Transparência | UN58 | <i>El país escaló 34 lugares en un ránking de transparencia</i> |
| <i>La Nación</i> | Economia | UN59 | <i>Venta de autos 0km: en agosto creció un 16,2% y acumula una suba del 30,4% en el año</i> |
| <i>La Nación</i> | Demografia | UN60 | <i>Qué porcentaje de parejas se divorcia y cuánto suelen durar los matrimonios</i> |

Fonte: A autora (2018)

O Quadro 3 reúne informações sobre o nome do veículo, a área de cobertura em que a notícia se insere, o código da notícia (aqui chamado de unidade de análise) e o título do texto. Os termos utilizados na coluna referente ao tema/editoria foram extraídos das próprias notícias, quando possível identificar a área da qual fazem parte, ou acrescentadas durante o processo de leitura e processamento dos dados. Conhecido o conteúdo com o qual ocorre o estudo de fontes, passa-se, agora, às explicações referentes ao estudo das notícias.

4.3 Etapas de organização dos dados

A organização das informações obtidas a partir da leitura das 60 notícias ocorreu em duas etapas: 1) inicialmente, os dados foram descritos em formato de relatório, em uma ficha de texto; 2) as informações foram sistematizadas em planilhas. As duas etapas serão explicadas nos tópicos seguintes.

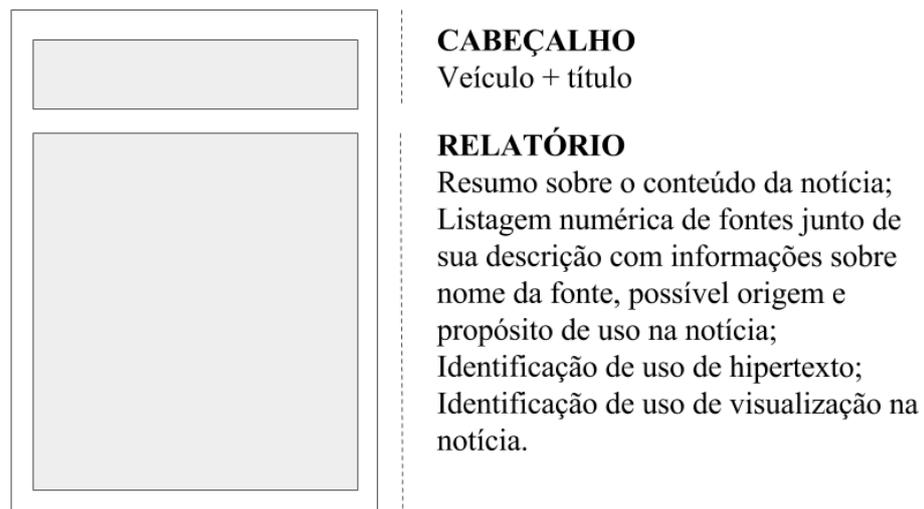
4.3.1 Leitura e descrição em fichas

Inicialmente, cada notícia foi descrita em uma ficha de leitura. O conteúdo foi dividido em duas partes: cabeçalho e relatório (Figura 2). O cabeçalho reúne dados de identificação da notícia, como o número da unidade de análise (uma espécie de código da notícia estabelecida para fins de organização do corpus, que será explicado adiante), o nome do veículo de comunicação e o título do texto, de modo que seja possível identificar rapidamente.

A descrição detalhada da notícia cabe ao espaço identificado como relatório. Com o objetivo de manter a uniformidade no processo descritivo, procurou-se seguir uma lógica para descrever os mesmos aspectos de cada notícia. Em um primeiro momento optou-se pela elaboração de um resumo da notícia, de modo que esteja evidente, em poucas linhas, o tema e

assunto específico sobre o qual a notícia se refere. Na sequência, os itens encontrados (como a ocorrência de visualizações e o uso de hipertexto, por exemplo), além das próprias fontes identificadas, foram listados de acordo com a ordem de surgimento no texto. Assim foi possível realizar apontamentos sobre as características de cada notícia.

Figura 2 – Modelo de ficha de leitura



Fonte: A autora (2018).

Para fins de organização, todas as fontes identificadas foram listadas seguindo uma lógica de número, nome da fonte/origem e informações a ela atribuídas (APÊNDICE A). O objetivo de criar uma ficha de texto inspirada em um relatório descritivo é reunir o máximo possível de informações sobre a notícia e as fontes acionadas – de acordo com as informações possíveis de serem extraídas por meio do texto da notícia. Foram priorizados elementos como hipertexto e visualização, bem como o detalhamento das fontes encontradas. Além das possibilidades apresentadas na Figura 2, são acrescentadas informações mais detalhadas quando perceptíveis no texto – é o caso da indicação, por exemplo, do formato de arquivo utilizado para embasar o conteúdo jornalístico (como atlas, notas, relatórios, livros e outros).

Para que o procedimento fique mais claro, passa-se, agora, a um exemplo prático da notícia *Pará tem 30% das mortes em conflitos por terra no Brasil*, publicada pelo jornal *O Globo* em 25 de maio de 2017. Na área intitulada cabeçalho (Figura 2) consta na ficha de leitura

o número da unidade de análise⁵¹, o nome do veículo de comunicação e o título da notícia. A saber:

Quadro 4 – Cabeçalho da ficha de leitura

| CABEÇALHO |
|--|
| <p>UNIDADE DE ANÁLISE 17</p> <p>O GLOBO Pará tem 30% das mortes em conflitos por terra no Brasil</p> <p>A notícia trata das mortes decorrentes de conflitos rurais no estado do Pará, apontando as áreas em que os crimes ocorrem com mais frequência.</p> |

Fonte: A autora (2018)

Os elementos do cabeçalho ficam estruturados da seguinte maneira: Unidade de Análise 17, *O Globo*, *Pará tem 30% das mortes em conflitos por terra no Brasil*. Na sequência, o resumo da notícia é acrescentado (Quadro 4).

Na área destinada ao relatório (Figura 2) está posta a relação das demais informações encontradas, como as fontes empregadas, o uso de hipertexto e de visualização de dados. Para facilitar a identificação do número de fontes acionadas no texto, cada uma recebeu um número, enquanto os demais elementos não. Desta forma:

Quadro 5 – Ficha de leitura em formato de relatório

| Fonte ou elemento | Descrição |
|---|--|
| 1. Relatórios da Comissão Pastoral da Terra (CPT) | Dados sobre conflitos de terra no Brasil |
| Hipertexto | “É o que mostram os números dos conflitos rurais em disputas por terra no Brasil, ampliados nesta quarta-feira com a <u>morte de mais dez pessoas</u> durante ação policial em uma fazenda no Pará, estado que concentra 30% do total desses assassinatos no país. Levantamento feito pelo <u>Núcleo de Dados</u> , com base em relatórios produzidos pela Comissão da Pastoral da Terra (CPT) desde 1985 [...]” [Link 1 = texto de <i>O Globo</i> sobre ação policial no Pará; Link 2 = link para a página do Núcleo de Dados]. |
| Visualização | Gráfico de barras sobre mortes em conflitos por terra Fonte: Comissão pastoral da terra |
| 2. Censo do IBGE 2010 | Dados sobre disputas agrárias no país |
| 3. Instituto Socioambiental (ISA) - oscip | Percentual de terras indígenas no Brasil |

⁵¹ Os números que representam a unidade de análise vão do 1 até o 60 e permitem que haja uma organização numerada das notícias analisadas.

| | |
|--|--|
| Visualização | Mapa das mortes por conflitos no Brasil Fonte: Comissão Pastoral da Terra |
| 4. Ariovaldo Umbelino, geógrafo e professor da USP | Fala sobre a raiz dos conflitos rurais no Brasil |
| 5. Nota do Instituto Nacional de Colonização e Reforma Agrária (Incra) | Nota sobre violência em áreas rurais e suas origens |

Fonte: A autora (2018)

Na notícia em questão, por exemplo, uma das fontes localizadas foi o Instituto Socioambiental (ISA). Como é a terceira a ser empregada no texto, leva o número três como identificação, seguida de dados sobre sua origem e informações a ela atribuídas. Em uma representação visual explicativa, esta foi a lógica respeitada: trechos com hipertexto foram identificados. Tais informações são transferidas para a ficha de leitura e as palavras ou frases contendo hiperlinks são sublinhados. Os links imediatamente são abertos e, no texto descritivo, identificados o seu lugar de destino. Desta forma, o trecho que leva escrita hipertextual, a exemplo de “[...] morte de mais dez pessoas [...]” (NIKLAS, 2017), tem em sua descrição que o hiperlink remete a outra notícia de *O Globo* que trata de uma ação policial no Pará.

As fichas ainda contam com identificação nos casos em que há visualização de dados. Com relação a esta notícia, observou-se a existência de um gráfico de barras, que na ficha foi caracterizado com o título, um resumo sobre o que significa naquele contexto e a fonte utilizada para sua elaboração.

A criação de detalhadas fichas de leitura permitiu que se perceba a existência de similaridades e diferenças no emprego de fontes em termos de quantidade e do tipo utilizado para dar sustentação aos eventos noticiados. A seguir serão trazidos alguns exemplos de cada veículo, com exploração qualitativa.

4.3.1.1 *O Globo*

A equipe de dados do jornal *O Globo* mostrou que é possível elaborar notícias com técnicas de JGD a partir de fontes pouco convencionais, como a quantidade de reproduções de determinada música a partir de dados da plataforma de streaming *Spotify*. Na publicação que leva o título de *O mapa de países onde ‘Despacito’ faz mais sucesso* (COUTO; JULIÃO; VASCONCELLOS, 2017), os autores apresentam números absolutos de quantidades de vezes em que a música foi ouvida e também de acordo com a população dos países, promovendo uma taxa. A partir de dados do *Spotify*, a equipe desenvolve a narrativa noticiosa e cria visualizações

como um mapa de ocorrências (Figura 3) e um gráfico de linhas em que compara a quantidade de reproduções de *Despacito* com outras músicas famosas.

Figura 3 – O mapa de reprodução da música *Despacito*



Fonte: COUTO, JULIÃO; VASCONCELLOS, 2017 | Captura de tela em 20 nov. 2017

O mapa criado pela equipe de *O Globo* e representado na Figura 3 mostra, a partir da escala de cores, os percentuais de reprodução da música de acordo com o número de habitantes de cada país – ou seja, a partir da taxa de reprodução.

Embora o processamento de dados e a criação de taxas integrem o processo de apuração e tratamento das informações, é preciso destacar que o veículo *O Globo*, quando o faz, deixa claro como o procedimento foi executado: Na notícia *Denúncia de violência doméstica e familiar contra a mulher crescem 133%* (VASCONCELLOS, 2017), o jornal utilizou como base de dados os números da Secretaria Especial de Políticas para as Mulheres e a Lei Maria da Penha. Juntas, essas duas fontes permitiram o emprego numérico dos casos e a aproximação com questões legais. Como a utilização de dados absolutos dificulta as comparações e não uniformiza parâmetros de análise, o Núcleo de Dados explica, no texto, que organizou as informações em um mapa e calculou a taxa dos casos: “O Núcleo de Dados tabulou os dados e organizou tudo em um mapa, segundo a taxa de denúncias de violência contra grupos de cada

100 mil mulheres” (VASCONCELLOS, 2017, s.p.). Trata-se de um cuidado metodológico e uma clareza necessária para que se saiba como tais resultados foram obtidos. A descrição de procedimentos, também chamada de transparência no método, é parte importante do jornalismo guiado por dados.

Mais um caso de *O Globo* diz respeito à notícia *Despesa com vices ultrapassa R\$ 64 milhões em 21 estados* (COUTO, 2017), que alia informações de portais da transparência com o contraponto de políticos e assessorias de imprensa relacionadas. Ao mesmo tempo, porém, poderia ter ficado mais claro quais foram as bases de dados empregadas – o que seria possível com o emprego de hiperlinks. O trecho que explica como foi feita a consulta informa que “Um levantamento feito pelo Núcleo de Jornalismo de Dados do GLOBO a partir dos portais de transparência e de informações dos próprios governos revela que [...]” (COUTO, 2017, s.p.). Enquanto no texto não fica tão claro que as informações foram extraídas de sites de governos estaduais, a fonte de uma das visualizações de dados cita os portais da transparência de governos estaduais e portal da transparência do Governo Federal, evidenciando que houve pesquisas em mais de uma base de dados e a combinação dessas informações. Além dos números levantados a partir de repositórios e contrastados com as versões oficiais de políticos e assessorias, *O Globo* preocupou-se em ouvir especialistas para fazer uma análise sobre o papel dos vice-governadores e sobre o cargo de vice em outros países. O que deixa a desejar é o fato de a equipe não explorar o uso de hiperlinks para que seja possível acessar a origem da fonte, ou seja, de onde os dados foram extraídos.

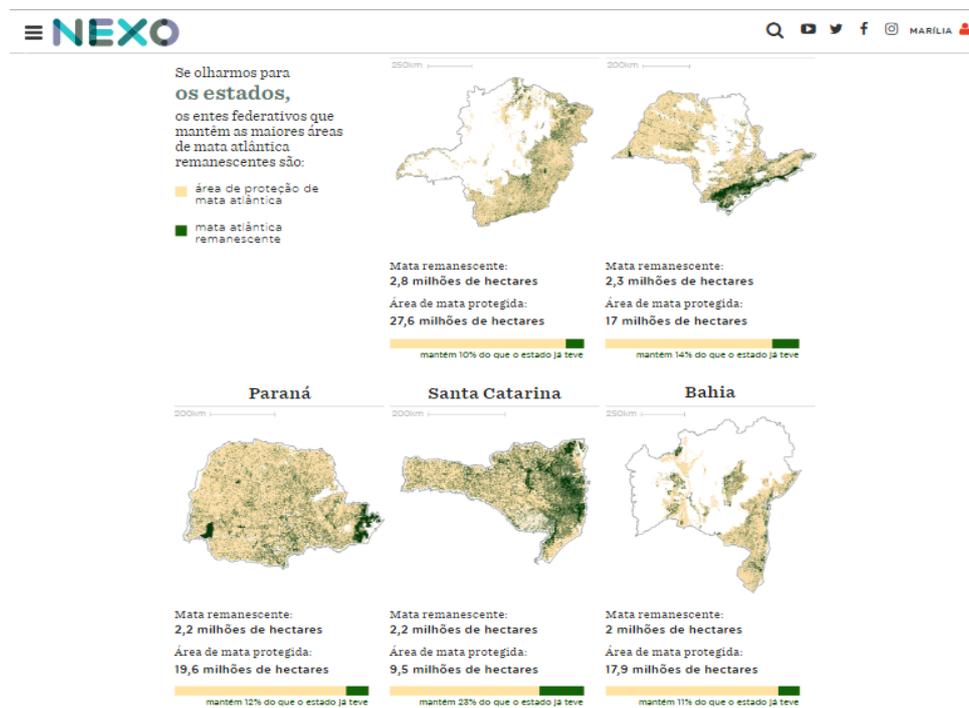
4.3.1.2 *Nexo*

Nativo digital brasileiro, o *Nexo Jornal* apresenta algumas notícias que não têm a participação humana como fonte. Os textos são sustentados a partir da leitura de documentos e de publicações editoriais. O produto revela, nesse caso, o que parece ser um exaustivo trabalho de pesquisa, a exemplo do que foi veiculado na publicação *Campeonato Brasileiro: história conturbada, disputas por dinheiro e paixão nacional* (PIMENTEL, 2017). Todas as fontes acionadas possuem caráter de documento, ou seja, não foi necessário adotar procedimentos tradicionais de pergunta-resposta em entrevista. Para que a origem dos documentos acompanhe a fonte, ou seja, de onde as informações foram extraídas, o *Nexo* costuma oferecer a escrita hipertextual como recurso, de modo que o destino final para outros locais seja demonstrado.

Com a autodeclarada missão de explicar as notícias, o *Nexo* acaba desenhando explicações complexas, isto é, oferece gráficos e mapas elaborados com base em relatórios,

atlas e outras publicações. Dentre as dez notícias estudadas neste trabalho, duas apresentam características semelhantes. Isso acontece nas notícias intituladas *Públicas e privadas: a divisão de terras no território brasileiro* (ALMEIDA; MARIANI, 2017), que teve como base o Atlas da Agropecuária Brasileira, o Instituto Nacional de Reforma Agrária e a Rádio CBN, e *O que resta da Mata Atlântica no Brasil* (ALMEIDA, 2017), que mostra em quais estados a mata segue preservada e em quais espaços a preservação está comprometida (Figura 4). Esse segundo texto mencionado foi produzido a partir do Atlas dos Remanescentes Florestais da Mata Atlântica – SOS Mata Atlântica e do Instituto Nacional de Pesquisas Espaciais (INPE).

Figura 4 – Visualização com mapas sobre a preservação da Mata Atlântica

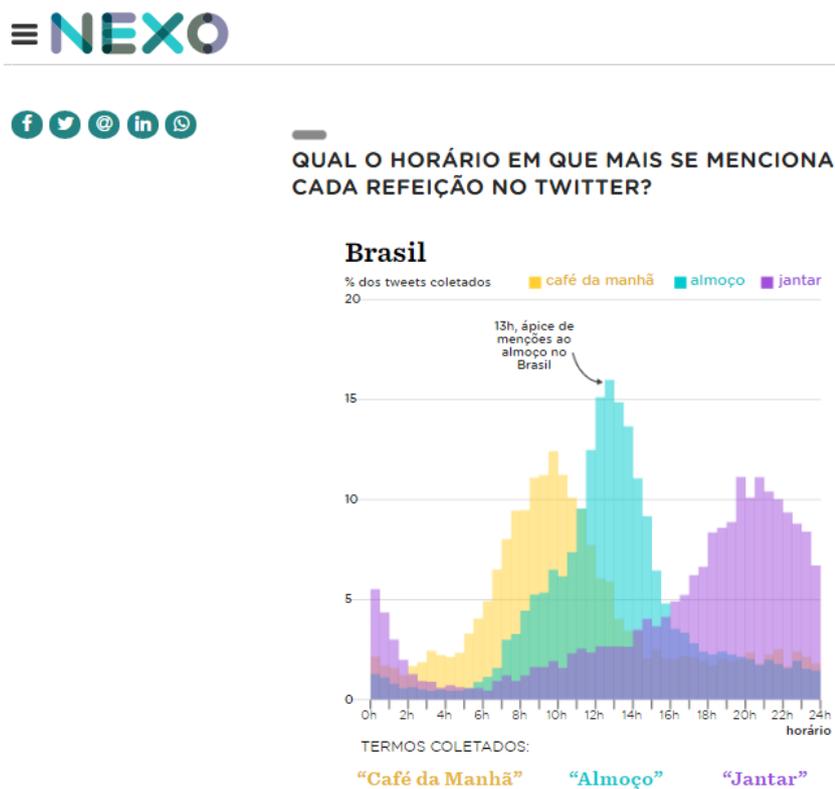


A Figura 4 mostra como as fontes se materializam em visualizações didáticas e explicativas. Mais do que informar os números referentes à mata remanescente em diferentes estados brasileiros, o *Nexo* se preocupou em representar graficamente tais números, facilitando a compreensão. O verde, que contrasta com a cor mais clara, serve para indicar o quanto ainda resta de Mata Atlântica.

O *Nexo* também costuma utilizar dados de sites de redes sociais para a elaboração de conteúdo jornalístico. A publicação *O horário das refeições no Brasil e no mundo, segundo o Twitter* (MARIANI; ALMEIDA, 2017) se utilizou de informações produzidas pelos usuários

para explorar gráficos baseados no horário em que as pessoas comem no Brasil (Figura 5) e em outros países. Além de exibir uma sequência de visualizações, o material apresenta uma nota metodológica elaborada pelos autores, que informa quando os dados foram coletados, quais os termos utilizados, em quais países e que tipos de usuários foram removidos para não comprometer o levantamento, já que o *Twitter* é uma rede exposta à criação de robôs que realizam postagens automáticas e, dessa forma, poderiam provocar um resultado enviesado.

Figura 5 – As refeições pelo Brasil e pelo mundo segundo o *Twitter*



Fonte: MARIANI; ALMEIDA, 2017. | Captura de tela em 20 nov. 2017

Na Figura 5 é possível perceber que os dados do *Twitter* como fonte geraram visualizações claras e compreensíveis sobre os horários mais mencionados de refeições pelo Brasil. A visualização, última etapa do jornalismo guiado por dados, por vezes traduz melhor as informações do que o texto.

Há, também, a combinação de fontes documentais e pessoais oferecidas no *Nexo*. Na notícia *Por que o Brasil não dá uma resposta efetiva ao aumento persistente dos homicídios?* (FÁBIO, 2017), o veículo se utiliza do Atlas da Violência, elaborado pelo Instituto de Pesquisa Econômica Aplicada (IPEA), como fonte principal junto com informações do Fórum Brasileiro

de Segurança Pública. Além de utilizar visualizações de dados para facilitar a localização dos números, a publicação traz a fala de dois especialistas para explicar os números do atlas, apresentando uma variedade de fontes empregadas.

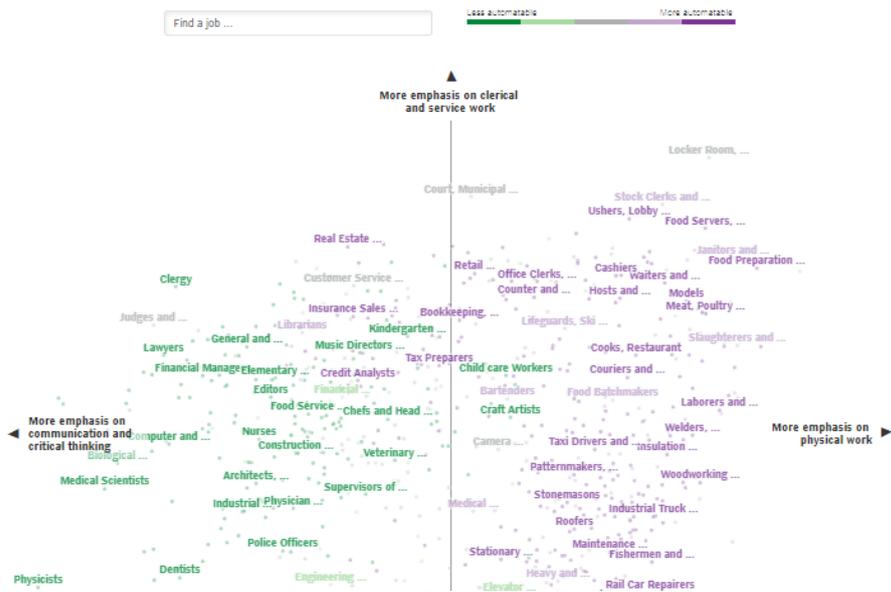
4.3.1.3 *The New York Times*

Este jornal apresenta, por sua vez, combinações de estudos científicos, bases de dados, falas de especialistas e visualizações nos conteúdos produzidos a partir de técnicas do JGD. Além disso, oferece conteúdo que combina o uso de dados com a personalização voltada ao leitor. A notícia intitulada *Switching careers doesn't have to be hard: charting jobs that are similar to yours* (MILLER; BUI, 2017), por exemplo, consiste na elaboração de um mapa de profissões a partir de bases de dados cujas atribuições profissionais são reunidas por similaridade, de acordo com as habilidades necessárias. Dessa forma, quando o leitor insere o nome de sua função, a base de dados permite que situe sua profissão com outras que possuem atividades parecidas. O infográfico (Figura 6) em questão foi elaborado com base em diferentes fontes, como o Departamento de Trabalho do governo e uma pesquisa envolvendo profissões. Além disso, foi complementada com histórias reais de pessoas que precisaram mudar de carreira e conseguiram encontrar soluções em espaços similares. Este exemplo combina o uso de bases de dados e de entrevistas com pessoas, além de oferecer algo diferente ao usuário, permitindo a interação a partir da inserção de seus próprios dados.

Figura 6 – Profissões de acordo com a similaridade de atividades

Where Do American Jobs Fall on the Skill Spectrum?

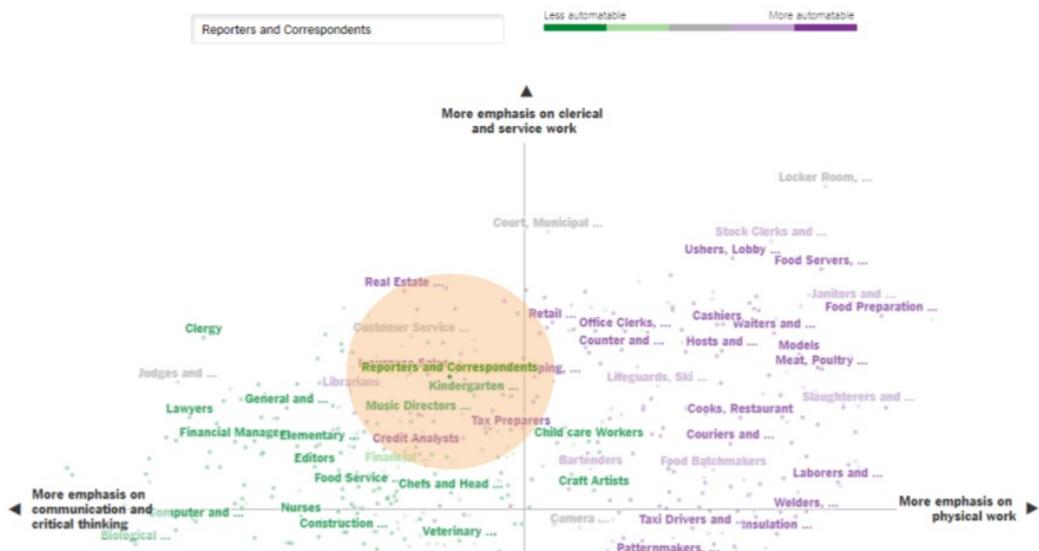
A look at detailed skills data from the Labor Department shows how similar and dissimilar jobs are in the American labor market.



Fonte: MILLER; BUI, 2017. | Captura de tela em 20 nov. 2017

A figura a seguir mostra uma captura de tela com inserção de dados:

Figura 7 – Inserção de profissão em infográfico interativo



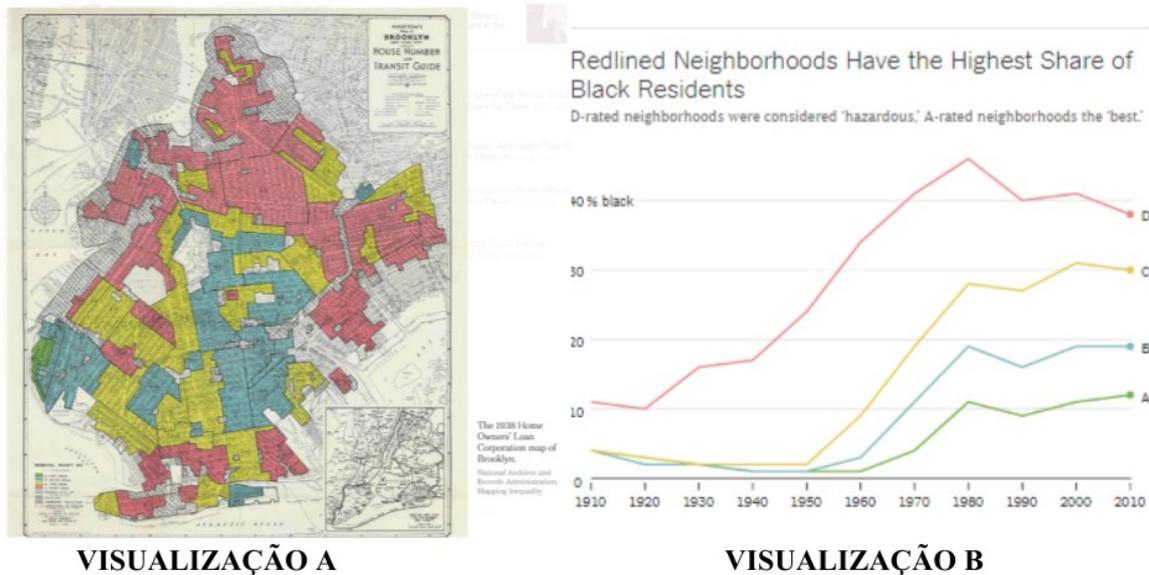
Fonte: MILLER e BUI, 2017. | Captura de tela em 20 nov. 2017

A Figura 6 mostra a barra de pesquisas, abaixo do título, em que há espaço para inserção das profissões e a descoberta de quais atividades são mais similares com os dados inseridos pelo leitor e calculados a partir da base de dados incorporada à notícia. A Figura 7 mostra o infográfico na prática: a inserção dos termos *Journalist – Reporters and Correspondents*, perceptível na área circular destacada, posicionou a profissão entre carreiras cujas atividades estão mais conectadas à comunicação e ao pensamento crítico.

Outra notícia explica o algoritmo que a polícia de Chicago utiliza para tentar antecipar os casos de violência na cidade, ou seja, quais são as chances de uma pessoa se envolver nessas ocorrências como autor ou vítima. O caso gerou polêmica nos Estados Unidos porque tenta prever situações a partir de uma série de pressupostos que não estão claros, já que o algoritmo foi desenhado a partir de dados não tornados públicos pela polícia. Dessa forma, questiona-se o que torna uma pessoa mais ou menos suspeita ou mais ou menos sujeita a ser vítima de alguma ocorrência. Na tentativa de compreender e desconstruir essa série de comandos, o veículo apresenta, no texto intitulado *Inside the algorithm that tries to predict gun violence in Chicago* (ASHER; ARTHUR, 2017), informações extraídas de bases de dados, como números sobre criminalidade, consulta a estudiosos do tema, notícias do próprio veículo e também de concorrentes. Até mesmo faz uma comparação com o filme *Minority Report*, com hiperlink para a sinopse, sobre um filme de ficção científica em que a polícia era capaz de prender criminosos a partir de conhecimento prévio e antes que o crime fosse concretizado.

Fontes documentais antigas e digitalizadas correspondem a outro recurso utilizado pelo veículo como base para sustentação em notícias. É o caso da notícia chamada *How redlining's racist effects lasted for decades* (BADGER, 2017), que utiliza mapas antigos, da década de 1930, para ajudar a descrever a região do Brooklyn (Figura 8) à época e mapear os grupos que lá moravam, como trabalhadores, estrangeiros e judeus. Com base no público que residia em cada região, uma organização apoiada pelo governo classificava as áreas entre mais e menos nobres ou então mais ou menos perigosas, codificadas segundo padrões raciais e de acesso a crédito. Por meio da fala de estudiosos da área de economia, descobriu-se, por exemplo, que maior parte dos negros eram moradores de áreas consideradas inferiores, de menor qualidade de moradia. Por isso é possível visualizar, a partir de um mapa, a desigualdade que perdurava sobre os bairros. O uso de um documento antigo como fonte sugere que materiais de arquivo podem ser utilizados para explicar questões do presente e do futuro.

Figura 8 – Mapa antigo do Brooklyn



Fonte: BADGER, 2017 | Captura de tela em 20 nov. 2017

O jornal mostrou, nas duas visualizações, que é possível aliar documentos antigos com novas visualizações para explicar um contexto. A visualização A da Figura 8 mostra como era o mapa do Brooklyn nos anos de 1930, com áreas mais e menos valorizadas, sendo que a área D, habitada por minorias, era considerada a mais perigosa. Já a visualização B da mesma figura exibe uma visualização de dados criada na atualidade, porém com dados desde 1910. O gráfico de linhas mostra que as áreas classificadas como “D”, consideradas as menos nobres e mais perigosas, eram habitadas por pessoas negras, em sua maioria.

Outro caso publicado pelo jornal diz respeito à notícia intitulada *Actually, income in rural america is growing, too* (BUI, 2016a). Em resumo, dados divulgados pelo Census Bureau, divisão de governo responsável por trabalhar com os dados populacionais, inicialmente informaram que havia uma elevação dos rendimentos da classe média americana e redução da pobreza, só que esses resultados não se aplicavam à área rural dos Estados Unidos. No entanto, o levantamento começou a ser questionado por outras pessoas, que elaboraram artigos para discutir a questão. Descobriu-se, então, que os rendimentos na área rural sequer caíram. O problema era metodológico: a notícia explica que a definição do termo rural é muito fechada no censo do país. Houve, ainda, um problema de amostragem entre pesquisas realizadas pelo Census Bureau. A notícia acaba discutindo, então, mais do que números, mas problemas metodológicos das próprias pesquisas e que acabam sendo replicados pelo jornalismo por vezes sem a devida análise crítica. Inicialmente, os jornalistas embasaram seus textos na primeira

planilha liberada, a que havia problemas. Mesmo sem haver significância estatística, os dados foram reproduzidos nas notícias. Gera, portanto, reflexão sobre o cuidado metodológico.

4.3.1.4 *FiveThirtyEight*

O *FiveThirtyEight* apresenta algumas diferenças em relação aos demais veículos. Uma de suas particularidades é oferecer o uso de notas explicativas, semelhantes a notas de rodapé, para explicar expressões ou situações que demandam complemento. A notícia intitulada *Data on drug use is disappearing just when we need it most* (CASTEEL, 2017, s.p.) aborda a epidemia no consumo de heroína nos Estados Unidos. No trecho “Mortes por overdose de heroína cresceram mais do que seis vezes em menos de uma década e meia⁵²” há o vínculo com uma dessas notas explicativas. Quando aberta, tal nota se transforma em um parágrafo com dados do Centro de Controle e Prevenção de Doenças dos Estados Unidos informando que o número de mortes saltou de 2 mil, em 2002, para 13 mil, em 2015. Ainda na nota explicativa foi citado um levantamento feito pelo *The New York Times* que sugere que o número de overdoses de diferentes tipos de drogas, não só heroína, atingiu o patamar de 65 mil casos em 2016 (Figura 9). As notas explicativas utilizadas pelo veículo fornecem autonomia ao leitor, que pode escolher se pretende ampliar o nível de detalhamento em determinada informação. Além dos dados de governo e de levantamento próprio do NYT, esta notícia foi baseada em pesquisas e na fala de especialistas que trabalham com o tema.

⁵² No original “Heroin overdose deaths have risen more than sixfold in less than a decade and a half”.

Figura 9 – Uso de notas pelo *FiveThirtyEight*

JUN. 29, 2017 AT 10:00 AM

Data On Drug Use Is Disappearing Just When We Need It Most

By [Kathryn Casteel](#)
Filed under [Opioids](#)

[f](#) [t](#) [e](#)

It's no secret that [heroin](#) has become an epidemic in the United States. Heroin overdose deaths have risen more than sixfold in less than a decade and a half.^x

From around [2,000 in 2002](#) to [roughly 13,000 in 2015](#), according to data from the Centers for Disease Control and Prevention. [Data collected and analyzed by The New York Times](#) suggests that deaths from all drug overdoses, not just heroin, could have risen to 65,000 in 2016.

Yet according to one of the most widely cited sources of data on drug use, the number of Americans using heroin has risen far more slowly, roughly doubling during the same time period.^x

The Substance Abuse and Mental Health Services Administration defines "current users" as those who have used heroin in the past month.

RECOMMENDED

Will The GOP Pass Anything Before Midterms?

The Values That 'Values Voters' Care About Most Are Policies, Not Character Traits

The GOP's Abortion Ban Is About Politics, Not Science

Fonte: CASTEEL, 2017 | Captura de tela em 20 nov. 2017

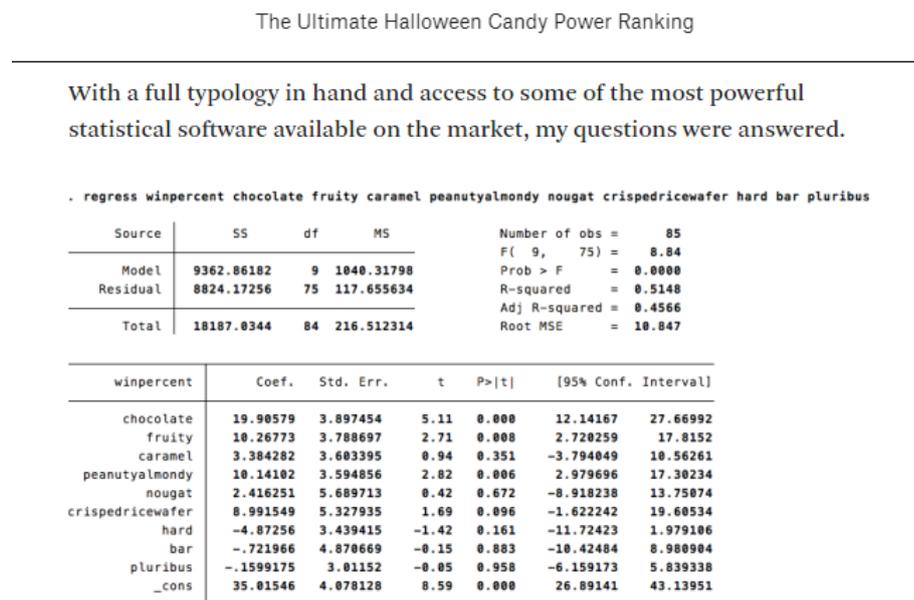
Os termos sublinhados no texto representado na Figura 9 mostram o uso de notas explicativas no texto do veículo, o que auxilia na compreensão de termos às vezes não tão claros ou que demandam mais explicações.

Outro recurso utilizado pelo *FiveThirtyEight* é a geração de dados próprios para a elaboração de notícias, o que geralmente é feito por meio de pesquisas com a participação dos leitores. Um dos exemplos está no texto intitulado *Who hates The Patriots the most?* (PAINE, 2017), sobre o bom desempenho do time de futebol americano *The Patriots* na *National Football League* (NFL) e o sentimento que resultados tão positivos podem provocar em torcedores de times rivais. Para descobrir quais times são mais adorados e detestados na liga NFL, o *FiveThirtyEight* conduziu uma pesquisa online aberta por seis dias com adultos autodeclarados fãs da liga. A pesquisa foi a única fonte identificável utilizada ao longo do texto, então a narrativa foi conduzida basicamente em torno dela. Houve utilização de escrita hipertextual com informações a título de curiosidade remetendo a outros veículos de comunicação.

Caso semelhante é do texto *The ultimate Halloween candy power ranking* (HICKEY, 2017). A partir de pesquisa online, o *FiveThirtyEight* buscou descobrir quais são os doces preferidos do público a fim de identificar o que faz um mais desejável do que o outro. A partir das descobertas da pesquisa, o autor do texto trabalhou com os ingredientes dos doces eleitos

pelo público. Utilizando uma regressão linear estatística (Figura 10), apontou quais ingredientes são decisivos para um doce ser considerado melhor ou pior na concepção do público e que tendo chocolate na composição as chances de ele ser bem aceito pelo público são maiores. Além de empregar estatística para finalidade pouco comum em texto jornalístico, o autor consegue estabelecer, ao final do texto, uma espécie de doce *Frankenstein*: a partir da preferência do público, estabelece qual seria a mistura perfeita para o doce de Halloween, como a quantidade de açúcar existente em determinada marca associada à quantidade de chocolate presente em outra.

Figura 10 – Uso de estatística para entender a escolha por chocolate



That's a lot to take in! In general, here's what this information says. According to the regression, about half the variance observed in the quality can be explained by these nine properties of candy, which isn't great but also isn't awful — and is thus enough for us to work with.

Fonte: HICKEY, 2017 | Captura de tela em 20 nov. 2017

Por meio de análise estatística, na Figura 10, foi possível definir a significância dos componentes dos doces.

Outra abordagem do *FiveThirtyEight* tem a ver com uma notícia sobre questões demográficas, intitulada *'Normal America' is not a small town of white people* (KOLKO, 2016). A partir de cálculos estatísticos, a notícia mostra que os Estados Unidos não são o que pensa o senso comum. Para fazer o cálculo, o autor compara quanto demograficamente são similares

áreas metropolitanas e outras áreas com base em índices como idade, educação, raça e etnia. Por meio de tabelas, a publicação mostra quais áreas metropolitanas são demograficamente mais parecidas com o país hoje e quais se assemelham ao que os Estados Unidos eram nos anos de 1950. Esse tipo de abordagem mostra que é possível criar comparações com base nos dados e que é preciso considerar metodologicamente índices comparáveis.

4.3.1.5 *La Nación*

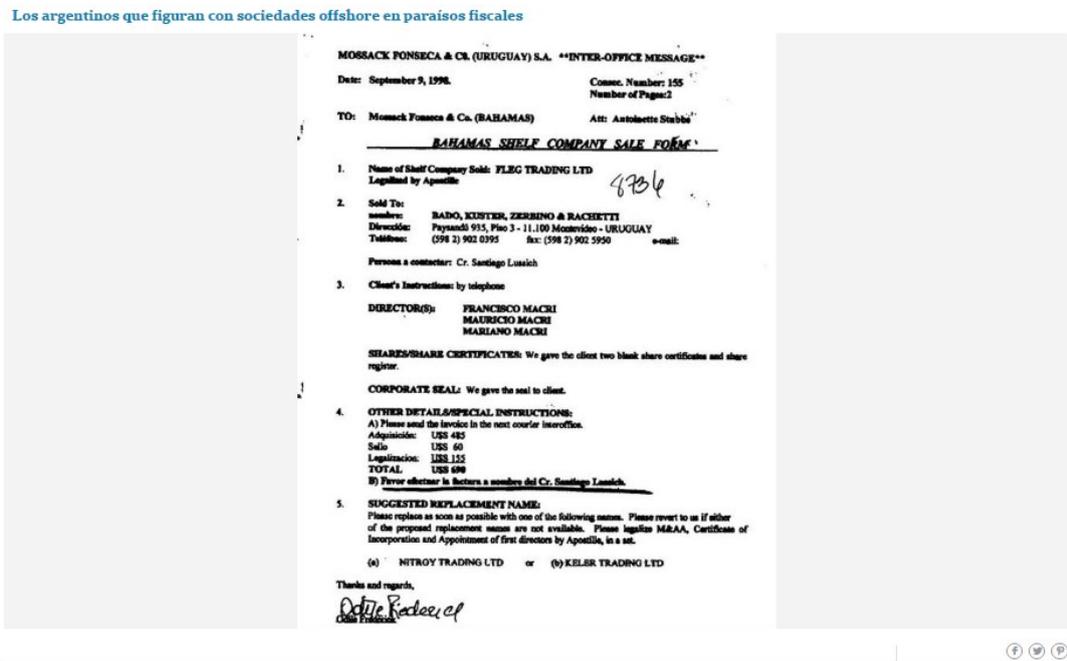
Um dos casos apresentados pelo *La Nación* é uma versão do *Panama Papers* – investigação jornalística em rede a partir de documentos vazados por um escritório sobre a participação de pessoas públicas em empresas *offshores*. O veículo informa que, na Argentina, com a colaboração do Canal 13⁵³, analisou mais de 11 milhões de documentos de sociedades offshore, deixando claro que a investigação é em rede e via Consórcio Internacional de Jornalistas Investigativos. Apresenta interessante uso de fontes. Para aproximar o caso de envolvidos na Argentina, o veículo criou uma notícia específica intitulada *Los argentinos que figuran com sociedades offshore em paraísos fiscales* (RUIZ; JASTREBLANSKY; MON, 2016).

A partir dos documentos vazados, identificou a participação de personalidades como o presidente Mauricio Macri, o ex-presidente Néstor Kirchner e o jogador de futebol Lionel Messi. Um documento vazado e veiculado junto à notícia (Figura 11) mostra o nome do presidente e de familiares em uma empresa que funcionou até o ano de 2009, mas o próprio *La Nación* afirma no texto que não é possível concluir, somente a partir do documento, se o presidente continuava vinculado à sociedade até aquele momento ou se havia se retirado. De qualquer forma, o documento digitalizado é disponibilizado aos leitores, de modo que possam examiná-lo. A resposta da presidência da República ocorreu por meio de nota oficial, explicando que a participação de Mauricio Macri estava ligada a um grupo empresarial familiar e que havia sido designado ocasionalmente como diretor, porém sem participação acionária. O veículo ainda ouviu porta-vozes do governo, que confirmaram o caráter circunstancial de tal cargo, e outras pessoas que possuem algum tipo de ligação com políticos que apareceram na lista do *Panama Papers*. Esse caso apresenta um documento vazado, nota oficial enviada em nome do presidente, tentativas frustradas de entrar em contato com outras fontes e o uso de

⁵³ Entre os veículos argentinos, a investigação do Panama Papers foi realizada pelo La Nación e o Canal 13, integrantes do Consórcio Internacional de Jornalistas Investigativos (ICIJ, na sigla em inglês). Mais detalhes: <http://www.lanacion.com.ar/1885774-panama-papers-documentos-secretos-revelan-operaciones-de-lideres-mundiales-en-paraísos-fiscales>

escrita hipertextual como mais um recurso adotado, especialmente para explicar a investigação internacional e os veículos mobilizados.

Figura 11 – Documento ligado ao *Panama Papers*



Fonte: RUIZ; JASTREBLANSKY; MON, 2016. | Captura de tela em 20 nov. 2017

O documento publicado pelo jornal e reproduzido na Figura 11 procura apresentar, de forma transparente, o documento utilizado como fonte na notícia.

Outro exemplo diz respeito ao uso de dados qualitativos em uma notícia com técnicas de jornalismo guiado por dados. No texto *Mirá el comparativo entre los discursos de Obama y Trump*, o veículo transcreveu, traduziu para o espanhol e comparou os discursos de posse do ex-presidente dos Estados Unidos, Barack Obama, e o atual presidente, Donald Trump. O vídeo de cada discurso, em inglês, foi disponibilizado junto ao conteúdo, oferecendo ao leitor a fonte de onde ambos os discursos podem ser conferidos na íntegra, na versão original. Na notícia, as palavras viraram números e infográficos: há informações sobre o tempo de duração de cada discurso e o percentual ocupado com temas abrangentes como o governo anterior, a política nacional, a política exterior e os valores do povo dos Estados Unidos (Figura 12). Em um esquema de cores, os trechos correspondem aos temas abordados nas falas de Obama e Trump. Há, ainda, ao final, uma nuvem de palavras contendo os termos mais mencionados.

Figura 12 – Dados qualitativos transformam-se em números



Fonte: LA NACIÓN, 2016 | Captura de tela em 20 nov. 2017

É possível perceber, na Figura 12, que os dados qualitativos ligados ao discurso dos presidentes foram transformados em números e quantificados, tornando-se matéria-prima para o jornalismo guiado por dados.

Mais um exemplo diz respeito a uma notícia sobre 29 mil nomes originais na Argentina (COSTA, 2017). A publicação combina duas fontes principais: o Registro Nacional das Pessoas, que representa a base de dados usada para construir a notícia, e fontes ligadas ao Registro Civil, sem citar nomes. O texto ainda traz informações sobre uma lei argentina de 1969 que passou a estabelecer limite de no máximo três nomes para as pessoas. Esta notícia é um exemplo interessante do uso de uma extensa base de dados como fonte e sua aplicação junto ao público, que pode utilizar as informações e buscar por inicial do nome ou fazer pesquisas gerais. Ainda que a quantidade de nomes inéditos esteja na casa dos 29 mil, a base de dados, segundo informa a notícia, tem 60 milhões de pessoas cadastradas – o que torna um desafio para os jornalistas trabalharem com tanto volume de informação.

Os sequestros relâmpagos também foram tema de notícia e de visualizações de dados. O *La Nación* utilizou os dados da Unidade Fiscal Especializada em Sequestros Extorsivos para

reunir as principais características das vítimas: homens adultos, ocorrências entre 20 horas e meia-noite, com duração de até quatro horas e com a vítima mantida em um veículo (AMAYA; PALLARO, 2017). Além das estatísticas que dão sustentação à notícia, o jornal também teve fontes da instituição e menciona casos de sequestro que acabaram em morte, oferecendo o hiperlink para que seja possível acessar a origem das informações em publicações anteriores.

4.3.1.6 *Chequeado*

O *Chequeado*, da Argentina, igualmente trabalhou com fontes diferentes em seus textos. Uma das notícias tem como motivo principal verificar o cumprimento da chamada Lei de Florestas, o que faz com que a legislação seja a principal fonte da notícia. Além da própria lei, um decreto assinado pela então presidente Cristina Kirchner autorizou a criação de um fundo para assegurar o cumprimento da lei. A notícia, intitulada *A nueve años de su sanción, ¿se cumple la llamada 'ley de bosques'?* (MIRANDE, 2016), se utiliza de documento assinado pela chefia de gabinete de ministros da Argentina com explicações sobre o valor que seria destinado para o cumprimento da lei e acabou sendo empregado com outra finalidade. Há, ainda, outras fontes de caráter documental que foram acionadas, como a própria assessoria de imprensa do Ministério do Meio Ambiente e Desenvolvimento Sustentável, que enviou nota ao *Chequeado*. Como este texto jornalístico está baseado em diferentes informes, empregados como fonte, o *Chequeado* procura oferecer ao leitor, por meio de escrita hipertextual, os arquivos originais, quando possível.

Na notícia intitulada *¿Que muestran los datos sobre los juicios laborales?* (DI SANTI, 2017), sobre a ocorrência de processos trabalhistas na Argentina, apresenta um caso diferenciado sob o ponto de vista de uso de fontes. Pelas marcas do produto, é possível inferir que a pauta surgiu a partir de dois documentos: uma nota emitida pelo presidente argentino, Mauricio Macri, sobre o que chamou de máfia dos processos trabalhistas e os danos às pequenas empresas, e uma publicação no Twitter por parte de Florencio Randazzo, pré-candidato ao senado, que questionou a fala do presidente. A partir dessas informações, dispostas no texto (Figura 13), e da possibilidade de contraponto, o *Chequeado* desenvolveu um texto em cima de bases de dados da Superintendência de Riscos do Trabalho. A propósito, o *Chequeado* divulgou, ao público, a planilha utilizada nesse caso. O oferecimento das planilhas aos leitores, de modo que seja possível conferir os dados publicados e sua origem, é prática frequente no veículo.

Figura 13 – Nota e publicação no *Twitter*

The image shows a screenshot of a Twitter post by Florencio Randazzo (@Randazzo) and a linked article from the newspaper El Puntal. The article is titled "¿Qué muestran los datos sobre los juicios laborales?" and discusses the impact of labor lawsuits on small businesses and employees. The article text includes: "El presidente, Mauricio Macri, volvió a apuntar contra 'la mafia que tanto daño le hace a las pequeñas empresas', y el precandidato a senador Florencio Randazzo le contestó por Twitter. Los juicios crecieron y los accidentes no, pero tampoco se advierte un impacto inmediato por la modificación de la ley de ART." and "En una [nota de opinión](#) publicada en el diario cordobés El Puntal, el presidente de la Nación, **Mauricio Macri**, apuntó contra 'la mafia de los juicios laborales, que tanto daño les hacen a las pequeñas empresas y sus trabajadores'." The Twitter post shows the user's profile, the text of the tweet, and the article preview with a photo of Florencio Randazzo.

Fonte: DI SANTI, 2017 | Figura elaborada a partir de captura de tela

A Figura 13 mostra, à esquerda, o hipertexto em vermelho que remete à publicação de Randazzo no *Twitter*, à direita, que acabou sendo utilizada pelo veículo como uma das fontes de informação.

A abordagem sobre a importância do uso de dados também pode se tornar conteúdo jornalístico. Os veículos argentinos parecem importar-se com o desempenho do país em rankings de transparência. O *Chequeado*, por exemplo, publicou uma notícia intitulada Índice de *Datos Abiertos: em qué mejoró y em qué no la Argentina para subir en el ranking* (MARTÍNEZ, 2017). E, para tanto, utilizou dados da Fundação *Open Knowledge International* sobre as posições de países no ranking, além de uma publicação, no *Twitter*, do ministro da Modernização da Argentina, Andrés Ibarra. Mais uma vez o *Twitter* serviu como fonte, não sendo necessário entrevistar a pessoa que escreveu a publicação. Quando extrai informações da rede, o *Chequeado* costuma oferecer ao público a fonte original, como o link que remete à postagem.

Publicações sobre questões de gênero e política integram o escopo do veículo. Em uma notícia sobre a Argentina ser o quarto país da América do Sul com mais mulheres no Congresso (SOHR, 2017), o *Chequeado* se utiliza de levantamentos estatísticos do Diretório Legislativo, que trabalha com informações sobre a participação de mulheres na política em países da

América Latina. As informações ainda são complementadas com outras bases de dados, como a do Observatório Eleitoral Argentino e o Informe da União Interparlamentar. Para completar o processo de contextualização, ainda são ouvidas especialistas mulheres que trabalham com o tema e explicam que os partidos ainda são misóginos e há oportunidades limitadas para mulheres na política.

O *Chequeado* também elabora notícias apenas com uma ou duas fontes: um texto sobre lares em que os pais têm apenas a formação primária tem mais chances de níveis de pobreza, a principal fonte utilizada é o Instituto Nacional de Estatística e Censos, conectada aos problemas monetários que enfrentam as crianças e como isso se reproduz mais tarde (MARTÍNEZ, 2017). Para complementar os números, um especialista em monitoramento de políticas públicas da Argentina é ouvido para comentar os casos.

4.3.1.7 O que mostram os exemplos

Todos os veículos jornalísticos aqui estudados utilizam dados como fonte no trabalho de reportagem, mas os procedimentos se materializam de formas diferentes no produto final. Alguns empregam somente fontes de origem numérica, como bases de dados, enquanto outros complementam ouvindo especialistas e recorrendo a outras fontes para dar vida às histórias. O número de fontes varia em cada caso, mas o menor encontrado foi uma única fonte, em notícias de diferentes veículos, como os registros noticiosos do *Nexo* sobre o horário de refeições (MARIANI; ALMEIDA, 2017) e sobre o quadro de investimento estrangeiro no Brasil (CASTRO, 2017); e o maior foi 17, em uma publicação do *The New York Times* sobre mudanças de carreira (MILLER; BUI, 2017).

Em linhas gerais, o jornal *O Globo* se utiliza principalmente de bases de dados públicas para dar sustentação às notícias e cria taxas e índices em cima dessas informações. Já o *Nexo* emprega, além das bases de dados, estudos científicos e fontes menos convencionais, como sites de rede social. O *The New York Times*, por sua vez, adota em todas as notícias algum tipo de escrita hipertextual ligado à fonte, de modo que seja possível consultar a origem das informações. O *FiveThirtyEight* tem como principal diferencial o uso de notas explicativas ao longo do texto, o que facilita, enquanto fonte, explicações sobre termos e procedimentos adotados na notícia. O *La Nación* conseguiu trabalhar com técnicas de jornalismo guiado por dados a partir de casos pouco comuns, como o vazamento de informações por meio de documento e a transformação do discurso, de caráter qualitativo, em quantidade e dados. O *Chequeado*, que executa checagem de discurso e jornalismo guiado por dados, parece destacar

sobretudo a transparência no método enquanto procedimento de uso de fonte, pois deixa clara a origem dos dados e costuma usar a escrita hipertextual para mostrar o caminho de onde a informação foi extraída.

4.3.2 O uso de planilhas na organização dos dados

A partir das descrições das notícias e com o intuito de organizar os dados de forma estruturada, de modo que se possa fazer cálculos e cruzamentos, criou-se uma planilha geral tendo como ponto de partida as informações anotadas nas fichas de leitura. Cada uma das 315 fontes⁵⁴ encontradas foi descrita em uma linha, enquanto suas especificidades foram descritas nas colunas (APÊNDICE C).

Figura 14 – Planilha na organização dos dados

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|----|--------------|-------------------------------|-------|-------------------|------------------|----------------------|--------------------|------------------------|---------------|-----------|-------------------|-------------------|---|---|
| 1 | Código geral | Veículo da Unidade/veículo da | Fonte | Descrição fonte | Link (S ou N) | Termo sublinhado | Destino link | Tag | Tag | Tag | Título matéria | Assuntoreditorial | Link | |
| 2 | 0N1 | NK1 | F1 | Dados do Censo | números do censo | 3 milhões de orang | Link para aprese | Governo | Base de dados | Documento | 7 dados que mo | Educação | https://www.inec.gov.br/ | |
| 3 | 0N1 | NK1 | - | - | - | anuncio | Link para página | Link de complement | - | - | 7 dados que mo | Educação | https://www.inec.gov.br/ | |
| 4 | 0N1 | NK1 | F2 | Declaração do m | "É impossível ou | apresentou os núm | Link para matéri | Governo | Declaração | Pessoa | 7 dados que mo | Educação | https://www.inec.gov.br/ | |
| 5 | 0N1 | NK1 | - | - | - | dados do censo | Link para aprese | Link de complement | - | - | 7 dados que mo | Educação | https://www.inec.gov.br/ | |
| 6 | 0N1 | NK1 | F3 | Plano Nacional d | informações sob | aprendizado sufiar | Link para site d | Governo | Pesquisa | Documento | 7 dados que mo | Educação | https://www.inec.gov.br/ | |
| 7 | 0N1 | NK1 | F4 | Pesquisa Nacion | relaciona o núme | - | - | Governo | Pesquisa | Documento | 7 dados que mo | Educação | https://www.inec.gov.br/ | |
| 8 | 0N1 | NK1 | F5 | Instituto Paulo M | Dados sobre a p | - | - | Instituições.org | Pesquisa | Documento | 7 dados que mo | Educação | https://www.inec.gov.br/ | |
| 9 | 0N1 | NK1 | - | - | - | das universitários | Link para página | - | - | - | 7 dados que mo | Educação | https://www.inec.gov.br/ | |
| 10 | 0N1 | NK1 | - | - | - | reduzir pela metade | link para o site d | Link de complement | - | - | 7 dados que mo | Educação | https://www.inec.gov.br/ | |
| 11 | 0N1 | NK1 | F6 | Organização par | organização que | países avaliados | Link para relat | Link de complement | - | - | 7 dados que mo | Educação | https://www.inec.gov.br/ | |
| 12 | 0N2 | NK2 | F1 | Twitter (postage) | Uso de postagem | - | - | Instituições.org | Base de dados | Documento | O horário das rel | Cultura | https://www.inec.gov.br/ | |
| 13 | 0N3 | NK3 | F1 | Blog do Mauro B | Conteúdo do fu | estava atrasado | Link para o blog | Instituições.org | Blog | Documento | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 14 | 0N3 | NK3 | F2 | Confederação B | informações sob | a CBF considera | Link para docum | Instituições.org | Resolução | Documento | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 15 | 0N3 | NK3 | F3 | Coluna assinada | informações sob | coluna | Link para coluna | Instituições.org | Blog | Documento | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 16 | 0N3 | NK3 | - | - | - | motivos mencion | Link para docum | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 17 | 0N3 | NK3 | F4 | Fonte ouve | Aparece no h | argumentos favor | Link para estudo | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 18 | 0N3 | NK3 | F5 | Declaração de L | Declaração em d | em 2009 | Link para matéri | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 19 | 0N3 | NK3 | - | - | - | desavergas | Link para matéri | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 20 | 0N3 | NK3 | - | - | - | campeões de 1987 | link para public | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 21 | 0N3 | NK3 | - | - | - | uma decisão do Su | link para matéri | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 22 | 0N3 | NK3 | - | - | - | "dentro de campo" | link para matéri | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 23 | 0N3 | NK3 | - | - | - | levantar dívidas | Link para matéri | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 24 | 0N3 | NK3 | - | - | - | R\$ 233 milhões | Link para public | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 25 | 0N3 | NK3 | - | - | - | "esportização" | Link para blog d | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 26 | 0N3 | NK3 | - | - | - | mudança | Link para o blog | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 27 | 0N3 | NK3 | - | - | - | premição | Link para site d | Link de complemento | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 28 | 0N3 | NK3 | F6 | Revista Forbes | Ranking feito pe | ranking | Link para site d | Instituições.org | Ranking | Documento | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 29 | 0N3 | NK3 | F7 | Globo (emissora) | Questões sobre | "mais valioso ter fu | Link para matéri | Instituições.org | Declaração | Documento | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 30 | 0N3 | NK3 | F8 | Levantamento d | Levantamento s | um levantamento | Link para o site | Instituições.org | Blogsite | Documento | Campeonato Br | Esporte | https://www.inec.gov.br/ | |
| 31 | 0N3 | NK3 | - | - | - | protestos | Link do vídeo do | Link de exemplificação | - | - | Campeonato Br | Esporte | https://www.inec.gov.br/ | |

Fonte: A autora (2018)

A **Área A** da planilha, destacada em roxo na **Figura 14**, é constituída por três colunas de códigos que servem para identificar a notícia, o veículo e a fonte. São eles:

⁵⁴ Mesmo que acionada mais de uma vez em uma mesma notícia, a fonte só foi contabilizada uma única vez neste estudo. Neste trabalho, o interesse está em torno dos tipos de fontes usadas como um todo e quantas fontes diferentes são usadas em um mesmo trabalho de apuração, mas não a quantidade de citações. Além disso, os trechos de escrita hipertextual que não detêm caráter de fonte foram armazenados e serão utilizados em estudos futuros.

- a) **Código geral** – número que indica a unidade de análise referente à notícia no universo global do corpus analisado. Tem como identificação a sigla UN (que remete à unidade noticiosa) e vai do 1 ao 60. Exemplo: UN1 refere-se à primeira notícia da lista.
- b) **Código geral da unidade/veículo** – número que indica a ordem de análise da notícia dentro do seu próprio veículo; leva a sigla de identificação do veículo e um número de 1 a 10. O *Nexo* é representado pelas iniciais *NX*, *O Globo* por *GL*, o *The New York Times* por *NT*, o *FiveThirtyEight* por *FT*, o *La Nación* por *LN* e o *Chequeado* por *CH*. Exemplo: FT4 significa a quarta notícia estudada do *FiveThirtyEight*.
- c) **Código da fonte** – número que indica a ocorrência de fontes apenas dentro da notícia em que ela está inserida. Dessa forma, começa no número 1 e vai até o algarismo necessário para indicar o total de fontes. Exemplo: F64 é a fonte de número 64 na lista final de notícias.

Desta forma, um código denominado UN18 GL8 F71 refere-se à unidade de análise de número 18, oitava de *O Globo*, e fonte de número 71.

A **Área B** da coluna identificada em verde na **Figura 14** é destinada à descrição das características de cada fonte localizada. Duas colunas operam com essa finalidade: a primeira reúne a identificação da fonte acionada e a segunda o detalhamento de atribuição da fonte, ou seja, a informação a ela relacionada. Dessa forma:

- d) **Fonte** – identifica nominalmente a fonte conforme aparece no texto. Por isso, pode menos específica, como “Instituto Paulo Montenegro” e “Plano Nacional de Educação (PNE)” ou mais específica, como “Coluna assinada pelo jornalista esportivo Juca Kfoury e publicada no jornal Folha de S.Paulo”, dependendo da clareza com que é identificada na notícia.
- e) **Descrição fonte** – apresenta as informações atribuídas à fonte na notícia, ou seja, o que ela sustenta ou explica no texto veiculado.

Desta forma, na notícia *Mudança de faixa populacional salva receita do FPM de 13 municípios* (VASCONCELLOS, 2016), de *O Globo*, o nome da fonte aparece como Tesouro Nacional, enquanto a descrição diz respeito à atribuição da fonte na notícia, ou seja, de informações sobre faixas populacionais/valores recebidos no Fundo de Participação dos Municípios (FPM).

A **Área C**, identificada em azul na **Figura 14**, é destinada aos detalhes da escrita hipertextual. As três colunas apresentam informações sobre o uso (ou não) de hiperlinks, a palavra ou frase destacada e, por fim, o destino do link. Sendo assim:

- f) **Link (S ou N)** – aponta a existência ou não de escrita hipertextual junto à fonte identificada;
- g) **Termo sublinhado** – inserção do termo que contém hiperlink. Vale para as fontes que apresentam escrita hipertextual;
- h) **Destino link** – descrição do destino do hiperlink.

Um exemplo diz respeito à notícia do Nexo sobre o Campeonato Brasileiro, em que é disponibilizado o link para documento da Confederação Brasileira de Futebol (CBF). Desta forma, a coluna com o título Link leva o “S” para indicar a existência de hipertexto e o termo/trecho sublinhado, como indica a outra coluna, é “a CBF considera”. A coluna sobre a descrição do link, por sua vez, é preenchida com a descrição “Link para documento em PDF hospedado no site da CBF com uma resolução da presidência para informar quem são considerados os campeões brasileiros”.

A **Área D** da **Figura 14** é composta por três colunas que serviram para descrever, em palavras-chave, os atributos de cada fonte. A teoria revisada até então auxiliou neste processo de pré-classificação, pois foram atribuídas à fonte características de sua natureza, de caráter, de competência, quanto ao tipo ou material e à escrita hipertextual desenvolvida, quando é o caso. Desta maneira:

- i) **Tag 1** – primeira palavra-chave (ou espécie de etiqueta) capaz de descrever a fonte acionada no texto jornalístico;
- j) **Tag 2** – segunda palavra-chave (ou espécie de etiqueta) capaz de descrever a fonte acionada no texto jornalístico;
- k) **Tag 3** – terceira palavra-chave (ou espécie de etiqueta) capaz de descrever a fonte acionada no texto jornalístico. Exemplo: pessoa, documento, oficial, instituição, alternativo, base de dados, blog, legislação, resolução, site, especialista, pesquisa, governo, político, personagem e outros.

Na notícia veiculada pelo *The New York Times* sobre casos de terrorismo (SANGER-KATZ, 2016), por exemplo, a fonte *Global Terrorism Database*, da Universidade de Maryland,

é classificada com as seguintes tags: base de dados (com relação ao formato), instituições/org (de onde as informações são extraídas) e documento (quanto à natureza).

Assim, as três colunas da **Área D**, denominadas *tags* (etiquetas), foram o ponto de partida para a identificação de características em comum nas 60 notícias analisadas, permitindo o posterior agrupamento de acordo com a maior quantidade de ocorrências e, por consequência, gerando padrões. É desta área que derivou a classificação proposta nesta dissertação e que será discutida na próxima seção.

Para finalizar as explicações das organizações dos dados nas planilhas, tem-se, ainda, a **Área E** da **Figura 14**, destacada em amarelo. Essa área, formada por três colunas, tem como principal propósito descrever informações gerais da notícia. A saber:

- l) **Título da matéria** – inclusão do título da notícia analisada;
- m) **Assunto/editoria** – área/setor de cobertura da notícia. Em alguns casos, a editoria ou o tema está evidente da notícia, em outros é necessário atribuir alguma nomenclatura;
- n) **Link** – endereço eletrônico da notícia estudada para consultas posteriores.

Até aqui foram explicados os procedimentos metodológicos que geraram a proposta de uma classificação de tipos de fontes ainda não contemplada pela teoria e gerada especificamente a partir de notícias com técnicas do jornalismo guiado por dados. Pretende-se detalhar, na próxima seção, como ocorreu a aproximação das categorias encontradas e como se chegou à tipologia final.

4.4 Proposta de classificação de tipos de fontes para o jornalismo guiado por dados

O trabalho de descrição e organização do material empírico em fichas e planilhas permitiu, ao final do processo, a criação de uma classificação ligada aos **tipos** de fontes comumente acionados em notícias que se utilizam de técnicas de jornalismo guiado por dados. A **Área D** da **Figura 14** revelou-se fundamental para essa finalidade. Após a primeira estruturação dos dados, observou-se a recorrência de alguns padrões e regularidades no corpus estudado. As descrições atribuídas às fontes, em especial nos campos destinados às etiquetas (*tags*), começaram a se repetir. Sob a inspiração da sistematização de categorias proposta pela análise de conteúdo (BARDIN, 2016), ocorreu o refinamento por meio de características similares entre as fontes.

O **tipo** diz respeito às características de arquivo ou suporte em que houve a disponibilidade do material consultado ou da pessoa ouvida. O quadro a seguir mostra como funcionou o processo de aproximação e união dos tipos encontrados na análise preliminar em direção à listagem final dos tipos, classificando as fontes encontradas nos produtos de jornalismo guiado por dados.

Quadro 6 – Das palavras-chave aos agrupamentos

| Palavras-chave encontradas | Agrupamentos em tipos de fontes |
|---|---|
| Análise Assessoria Bases de dados Bases de dados própria Comunicado Comunicado Declaração Declaração Decreto Deputado Documento Documento Especialista Especialista Governo Nota Nota Nota técnica Pesquisa Ranking Relatório Relatório Representante Resolução Site/blog | Arquivo documental Estatística Personagem Programa Reprodução |

Fonte: A autora (2018)

A coluna da esquerda referente ao Quadro 6 reúne a lista inicial de *tags* atribuídas às fontes: são 19 tipos diferentes identificados em um primeiro momento. Em alguns casos, as palavras-chaves se repetem, pois podem ter aproximação com mais de uma categoria. A coluna da direita reúne os tipos já agrupados, a partir das características encontradas de forma comum entre os elementos da coluna da direita.

Em ambas as colunas, as cores que identificam os termos representam a sua natureza. As palavras destacadas em **azul** no Quadro 6 referem-se às fontes **documentais**; os termos coloridos em **vermelho** representam as fontes **pessoais**. Já as palavras escritas em **verde** dizem respeito aos termos híbridos, que reúnem tanto fontes pessoais quanto documentais. Dessa

forma, a categoria **arquivo documental** é exclusivamente documental, sem a interferência de fontes humanas. A categoria **estatística** apresenta formato híbrido por conta da palavra-chave especialista, que é pessoal; **reprodução**, por sua vez, também é híbrida, pois conta com fontes de natureza pessoal, documental e híbrida em sua composição. **Personagem** é uma subdivisão formada exclusivamente por fontes pessoais, pois envolve testemunhos de pessoas que contaram suas histórias. Por fim, a categoria **programa** é composta somente por fontes de natureza documental.

Assim, as categorias identificadas até aqui são documentais, pessoais ou híbridas, compostas por palavras-chave que igualmente podem fazer parte de uma dessas definições. Antes de definir as especificidades dos tipos propostos nesta dissertação, são apresentadas as nomenclaturas inicialmente utilizadas.

4.4.1 Classificação inicial incluindo fontes pessoais, documentais e híbridas

Inicialmente, as 315 fontes encontradas no levantamento foram classificadas em 19 tipos, que incluem fontes pessoais, documentais e híbridas. A lista das fontes (APÊNDICE B) foi elaborada com base nas palavras-chave localizadas no *corpus* (Quadro 6).

4.4.2 Categorias preliminares

A partir da proximidade conceitual e de características, as categorias foram agrupadas em cinco tipos, de modo que seja operacionalmente mais viável.

- a) **Arquivo documental** – trata-se da categoria que envolve o uso de documentos como elemento autêntico de informação, de modo que o arquivo vire uma espécie de atestado verídico sobre um fato. Esse tipo abrange fontes documentais as mais variadas, como notas, memorandos, relatórios, decretos, *rankings*, resoluções e leis. Por vezes são disponibilizadas também em hipertexto. Entre os exemplos encontrados estão um pedido feito pelo Conselho Federal do Meio Ambiente e reproduzido pelo *Chequeado* em notícia sobre o cumprimento da Lei de Florestas (MIRANDE, 2016); uma cartilha de adoção intitulada *Child Welfare Information Gateway* do governo dos Estados Unidos, reproduzida em uma notícia do *The New York Times* sobre o mapeamento de casamentos gays no país (BUI, 2016b); e o Atlas dos Remanescentes Florestais da Mata Atlântica, utilizado como fonte pelo *Nexo* para a construção de uma notícia sobre a área restante da mata no Brasil (ALMEIDA, 2017).

- b) **Estatística** – compreende o uso de números para descrever um fenômeno. Em geral, as estatísticas empregadas no conteúdo jornalístico provêm de dados estruturados em planilhas simples ou mais complexas, como bases de dados. Também podem ser provenientes de relatórios, cartas de conjuntura com dados, mapas, pesquisas e outros. Optou-se por utilizar esse nome por conta da dificuldade de visualizar, no produto final, a origem dos dados – isto é, se viraram notícia a partir de estruturação em planilhas ou relatórios, por exemplo. Em geral são fontes documentais. Entre os exemplos localizados no corpus estudado estão o uso de uma base de dados sobre terrorismo, vinculada à Universidade de Maryland, em notícia publicada no *The New York Times* (SANGER-KATZ, 2016); dados do censo da Educação no Brasil, publicados em notícia do *Nexo* (DIAS, 2016); e números divulgados pela Organização para o Desenvolvimento Econômico e de Cooperação (OECD, na sigla em inglês) sobre desigualdade de gênero em termos de salário, em publicação no *FiveThirtyEight* (CASTEEL; CHIN, 2017).
- c) **Personagem** – reúne pessoas ouvidas como testemunhas ou fonte de relato da própria história. Entre os exemplos estão a história de Cristian Bebebino, paciente de um hospital à espera de transplante, ouvido em notícia do *La Nación* sobre transplantes de rim e fígado (PALLARO, 2017); a história de Maria Eugenia Mayorga, empreendedora na Nicarágua, em publicação do *FiveThirtyEight* sobre a economia daquele país (HE, 2016); e outra personagem ouvida como fonte é Laurel Yoeder, que é enfermeira e doula, na publicação do *The New York Times* sobre mudanças de carreira (MILLER; BUI, 2017).
- d) **Programa** – esse tipo abrange informações gerais sobre programas, projetos e iniciativas cujas informações que servem de base para a notícia fogem às estatísticas. Costumam ser informações genéricas e descritivas, em geral fontes documentais, sem o uso de dados numéricos. Entre os exemplos encontrados estão informações sobre o Plano Nacional de Educação (PNE), publicadas em notícia do *Nexo* sobre índices econômicos no Brasil (MARIANI; LUPION; ALMEIDA, 2016); um documento explicativo sobre o programa estadunidense *Food and Nutrition Service*, publicado pelo *The New York Times* em uma notícia sobre obesidade e um programa de governo (CARROLL, 2016); e um documento com perguntas e respostas, com hiperlink para o site, sobre a marcha *A Day Without a Woman*, em publicação do *FiveThirtyEight* (CASTEEL; CHIN, 2017).

- e) **Reprodução** – envolve declarações emitidas por pessoas e organizações, podendo ter natureza pessoal ou documental. Publicações de artigos de opinião são consideradas reprodução. Falas transcritas a partir de diálogo direto com a fonte ou em coletivas de imprensa também fazem parte desse grupo. Nesta categoria também estão incluídas publicações de outros veículos de comunicação e postagens em sites de redes sociais. Entre os exemplos encontrados no corpus estudado estão a fala de John Carnevale, especialista em política de drogas, em uma notícia do *FiveThirtyEight* sobre a falta de dados relacionados ao consumo (CASTEEL, 2017); fala do diretor de Consumidores e Administradores da Defensoria do Povo, Gerardo Gullelmo, em notícia publicada pelo *La Nación* sobre reclamações nos serviços de telefonia (MUSSE, 2017); e notícia publicada pelo *The Japan Times* e citada pelo *Nexo* para reproduzir informações do Ministério da Agricultura do Japão. A publicação do *Nexo* tem a ver com explicações de por que aquele país tem as menores taxas de obesidade do mundo (ORENSTEIN, 2017).

4.4.3 Categorias refinadas: filtro para fontes documentais

As categorizações propostas até o momento incluíram fontes de uso documental e pessoal. Considerando que as fontes pessoais já foram exploradas nos estudos de jornalismo, como visto no capítulo 2, optou-se por isolar as fontes de natureza documental e considerar apenas esses registros para a classificação de fontes, não considerando fontes de origem pessoal ou híbrida (pessoa e documento). Nas 60 notícias analisadas, foram encontradas 213 fontes de natureza documental. Elas podem ser enquadradas em três tipos: **arquivo documental**, **estatística** ou **reprodução**. A categoria personagem é removida porque envolve apenas pessoas, que não são o foco deste estudo. Já a categoria programa, que contava apenas com quatro casos, pode ser integrada ao arquivo documental sem sofrer prejuízos.

A seguir, os três tipos serão detalhados e discutidos. Além da informação sobre a autoria, evidenciada entre parênteses, serão indicados, entre colchetes, os códigos gerais da unidade de cada um, de modo que seja mais rápida a sua localização (APÊNDICE C), que compreende a planilha geral das notícias.

4.4.3.1 Arquivo documental

Envolve fontes documentais as mais variadas, como notas, relatórios, decretos, memorandos, rankings, estudos, resoluções, leis e informações gerais sobre programas, projetos e iniciativas. Abrange, desta forma, um grande número de documentos. O nome da categoria é atribuído à utilização do arquivo como uma espécie de atestado verídico sobre um fato. Não raro são disponibilizados em formato hipertextual e na íntegra, de modo que a origem das informações obtidas esteja clara nas notícias.

Entre os exemplos classificados como tal estão o documento emitido pela chefia de Gabinete de ministros da Argentina por conta de mudanças orçamentárias com relação à Lei de Florestas [UN50 CH10 F265] em notícia publicada pelo *Chequeado* (MIRANDE, 2016); uma dissertação de mestrado sobre microcrédito e desenvolvimento econômico [UN32 FT2 F177], que auxiliou o *FiveThirtyEight* no processo de contextualização da situação econômica da Nicarágua em relação às mulheres empreendedoras; e a nota da presidência da República [UN53 LN3 F277], reproduzida na íntegra em notícia do *La Nación* sobre os argentinos que detêm sociedades offshore, como revelou o processo de investigação conhecido como *Panama Papers* (RUIZ, JASTREBLANKSY e MON, 2016).

4.4.3.2 Estatística

Compreende o uso de números para descrever um fenômeno. Em geral, provêm de informações estruturadas em planilhas simples ou mais complexas, como as bases de dados disponíveis em portais da transparência. Números extraídos de pesquisas sociais também integram essa categoria, que apresenta abordagem quantitativa. As estatísticas fazem parte de materiais consultados pelos jornalistas sem a necessidade de entrevista. O termo foi escolhido no lugar de bases de dados, por exemplo, por conta da dificuldade de observar, no produto final, qual o tipo de arquivo utilizado.

Esse é o tipo que apresenta a maior quantidade de exemplos em todo o corpus estudado. Entre os exemplos estão a pesquisa de salário e pobreza nos Estados Unidos [UN35 FT5 F199], elaborada pelo Census Data, em notícia publicada pelo *FiveThirtyEight* sobre o protesto de mulheres nos anos de 1970 (CASTEEL; CHIN, 2017); estudo sobre crimes em 2016 [UN29 NNT9 F148], utilizado pelo *The New York Times* em notícia sobre o algoritmo utilizado pela polícia de Chicago para tentar antecipar crimes (ASHER; ARTHUR, 2017); e a Corte Suprema

Argentina [UN55 LN5 F284], que foi acionada como fonte pelo *La Nación* por conta dos registros de violência contra a mulher.

4.4.3.3 Reprodução

Reúne comunicados ou declarações emitidas por pessoas e organizações e que são posteriormente consultadas como registro, por isso não ocorre por meio de entrevista direta. Publicações de artigos de opinião e outros textos autorais, como postagens em blogs, são consideradas reprodução. Nesta categoria também estão incluídas publicações de outros veículos de comunicação e postagens em sites de redes sociais que, sem o caráter de entrevista, recebem status de documento quando acessado pelos jornalistas. Pressupõe maior mediação do que os outros tipos.

São considerados exemplos de fontes do tipo reprodução trechos de entrevistas publicadas em outros veículos de comunicação, como fez o *Nexo* na notícia sobre o Campeonato Brasileiro (PIMENTEL, 2017); também utilizada pelo *Nexo*, uma reportagem do jornal *The Washington Post* [UN8 NX8 F56] sobre a situação do Japão após a Segunda Guerra Mundial, em notícia que trata dos baixos índices de obesidade naquele país (ORENSTEIN, 2017); e, por fim, outro exemplo a ser considerado neste tipo é de uma notícia de *O Globo* sobre gastos de vice-governadores (COUTO, 2017), em que os comunicados emitidos por assessorias de imprensa, como a dos políticos Renato Santana e Rose Modesto [UN15 GL5] são declarações.

Esta proposta de categorização de fontes encontradas no jornalismo guiado por dados parte da premissa de oferecer uma contribuição inédita e sustentada pelo diálogo entre a teoria e a análise empírica. Assim, não explora com detalhes classificações já existentes, ainda que admita a possibilidade de consultas cruzadas. Significa, por exemplo, que uma fonte pode ser do tipo **estatística** e apresentar caráter oficial ou não oficial, estar ligada ou não a um órgão público ou a uma instituição e apresentar o uso de escrita hipertextual remetendo à origem dos dados.

4.5 Fontes em números: os resultados globais da pesquisa

As 60 notícias estudadas apresentam 315 fontes. Dessas, 213 (67,62%) são **documentos** e 102 (32,38%) são **pessoas**. Esta é uma amostra não probabilística, ou seja, não possui significância estatística. Dessa forma, os números aqui apresentados são a título de curiosidade

e não serão explorados em exaustão de forma quantitativa, pois não permitem fazer inferências generalizantes, que se repetiriam em outras amostras.

Tabela 1 – Ocorrências de fontes por veículo

| Veículo | Total de ocorrências | Percentual* | Documento | %** | Pessoa | %** |
|---------------------------|----------------------|-------------|-----------|--------|--------|--------|
| <i>O Globo</i> | 28 | 8,89% | 21 | 75,00% | 7 | 25,00% |
| <i>Nexo Jornal</i> | 49 | 15,56% | 44 | 89,80% | 5 | 10,20% |
| <i>The New York Times</i> | 82 | 26,03% | 53 | 64,63% | 29 | 35,36% |
| <i>FiveThirtyEight</i> | 62 | 19,68% | 41 | 66,13% | 21 | 33,87% |
| <i>La Nación</i> | 43 | 13,65% | 19 | 44,19% | 24 | 55,81% |
| <i>Chequeado</i> | 51 | 16,19% | 35 | 68,63% | 16 | 31,37% |
| Total | 315 | 100% | 213 | 67,62% | 102 | 32,38% |

* Percentual em relação ao apanhado global de veículos

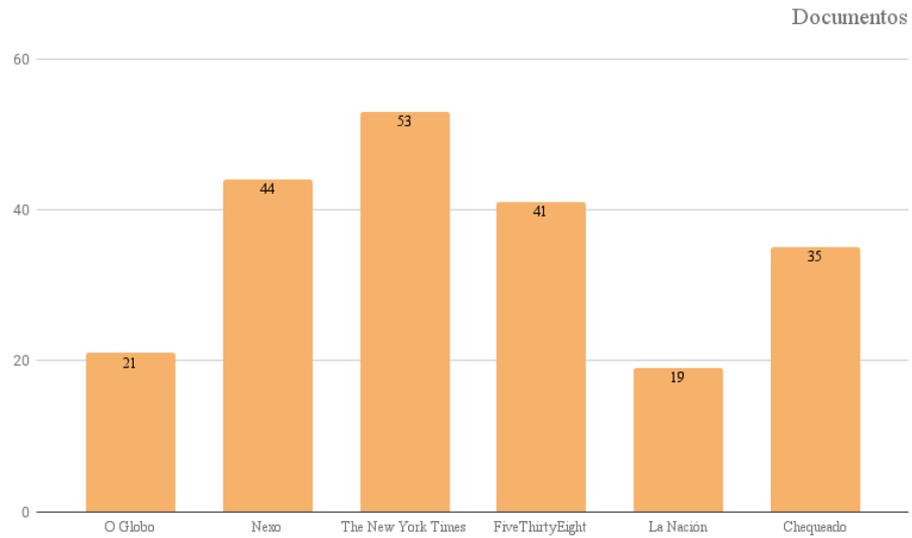
** Percentual de representatividade dentro do próprio veículo

Fonte: A autora (2018)

Nota-se, pela leitura da Tabela 1, que o jornal *The New York Times* é o veículo que apresenta o maior número de fontes: são 82 (26,03%) e o *FiveThirtyEight* aparece logo na sequência, com 62 (19,68%). Isso faz com que, juntos, os veículos dos Estados Unidos concentrem quase metade do total de fontes acionadas nas 60 notícias, com 45,71% do total. Os veículos argentinos, por sua vez, reúnem, juntos, 94 (29,84%) do total de fontes deste levantamento: são 51 (16,19%) no *Chequeado* e 43 (13,65%) no *La Nación*. Já os jornais brasileiros representam 24,44% da fatia total de fontes analisadas nesta pesquisa: 49 (15,56%) fontes aparecem no *Nexo* e 28 (8,89%) em *O Globo*.

Em se tratando de fontes documentais, a maior parte está concentrada no *The New York Times*, com 53 (24,88% das 213) ocorrências, seguido do *Nexo*, com 44 (20,66%), e o *FiveThirtyEight*, com 41 (19,25%) do total. Os demais veículos aparecem em menor escala: o *Chequeado*, com 35 (16,43%), *O Globo*, com 21 (9,86%) e o *La Nación*, com 19 (8,92%). Representadas graficamente, essas ocorrências ficam assim distribuídas:

Figura 15 – Gráfico da distribuição de fontes documentais por veículo



Fonte: A autora (2018).

A seguir, as fontes documentais acionadas pelos veículos serão detalhadas de acordo com os tipos que representam. Fontes repetidas pelo mesmo veículo não são consideradas nos quadros.

4.5.1 *O Globo*

O jornal *O Globo* apresenta um total de 21 fontes documentais, que ao serem listadas na íntegra são 20 inéditas. Ao todo são duas na categoria arquivo documental, 16 em estatística e duas na categoria reprodução. Assim detalhadas:

Quadro 7 – Fontes documentais acionadas em *O Globo*

| O Globo | |
|---------------------------|---|
| Arquivo documental | <ul style="list-style-type: none"> - Ministério da Fazenda – nota com explicações sobre a redução no Fundo de Participação dos Municípios; - Nota oficial do Instituto Nacional de Colonização e Reforma Agrária (Incra) sobre casos de violência. |
| Estatística | <ul style="list-style-type: none"> - Instituto de Segurança Pública do Rio (ISP) – Dados sobre apreensões de armas de fogo; - Dados da plataforma de streaming Spotify; - Departamento Intersindical de Estatística e Estudos Socioeconômicos (Dieese), com dados de greves/paralisações ocorridas no Brasil; - Ligue 180 – dados de casos de violência contra a mulher; - Portal da transparência do governo federal; - Portais da transparência de governos estaduais (a fonte é citada desta forma); - Conselho Nacional de Justiça (CNJ) – Dados sobre quantidades de prisões; - Comissão Pastoral da Terra (CPT) – Dados sobre conflitos de terra no Brasil; - Censo 2010 do Instituto Brasileiro de Geografia e Estatística (IBGE); - Instituto Socioambiental (ISA) – Números de terras indígenas no Brasil; - Relação Anual de Informações Sociais (Rais) 2015 do Ministério do Trabalho; - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) – número de analfabetos; - Levantamento (espécie de mapa da violência) feito por Julio Jacobo Waiselfsz – taxa de mortos por arma de fogo; - Tesouro Nacional – dados de informações sobre faixas populacionais/valores recebidos por municípios via FPM; - Observatório de Informações Municipais – dados de municípios; - Ministério da Cultura – base de dados com informações sobre quantidades de projetos da Lei Rouanet. |
| Reprodução | <ul style="list-style-type: none"> - Assessoria de imprensa do vice-governador Renato Santana; - Assessoria de imprensa da vice-governadora do Mato Grosso do Sul, Rose Modesto; |

Fonte: A autora (2018)

O Quadro 7 permite inferir que *O Globo* dá preferência a bases de dados públicas, especialmente ligadas ao governo federal ou a órgãos de governo, para compor suas notícias a partir de técnicas de jornalismo guiado por dados. Assim, maior parte das fontes está concentrada no quadro de estatística à medida que são consultados portais da transparência para a obtenção de dados numéricos, em especial de investimentos e despesas do poder público.

4.5.2 Nexo

O *Nexo* apresenta um total de 44 fontes documentais, que ao serem listadas na íntegra são 43 inéditas. Ao todo são dez na categoria arquivo documental, 18 na estatística e 15 na categoria reprodução. Assim detalhadas:

Quadro 8 – Fontes documentais acionadas pelo *Nexo*

| Nexo | |
|---------------------------|---|
| Arquivo documental | <ul style="list-style-type: none"> - Plano Nacional de Educação (PNE) e metas cumpridas; - Documento da Confederação Brasileira de Futebol (CBF) sobre os times que considera campeões brasileiros; - Estudo científico sobre a unificação dos títulos de campeonatos brasileiros; - Ranking criado pela Revista Forbes sobre os times mais valiosos; - Informações da Rede Globo sobre preço e horário de exibição do futebol; - Atlas dos Remanescentes Florestais da Mata Atlântica (2015-2016) – SOS Mata Atlântica; - Instituto Nacional de Pesquisas Espaciais (INPE); - Observatório do Plano Nacional de Educação (PNE) – informações sobre o trabalho de acompanhamento de índices do plano; - Estudo publicado no periódico <i>The New England Journal of Medicine</i>; - Estudo científico do <i>The Food Industry Center</i>; |
| Estatística | <ul style="list-style-type: none"> - Dados do Censo da Educação, do Ministério da Educação; - Pesquisa Nacional por Amostra de Domicílios (PNAD) – número de pessoas que não sabem ler nem escrever; - Instituto Paulo Montenegro – número de pessoas que não sabem interpretar textos simples; - Dados da Organização para a Cooperação e o Desenvolvimento Econômico (OECD); - Base de dados do <i>Twitter</i> com postagens sobre horário de refeições; - Levantamento realizado pelo site Verminosos por Futebol com relação à quantidade de público nos jogos; - Levantamento feito pela Revista <i>Época</i> sobre o sistema de liga de futebol no Brasil; - Dados do Departamento Intersindical de Estatística e Estudos Socioeconômicos (Dieese) sobre o salário mínimo; - Pesquisa Mensal de Emprego (PME); - Pesquisa Nacional por Amostra de Domicílios (PNAD); - Base de dados do Banco Mundial – dados econômicos sobre o Brasil; - Comissão Econômica para a América Latina e o Caribe (Cepal) – números sobre casos de extrema pobreza no Brasil; - Organização das Nações Unidas (ONU) – dados sobre o Índice de Desenvolvimento Humano (IDH); - Dados do Atlas da Agropecuária Brasileira sobre a composição de áreas públicas e privadas no Brasil; - Dados do Instituto Nacional de Colonização e Reforma Agrária; - Planilhas do Supremo Tribunal Federal (STF) com dados de votação dos ministros; - Dados da Organização Mundial da Saúde (OMS) sobre IMC; - Números do Atlas da Violência, elaborado pelo Ipea e Fórum Brasileiro de Segurança Pública; - Informações sobre números de investimentos via relatório da Conferência das Nações Unidas sobre Comércio e Desenvolvimento. |

| | |
|-------------------|---|
| Reprodução | <ul style="list-style-type: none"> - Publicação no Blog do Mauro Beting sobre o contexto do futebol brasileiro; - Coluna assinada pelo jornalista Juca Kfourri na <i>Folha de S.Paulo</i>; - Declaração de Luiz Gonzaga Belluzzo, ex-presidente do Palmeiras, à CBF; - Texto de Luis Filipe Chateaubriand defendendo a sincronização de calendários de futebol; - Publicação de relato e entrevista no blog do jornalista Emerson Gonçalves; - Declaração do ex-presidente da CBF, Marco Polo Del Nero, publicada no site Cidade Verde; - Entrevista do ex-jogador Zico publicada pelo jornal O Lance; - Entrevista da professora Silvana Goellner, da UFRGS, à <i>Zero Hora</i>; - Entrevista do vice-presidente de esportes da Turner, Leonardo Lenz Cesar, à <i>Folha de S.Paulo</i>; - Informações do Blog do jornalista Luiz Antônio Prósperi; - Notícia veiculada pela <i>Folha de S.Paulo</i> sobre a desigualdade de renda – contém fala do economista Marcelo Neri; - Notícia veiculada pela <i>Rádio CBN</i>; - Notícia do <i>Japan Times</i> com informações do ministério da Agricultura; - Notícia do jornal <i>The Washington Post</i> sobre o Japão; - Notícia da BBC sobre leis japonesas; |
|-------------------|---|

Fonte: A autora (2018)

É possível perceber, ao analisar o Quadro 8, que o *Nexo* utiliza com frequência informações de bases de dados públicas e de documentos de organizações internacionais como fontes documentais para a composição de seu conteúdo. A listagem maior de fontes está no campo destinado à estatística. Notícias publicadas por outros veículos ou mesmo colunas de opinião e publicações em blogs e sites também acabam acionadas.

4.5.3 *The New York Times*

O *The New York Times* apresenta um total de 53 fontes documentais, que ao serem listadas na íntegra são todas inéditas. Ao todo são 11 na categoria arquivo documental, 41 em estatística e apenas uma na categoria reprodução. Assim detalhadas:

Quadro 9 – Fontes documentais acionadas pelo *The New York Times*

| <i>The New York Times</i> | |
|---------------------------|---|
| Arquivo documental | <ul style="list-style-type: none"> - Boletim sobre alimentos – informações sobre o <i>Food Stamp Program</i> sobre o consumo de alimentos; - Estudo sobre obesidade; - Estudo sobre restrições do SNAP e o impacto que gera sobre a alimentação; - Estudo sobre incentivos e restrições pelo programa; - Documento do USDA sobre as restrições do SNAP; - Documento com explicações sobre o programa <i>Food and Nutrition Service</i>; - Pesquisa do Federal Reserve Bank of Chicago sobre efeitos dos mapas nos anos 1930; - Mapa digitalizado de regiões de Nova York; - Nota do Departamento de Polícia de Chicago; - Cartilha de adoção criada pelo <i>Child Welfare Information Gateway</i>, do governo dos Estados Unidos; - Pesquisa do <i>Pew Research Center</i> que discute os levantamentos sobre casamento gay. |

| | |
|-------------|---|
| Estatística | <ul style="list-style-type: none"> - Pesquisas científicas sobre auxílio financeiro dos pais a pessoas jovens – valores de assistências fornecidas; - Pesquisa <i>Historial Trends in Parental Financial Support of Young Adults</i>, assinada pelo Population Studies Center – números e tendências; - Base de dados do P.S.I.D Transition to Adulthood Supplement; - Census Bureau – divisão responsável pelo Censo nos EUA. - Pesquisa American Community Survey, compilada pelo Census Bureau; - American Factfinder – buscador que permitiu refinar pesquisas e bases de dados do censo; - Pesquisa <i>Current Population Survey's Annual Social and Economic Supplement</i> – números sobre áreas urbanas e rurais; - Relatório do Center on Budget and Policy Priorities – número de residentes em áreas rurais; - Base de dados O*Net Database, ligada ao departamento de Trabalho do governo americano; - Bureau of Labor Statistics – Dados de emprego; - Pesquisa <i>The Future of Employment: how susceptible are jobs to computerisation?</i>, com dados do futuro do trabalho. - Pesquisa <i>The Polarization of Job Opportunities</i>, do Departamento de Trabalho dos Estados Unidos; - Carta de conjuntura intitulada <i>Is there a skills mismatch in the labor Market?</i>, com dados sobre trabalho; - Estudo científico sobre o mercado chinês – números; - Estudo científico sobre situação de trabalho pós layoff - números; - Dois estudos científicos (sem especificar quais) – sustentam que ao final dos 20 anos e até a metade dos 30 é o período em que se acentua a lacuna de pagamentos díspares entre homens e mulheres; - Estudo de pesquisadora de Harvard sobre valores díspares nos salários de homens e mulheres; - Novo estudo (não especifica qual) – dados sobre desigualdade de gênero em salários; - <i>The American Economic Review Paper</i> – documento que traz números de séries históricas sobre questões de emprego; - Duas bases de dados (não especifica quais) – número de ataques terroristas; - Base de dados <i>Global Terrorism Database</i> (Universidade de Maryland); - IHS <i>Jane's Terrorism and Insurgency Center</i> – monitora dados de terrorismo; - Artigo científico sobre índices de mortalidade em hospitais; - Artigo científico sobre os parâmetros e padrões de mortalidade em hospitais; - Estudo científico sobre medidas de qualidade em hospitais; - Estudo científico sobre evidências em cuidados hospitalares – números; - <i>The Supplemental Nutrition Assistance Program (SNAP)</i> – estatísticas do programa; - Estudo científico sobre segurança alimentar e obesidade; - Estudo sobre os beneficiados pelo SNAP; - Estudo recente (não especifica) sobre saúde; - Estudo sobre promoção de saúde por meio de auxílio-alimentação; - Estudo com números sobre as implicações de remover doces e bebidas do SNAP; - Documento do USDA sobre incentivos e evolução de incentivos da saúde. - Mapa digitalizado da Universidade de Richmond com informações sobre desigualdade; - Artigo acadêmico com números sobre desigualdade nos anos de 1930; - Pesquisa sobre propriedade e moradia em Nova York – números; - Planilhas com os números de casos de violência por conta de questões raciais, de propriedade e de moradia; - Estudo com números de crimes violentos em 2016; - Documento do Departamento de Polícia de Chicago com estatísticas sobre criminalidade; - Tabelas do <i>National Center for Health Statistics</i> com informações sobre número de casamentos; - Estudo com informações estatísticas sobre casamentos gay. |
|-------------|---|

| | |
|-------------------|---|
| Reprodução | - Comunicado de que o Departamento de Polícia de Chicago decidiu não comentar sobre a declaração do chefe de Polícia. |
|-------------------|---|

Fonte: A autora (2018)

Uma consulta ao Quadro 9 permite inferir que o jornal utiliza com frequência estudos científicos que dão sustentação às notícias, além de bases de dados de órgãos do governo federal e de organizações internacionais. Dessa forma, maior parte das fontes está concentrada na categoria estatística.

4.5.4 *FiveThirtyEight*

O *FiveThirtyEight* apresenta um total de 41 fontes documentais, que ao serem listadas na íntegra são 38 inéditas. Ao todo são quatro na categoria arquivo documental e 34 em estatística. Assim detalhadas:

Quadro 10 – Fontes documentais acionadas pelo *FiveThirtyEight*

| FiveThirtyEight | |
|---------------------------|---|
| Arquivo documental | <ul style="list-style-type: none"> - Informações sobre a pesquisa nacional sobre uso de drogas (<i>National Survey on Drug Use and Health</i>); - Dissertação de mestrado sobre microcrédito e desenvolvimento econômico; - Documento com perguntas e respostas mais comuns no site <i>Women's March</i>; - Memorando da Casa Branca sobre agências não-governamentais; |

| | |
|-------------|--|
| Estatística | <ul style="list-style-type: none"> - <i>Centers for Disease Control and Prevention</i> – dados do governo sobre uso de drogas; - Dados sobre uso de drogas – coleta e análise pelo <i>The New York Times</i>; - Pesquisa científica com números sobre queda no consumo de cocaína; - Pesquisa científica com números sobre uso de drogas ilegais entre 2000 e 2010; - Pesquisa de monitoramento de dados do <i>Research and Development (Rand)</i>; - Base de dados do sistema de evidência de consumo (<i>System to Retrieve Information from Drug Evidence</i>); - <i>International Labor Organizartion</i> – Números sobre a força de trabalho de mulheres na Nicarágua; - <i>The World Bank Data</i> – Números sobre o Produto Interno Bruto (PIB); - Artigo científico intitulado <i>Aggregate effects of gender gaps in the labor Market: a quantitative estimate</i>, publicado pela Universidade de Chicago; - Boletim Nicaragua’s International Foundation for Global Economic Challenges (Fideg); - World Economic Forum – dados econômicos sobre a Nicarágua; - <i>The United Nations Economic Comission for Latin America and the Caribbean</i> – estimativas de ganhos salariais por homens e mulheres na Nicarágua; - Trabalho de graduação da Universidade Nacional Agrária – dados sobre contribuição de pequenas empresas para a economia da Nicarágua; - <i>Land Administration Project (Prodep)</i> – número de mulheres rurais donas de terra; - <i>Pew Research Center’s Global Attitudes & Trends</i> – pesquisa anual global de opinião pública; - <i>American Community Survey</i> – pesquisa que envolve dados do Censo; - <i>Bureau of Labor Statistics</i> – dados estatísticos; - <i>Census Bureau</i> – estimativas de pagamento para mulheres; - Pesquisa sobre salário e pobreza nos Estados Unidos (Income and poverty in the US: 2105), do Census Data; - <i>Organization for Economic Co-operation and Development (OECD)</i> – números sobre igualdade de gênero; - <i>Guttmacher Institute</i> – Dados sobre aborto; - Gallup survey – números de pesquisa sobre construção de muro na área de fronteira entre EUA e México; - Pew Research Center; - Pesquisa <i>McClatchy-Marist Poll</i> sobre aprovação ou reprovação do governo Trump; - Pesquisa da <i>Quinnipiac University</i> – respostas sobre questões do governo Trump; - Pesquisa American National Election; - Pesquisa conduzida pelo <i>FiveThirtyEight</i> sobre doces preferidos no Halloween; - <i>Equal Employment Opportunity Comission (EEOC)</i>, com dados de situações de assédio sexual em situações de trabalho; - Pesquisa do YouGov sobre casos de assédio em situações de trabalho; - <i>Workplace violence and harrassment of low-wage workers</i>, publicado pelo <i>Berkeley Journal of Employment & Labor Law</i>; - Artigo científico <i>Why Sexual Harassment Complaints Fall on Deaf Ears</i>, sobre casos de assédio para quem recebe baixos salários; - <i>National Women’s Law Center</i>, com informações sobre mulheres em trabalhos de baixa renda; - <i>Dynamics of Collective Action</i> – base de dados sobre protestos; - <i>Survey Monkey Audience</i> – pesquisa do <i>FiveThirtyEight</i> sobre esporte. |
|-------------|--|

Fonte: A autora (2018)

O Quadro 10 permite inferir que o *FiveThirtyEight* dá preferência a bases de dados públicas, especialmente ligadas a organizações internacionais, pesquisas – incluindo levantamentos próprios – e artigos científicos. Por isso, maior parte das fontes está concentrada no tipo estatística.

4.5.5 *La Nación*

O veículo apresenta um total de 19 fontes documentais, que ao serem listadas na íntegra são 18 inéditas. Ao todo são sete registros na categoria arquivo documental, dez em estatística e um na categoria reprodução. Assim detalhadas:

Quadro 11 – Fontes documentais acionadas pelo *La Nación*

| La Nación | |
|---------------------------|--|
| Arquivo documental | <ul style="list-style-type: none"> - Nota da presidência da República sobre caso <i>Panama Papers</i>; - Documento ligado ao presidente no caso <i>Panama Papers</i>; - Vídeo com os discursos de Barack Obama e Donald Trump; - Nota da empresa Movistar sobre reclamações de clientes; - Documento da Assessoria de imprensa do Instituto Nacional Central Único de Ablação e Implante (Incucai); - Lei N° 26.928, sobre lista de espera para órgãos; - Estudo sobre casamentos e divórcios. |
| Estatística | <ul style="list-style-type: none"> - Unidade Fiscal Especializada em Sequestros Extorsivos – dados sobre sequestros; - Registro Nacional de Pessoas – repositório com nomes; - A Casa de Encontro – dados estatísticos sobre a violência contra a mulher; - Corte Suprema Argentina – registros sobre violência contra a mulher; - Ministério da Justiça – registros sobre violência contra a mulher; - Ministério Público Fiscal de Buenos Aires – sistematização de dados sobre violência de gênero; - Defesa do consumidor em Buenos Aires – dados de reclamação; - Base de dados do Instituto Nacional Central Único de Ablação e Implante (Incucai) – números de transplantes de órgãos; - Índice Global de Dados Abertos, ligado à Fundação Open Knowledge; - Associação de Concessionários de Veículos da Argentina (Acara) – números sobre vendas de automóveis. |
| Reprodução | <ul style="list-style-type: none"> - Posicionamento do banco Clariden sobre o ex-ministro da Fazenda; |

Fonte: A autora (2018)

O Quadro 11 permite observar que o jornal argentino valoriza notas e documentos como fontes de embasamento para suas notícias. Em termos de estatística, as informações em geral são extraídas de bases de dados do governo federal ou de outras entidades e associações que divulgam informações estatísticas.

4.5.6 *Chequeado*

O *Chequeado* apresenta um total de 35 fontes documentais, que ao serem listadas na íntegra são 34 inéditas. Ao todo são 17 na categoria arquivo documental e 17 em estatística. Não há registros na categoria reprodução. As fontes são assim detalhadas:

Quadro 12 – Fontes documentais acionadas pelo *Chequeado*

| Chequeado | |
|---------------------------|--|
| Arquivo documental | <ul style="list-style-type: none"> - Documento assinado por Marcos Peña, chefe de gabinete do governo da Argentina, sobre o pagamento de impostos naquele país; - Decreto assinado pela então presidente Cristina Kirchner sobre o não pagamento de impostos; - Nota da presidência sobre questões de processos trabalhistas; - Publicação do ex-ministro de transporte Florencio Randazzo no Twitter com link para a postagem; - Nota oficial da assessoria de imprensa da Superintendência de Riscos de Trabalho; - Nota oficial do Centro de Estudos e Investigações Trabalhistas; - Arquivo do Ministério de Educação e Esportes intitulado “Aprender 2016”, sobre qualidade da informação; - Explicações sobre o Operativo Nacional de Evaluación (ONE) de 2013, sobre níveis em provas; - Documento do Fundo das Nações Unidas para a Infância (Unicef); - Documento oficial assinado pelo presidente da Sancor, sobre atividades no setor leiteiro; - Publicação do ministro da Modernização argentina no Twitter, Andrés Ibarra, sobre o desempenho do País no ranking da transparência; - Observatório eleitoral argentino, com informações sobre regras eleitorais; - Lei de Florestas; - Documento elaborado pela Chefia de Gabinete de ministros da Argentina sobre a lei de florestas; - Informe da assessoria de imprensa do Ministério do Meio Ambiente e Desenvolvimento Sustentável; - Pedido do Conselho Federal do Meio Ambiente; - Informe do Greenpeace. |

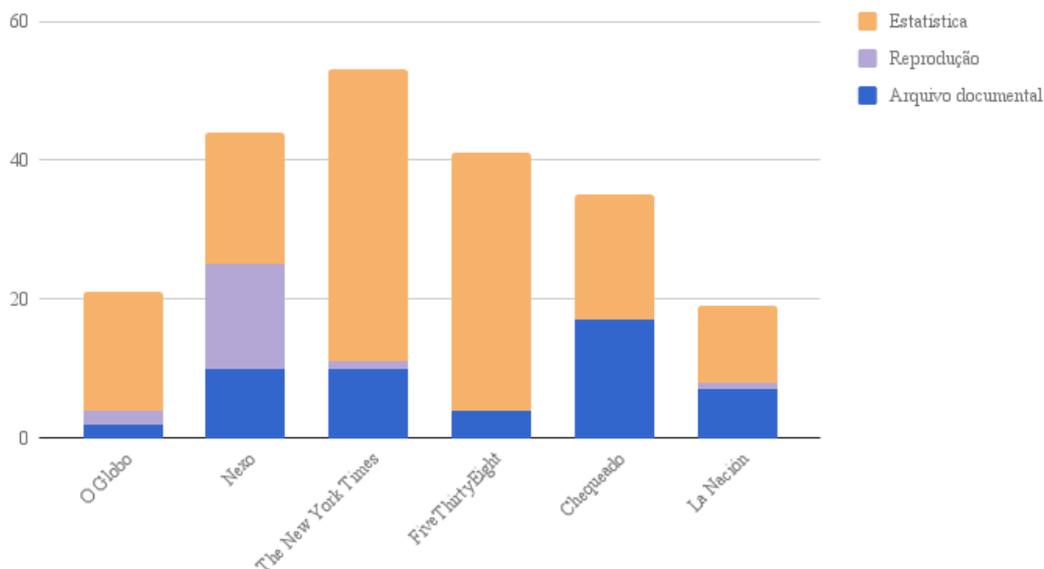
| | |
|--------------------|---|
| Estatística | <ul style="list-style-type: none"> - Chefia de gabinete de ministros na Argentina – números sobre a quantidade de pessoas que pagaram impostos; - Fundo Monetário Internacional (FMI) – Dados de economia sobre a Venezuela; Instituto Nacional de Estatística da Venezuela, com dados sobre inflação; - Superintendência de Riscos de Trabalho – números sobre processos de trabalhos por acidentes ou enfermidades; - União de Seguradoras de Riscos de Trabalho – números sobre notificações e demandas na Argentina; - Instituto Nacional de Estatística e Censo – dados socioeconômicos sobre mercado de trabalho; - Base de dados do Ministério da Educação; - Secretaria da Infância e da Adolescência – número de menores detidos; - Instituto Nacional de Estatística e Censo (Indec) – dados do censo; - Subsecretaria de Leiteria do Ministério da Agricultura – números sobre consumo de leite na Argentina; - Observatório da Cadeia Láctea Argentina (OCLA) – números da produção leiteira; - Direção geral de Estatística e Censo de Buenos Aires – dados sobre preço de leite; - Centro Cifra, com dados de salários de trabalhadores; - Fundação <i>Open Knowledge International</i>, com informações sobre índice global de abertura de dados; - Levantamento do Diretório Legislativo – número de mulheres no parlamento; - Informe da União Interparlamentar sobre o sistema eleitoral e o número de candidatas; - Ministério do Meio Ambiente e Desenvolvimento Sustentável – dados; - Associação Argentina de Orçamento e Administração Pública Financeira. |
|--------------------|---|

Fonte: A autora (2018)

O Quadro 12 permite inferir que o Chequeado utiliza com frequência notas, documentos, trechos de legislação, informes, informações de institutos e observatórios, levantamentos e bases de dados. As fontes são nacionais e internacionais, mostrando uma multiplicidade de origem dos dados.

Em uma representação gráfica, a distribuição quantitativa de fontes por veículo de comunicação pode ser vista da seguinte forma:

Figura 16 – Gráfico da distribuição dos tipos por veículo



Fonte: A autora (2018)

A Figura 16 revela a variedade de fontes documentais acionadas pelos veículos. Mostra que o tipo **estatística** é o mais utilizado. Em números absolutos, o *The New York Times* é o jornal que mais emprega estatística, com 42 ocorrências. Em termos percentuais, o *FiveThirtyEight* é o que mais usa esse recurso: são 37 (90,24%) das 41 fontes documentais acionadas. Na sequência está o jornal *O Globo*, com 17 (80,95%) de 21 casos e o *The New York Times*, com 42 (79,25%) das 53 fontes documentais. A estatística é o tipo que apresenta a maior parcela dos usos de fontes documentais: 144 (67,61%) das 213 fontes documentais encontradas no levantamento. As ocorrências de uso de **arquivo documental** aparecem em menor escala, com 50 (23,47%) casos, e a **reprodução** por último, com apenas 19 (8,92%) ocorrências.

É possível apontar termos conceituais, de modo geral, a partir dos atributos compreendidos em cada tipo de fonte documental. O **arquivo documental** abrange o conteúdo produzido pelo poder público ou instituições e disponibilizado por meio de estudos, notas, memorandos, cartilhas, legislação, planos, projetos, programas, rankings, publicações oficiais, resoluções, publicações editoriais e materiais digitalizados. Em geral, o conteúdo dessas publicações é não numérico ou não desenvolvido sob uma perspectiva analítica, mas serve para a obtenção de informações específicas sobre um programa, uma iniciativa. Documentos vazados e comprovadamente válidos também se enquadram nesta categoria. Tais documentos podem ser apresentados em formato de texto, áudio e vídeo.

A **estatística** é um tipo relacionado principalmente às publicações que levam conteúdo numérico, estatística descritiva, séries históricas, indicadores, taxas, relatórios, cartas de conjuntura e análises. Envolve o uso de bases de dados públicas ou privadas, portais da transparência, dados de pesquisas e estudos científicos, planilhas e levantamentos que geram estatística.

O tipo **reprodução**, por fim, é utilizado para classificar textos e publicações em colunas, blogs, sites de redes sociais outras plataformas. Comunicados e posicionamentos de órgãos do poder público e de instituições, incluindo os que são fornecidos via assessoria de imprensa, estão entre os tipos representados nesta categoria. Também estão incluídas entrevistas e notícias publicadas por veículos de comunicação e citados como fonte.

4.6 Considerações finais do estudo empírico

O cruzamento entre teoria e prática permitiu, após uma série de procedimentos envolvendo leitura e sistematização dos dados, o estabelecimento de características em comum entre as notícias dos seis veículos analisados. A saber:

Quadro 13 – Principais características encontradas no *corpus* estudado

| Veículo jornalístico | País | Principais achados | Tipos acionados |
|-----------------------------|----------------|--|---|
| <i>O Globo</i> | Brasil | Combinação de bases de dados para a criação de taxas e unidades comparáveis; exposição do método utilizado. | Arquivo documental, estatística e reprodução. |
| <i>Nexo Jornal</i> | | Uso de bases de dados, sites de redes sociais e estudos científicos como fonte para fins jornalísticos; técnicas de pesquisa aplicada ao uso de documentos como fonte no jornalismo; aposta em visualizações de dados. | Arquivo documental, estatística e reprodução. |
| <i>The New York Times</i> | Estados Unidos | Emprego de bases de dados e pesquisas científicas; uso acentuado de escrita hipertextual; criação de ferramentas que visam à personalização e utilização por parte do leitor. | Arquivo documental, estatística e reprodução. |
| <i>FiveThirtyEight</i> | | Utilização de notas explicativas no meio do texto, complementando as fontes acionadas; realização de pesquisas com os leitores para posterior uso dos dados em notícias; análise estatística. | Arquivo documental e estatística. |

| | | | |
|------------------|-----------|--|---|
| <i>La Nación</i> | Argentina | Transformação de dados qualitativos em dados numéricos; uso e publicação de documentos vazados envolvendo pessoas públicas; bases de dados e investigação em rede. | Arquivo documental, estatística e reprodução. |
| <i>Chequeado</i> | | Uso de sites de redes sociais como fonte documental; transparência no método a partir da disponibilização das planilhas consultadas e dados extraídos. | Arquivo documental e estatística. |

Fonte: A autora (2018).

O quadro mostra um resumo de como os veículos de comunicação se utilizaram das fontes de informação nas notícias produzidas a partir de técnicas do jornalismo guiado por dados. Praticamente todas as fontes documentais adotadas são inéditas considerando as dez notícias analisadas por veículo. De maneira geral, todos utilizam bases de dados públicas para compor as notícias, em especial repositórios de seus países de origem, como os portais da transparência, e também de organizações internacionais.

Cada um deles adota recursos específicos para oferecer a informação ao leitor. Enquanto *O Globo* priorizou o uso de bases de dados oficiais, o *Nexo* aproveitou o uso de documentos, estudos científicos e dados de sites de redes sociais. O *The New York Times*, por sua vez, também emprega estudos científicos como fonte e costuma trabalhar com a escrita hipertextual ao se referir a algum material complementar. O *FiveThirtyEight*, que ficou conhecido pelo trabalho com estatística, tem o hábito de criar pesquisas por conta e fazer análises em cima das informações obtidas. O *La Nación* transforma documentos vazados em notícia e analisa bases de dados públicas como fonte. O *Chequeado*, por fim, tem como principal característica a transparência no método, deixando clara a origem dos dados trabalhados.

Entende-se, ao final do levantamento, que as 213 fontes documentais localizadas nas 60 notícias estudadas apresentam características em comum que permitem agrupá-las em: **arquivo documental, estatística e reprodução**. Essa é a proposta de classificação de fontes documentais sugeridas nesta dissertação.

5 CONSIDERAÇÕES FINAIS

O cenário da produção jornalística está em transformação no que se refere ao uso de fontes. O acesso à transparência pública e a decorrente ampliação da disponibilidade de bases de dados e documentos, em um ambiente *big data*, fornecem subsídios para a construção de conteúdo baseado em dados e não apenas em declarações oficiais.

Em uma sociedade que opera em rede e com práticas profissionais que funcionam da mesma forma, alguns procedimentos dão abertura a outros e exigem conhecimentos específicos, a exemplo das demandas do jornalismo guiado por dados. A partir de recursos de *software*, proximidade com o rigor metodológico e conhecimento em ciências sociais, o jornalista se apropria de recursos que geram abordagens diferentes junto às fontes pessoais, que não podem simplesmente filtrar a sua versão dos fatos e ignorar os dados de transparência existentes.

Esta dissertação discutiu o uso de fontes documentais em notícias elaboradas a partir de técnicas de jornalismo guiado por dados. No capítulo 2, intitulado **Fontes: a origem da informação no jornalismo**, a revisão bibliográfica trouxe a abordagem de estudos em jornalismo principalmente da década de 1970 até os dias de hoje. Revelando que a preocupação com o uso de fontes e sua credibilidade sempre foi considerada importante na prática profissional, constatou-se que esses estudos enfatizaram muito a perspectiva de fontes pessoais, ainda que citem produções documentais como fonte.

O capítulo 3, denominado **A sociedade em rede, o fenômeno *big data* e o jornalismo**, levantou aspectos sobre as transformações sociais dos últimos anos, modificada pelo aumento na disponibilidade de informações e motivada pela proliferação da tecnologia, em especial por conta dos computadores e do acesso à internet comercial, na metade dos anos de 1990. Esse cenário contribuiu para que práticas jornalísticas mais específicas ganhassem espaço ou fossem aprimoradas, como é o caso do jornalismo guiado por dados. Ainda que suas raízes sejam anteriores até mesmo à inserção de computadores nas redações, o aumento na disponibilidade de bases de dados públicas é que permite maior ênfase nesta prática, atualmente.

Amparado pelas teorias que envolvem as fontes jornalísticas e em um cenário de transformações da sociedade em rede e da prática que se executa em meio a esse contexto, o Capítulo 4, intitulado **Estudo de fontes**, tem abordagem empírica. A partir da leitura e descrição de 60 notícias de jornalismo guiado por dados veiculadas nos jornais *O Globo*, *The New York Times*, *La Nación*, *Nexo Jornal*, *FiveThirtyEight* e *Chequeado*, identificou-se a utilização de 315 fontes e criou-se uma proposta de classificação de tipos para as fontes documentais, que somam 213 ocorrências. São eles: **arquivo documental, estatística e reprodução**.

Acredita-se que esta classificação possa se unir a uma lista de abordagens de teóricos que já discutiram o tema. Essa construção reflete o movimento de fechamento da dissertação, em que são resgatados aspectos teóricos combinados à análise empírica.

Quadro 14 – As fontes documentais a partir da teoria

| Autor | Proposta que leva em conta documento |
|--------------------------------------|---|
| Sigal (1973 apud Santos, 1997) | Canais de rotina, informal e de iniciativa. Os canais oficiais contam com releases, relatórios , discursos, cerimônias; os de iniciativa, por sua vez, incluem eventos testemunhados pelos repórteres, pesquisa envolvendo livros, análise estatística e algo feito pelos próprios jornalistas. |
| Ericson, Baranek e Chan (1987) | Canais altamente estruturados, menos estruturados e não estruturados, que variam de acordo com o acesso dos jornalistas a pessoas e documentos para obter informações. |
| Villafañe, Bustamante e Prado (1987) | Das 13 categorias propostas para classificar fontes de informação e procedência, só uma, denominada documentação e agenda do meio , é que leva em conta claramente o uso de documentos. Está ligada, segundo os autores, a fontes institucionalizadas, permanentes ou estáveis de caráter oficial ou institucional. |
| Erbolato (1991) | Informantes diretos, indiretos e adicionais. Entre os indiretos estão conteúdos de caráter documental publicados anteriormente. Livros, almanaques, enciclopédias e outros estão entre os exemplos citados dentro das fontes adicionais. |
| Pinto (2000) | Classifica as fontes, entre outras possibilidades, conforme a sua natureza: pessoal ou documental . |
| Alcântara, Chaparro e Garcia (2005) | Incluem em sua classificação as fontes documentais e as fontes bibliográficas , que não são muito detalhadas. |
| Jorge (2008) | Denomina de fontes secundárias formas de obter informações via relatórios, livros, revistas, recortes, lista telefônica, fitas e filmes e a internet como um todo. |
| Chaparro (2009) | Considera em sua classificação as fontes documentais , que incluem estudos, diagnósticos e avaliações , e fontes bibliográficas , que congregam saber científico, tecnológico, cultural e popular. |
| Schmitz (2011) | Leva em conta em sua classificação o que chama de fontes de referência , que incluem documentos, livros, artigos, teses, portais, sites e blogs . |
| Silva e Maia (2011) | Consideram fontes de segunda mão: agências de notícias, outros veículos jornalísticos, publicações científicas, documentos impressos e eletrônicos, ciberespaço, reedição e republicação . |
| Folha de S.Paulo (2010) | Trabalha com sua própria classificação de fontes, entre as quais aparece a fonte tipo zero , considerada a mais confiável, que inclui tradição em exatidão, como enciclopédias e documentos emitidos por instituições credíveis. |
| Gehrke (2018) - Dissertação | Arquivo documental, estatística e reprodução são tipos que reúnem fontes documentais identificadas em notícias que tiveram a utilização de técnicas de jornalismo guiado por dados. Leva em conta bases de dados, publicações em sites de redes sociais, pesquisas científicas, publicações editoriais, documentos vazados e geração própria de dados. |

Fonte: A autora (2018)

A proposta de inserção dos resultados desta dissertação no quadro teórico provém de um primeiro estudo, ou seja, de uma primeira tentativa de classificação, aberta a discussões e

processos de refinamento. Alguns aspectos das categorias propostas podem estar localizados em áreas borradas em se tratando de fronteiras conceituais. Essa é uma dificuldade com que as pesquisas envolvendo classificação se deparam. Vale, no entanto, como esforço inicial e de reflexão a partir dos estudos teórico e empírico realizados nesta dissertação.

O quadro teórico de classificações documentais (Quadro 12) mostra que esse tipo de fonte aparece nos estudos de jornalismo desde a década de 1970 até hoje, mas à época em menor escala e desde os anos 2000 de forma mais recorrente, porém não aprofundada. A chegada dos computadores às redações, seguida da internet comercial, ajudam a entender, em termos de suporte e disponibilidade, que o volume de documentos oferecidos para consulta dos jornalistas aumentou significativamente com a difusão de iniciativas de transparência do poder público e leis de acesso à informação. Ao mesmo tempo, a tecnologia facilitou a consulta de jornalistas a fontes pessoais, que podem ser entrevistadas por telefone, e-mail e ferramentas de comunicação instantânea.

Ainda que o volume de materiais e a facilidade de acesso a documentos via internet possam ter levado os teóricos a refletir com mais intensidade sobre o uso de fontes documentais no jornalismo, são raros os estudos aprofundados, que envolvem observação participante em redações ou mesmo uma análise densa de produto das quais emergem categorias e tipologias de fontes. Entre as bibliografias consultadas, a classificação de fontes costuma ocorrer de forma breve, com uma rápida apresentação e exemplificação. O contato humano entre jornalistas e fontes em geral é priorizado nessas abordagens e os documentos não protagonizam análises, mas surgem nas classificações como algo complementar às fontes pessoais.

Outro ponto é que geralmente as fontes documentais são chamadas de fontes secundárias ou de segunda mão, em geral porque pressupõem maior mediação ou filtro editorial antes da publicação. No entanto, entende-se que associar o uso de documentos a algo secundário já não faz muito sentido neste contexto, que têm nos documentos bases fundamentais e como principal fonte para o desenvolvimento de conteúdo jornalístico. Todos os estudos que discutem o emprego de fontes pessoais são importantes e descrevem a realidade de uma época, mas valorizar o uso de documentos como fontes no jornalismo é fundamental em um contexto em que essas consultas são cada vez mais recorrentes.

Se a informação é a matéria-prima na sociedade em rede, aliada ao contexto tecnológico na difusão dos computadores e da internet, é natural que os jornalistas utilizem as bases de dados como fonte. As próprias leis de acesso à informação e transparência, que igualmente são acionadas para a obtenção de informações, podem ser problematizadas enquanto práticas decorrentes de uma sociedade em rede e de conhecimento aberto com potencial para

mobilização de pessoas, tal qual o cenário descrito nesta dissertação. À medida que a sociedade em rede está conectada no acesso às bases públicas e de olho nas produções jornalísticas, há um processo de transformação.

Acredita-se, por fim, que este trabalho pode ser um ponto de partida para discussões mais incisivas sobre o uso de documentos como fontes pelo jornalismo. Trata-se de uma primeira contribuição em forma de classificação. A teoria consultada forneceu todo o embasamento necessário, inclusive para a criação de categorias, mas o empírico teve papel fundamental para se verificar, de fato, que tipos de fontes são empregadas a partir de técnicas do jornalismo guiado por dados. Em termos práticos, pode gerar contribuições para o ensino do JGD na elaboração de notícias e para a compreensão do conteúdo jornalístico já produzido.

REFERÊNCIAS

2017 PULITZER PRIZES. Journalism. Nova York: **Pulitzer Prizes**, 2017. Disponível em: <<http://www.pulitzer.org/prize-winners-by-year/2017>>. Acesso em 13 nov. 2017.

ABAD, Jimena. Entrevista a Florencia Coelho – Periodismo de datos: informar em la era digital. **Revista Dixit**, N. 22, janeiro-julho 2015. Montevideu: Universidade Católica do Uruguai, 2015.

AGÊNCIA PÚBLICA. O mapa do jornalismo independente, 2016. Disponível em: <<http://apublica.org/mapa-do-jornalismo/>>. Acesso em: 15 jun. 2017.

ALCÂNTARA, Norma S.; CHAPARRO, Manuel.; GARCIA, Wilson. **Imprensa na berlinda**: a fonte pergunta. São Paulo: Celebris, 2005.

ALLEGRO, Gabriela. Novidade no Ar. **Blog Na Base dos Dados**. Rio de Janeiro: O Globo, 2014. Disponível em: <<http://blogs.oglobo.globo.com/na-base-dos-dados/post/novidade-no-ar-546916.html>>. Acesso em: 15 jun. 2017.

ALMEIDA, Rodolfo; MARIANI, Daniel. Públicas e privadas: a divisão de terras no território brasileiro. São Paulo: **Nexo**, 2017. Disponível em: <<https://www.nexojournal.com.br/grafico/2017/04/07/P%C3%BAblicas-e-privadas-a-divis%C3%A3o-de-terras-no-territ%C3%B3rio-brasileiro>>. Acesso em: 15 jun. 2017

ALMEIDA, Rodolfo. O que resta da mata atlântica no Brasil. São Paulo: **Nexo**, 2017. Disponível em: <<https://www.nexojournal.com.br/grafico/2017/06/14/O-que-resta-da-mata-atl%C3%A2ntica-no-Brasil>>. Acesso em: 15 jun. 2017.

AMAYA, Sol; PALLARO, Bianca. Radiografía de los secuestros exprés: cuándo y dónde se producen y cómo hay que actuar. Buenos Aires: **La Nación**, 2017. Disponível em: <<http://www.lanacion.com.ar/1997361-radiografia-de-los-secuestros-expres-cuando-y-donde-se-producen-y-como-hay-que-actuar>>. Acesso em: 30 out 2017.

ANDERSON, C. W.; BELL, E.; SHIRKY, C. Jornalismo pós-industrial: adaptação aos novos tempos. **Revista de jornalismo ESPM**. São Paulo, Ano 2, N.5, Abr. Maio. Jun. 2013. p. 30-89.

ARES, Carlos. El periódico conservador argentino 'La Nación' ha cumplido 115 años. [S.l.]: **El País**, 1985. Disponível em: <https://elpais.com/diario/1985/01/06/sociedad/473814008_850215.html>. Acesso em 13 nov. 2017.

ASHER, Jeff; ARTHUR, Rob. Inside the algorithm that tries to predict gun violence in Chicago. Nova York: **The New York Times**, 2017. Disponível em: <<https://www.nytimes.com/2017/06/13/upshot/what-an-algorithm-reveals-about-life-on-chicagos-high-risk-list.html?rref=collection%2Fsectioncollection%2Fupshot>>. Acesso em 30 out. 2017.

ATWOOD, Roy; BEER, Arnold. The roots of academic news research: Tobias Peucer's 'De relationibus novellis' (1690). Routledge - Taylor and Francis Group, **Journalism Studies**, V. 2, N. 4, 2001 p. 485-496.

BAACK, Stefan. A new style of news reporting: Wikileaks and Data-driven Journalism. **Cyber Subjects**. [S.l.]: Social Science Open Access Repository, 2011. Disponível em: <<http://www.ssoar.info/ssoar/handle/document/40025>>. Acesso em 14 nov. 2017.

BADGER, Emily. How redlining's racist effects lasted for decades. Nova York: **The New York Times**, 2017. Disponível em: <<https://www.nytimes.com/2017/08/24/upshot/how-redlinings-racist-effects-lasting-for-decades.html?rref=collection%2Fsectioncollection%2Fupshot>>. Acesso em 30 out. 2017.

BARBOSA, Suzana. Jornalismo digital em base de dados (JDBD): um paradigma para produtos jornalísticos digitais dinâmicos. Tese (Doutorado em Comunicação e Cultura Contemporâneas), **UFBA**, Salvador, 2007.

BARBOSA, Suzana; TORRES, Vitor. O paradigma 'Jornalismo Digital em Base de Dados': modos de narrar, formatos e visualização para conteúdos. São Paulo: **Galáxia** (Online), N. 25, 2013. p. 152-164.

BARDIN, Laurence. **Análise de conteúdo**. São Paulo: Edições 70, 2016.

BELTRÃO, Luiz. **A imprensa informativa**. São Paulo: Editor Folco Masucci, 1969.

BOUNEGRU, Liliana. Data journalism and Computer-Assisted Reporting. In: GRAY, Jonathan; BOUNEGRU, Liliana; CHAMBERS, Lucy (Org.). **The data journalism handbook: how journalists can use data to improve the news**. Sebastopol: O'Reilly, 2012. p. 21.

BOYD, Danah; CRAWFORD, Kate. Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. **Information, Communication & Society**, V. 15, N. 5. [S.l.]: Routledge – Taylor and Francis Group, 2012. p. 662-679. Disponível em: <<http://www.tandfonline.com/doi/pdf/10.1080/1369118X.2012.678878>>. Acesso em 1 fev. 2017.

BRADSHAW, Paul. The inverted pyramid of data journalism. 2011. **Online Journalism**. Disponível em: <<https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/2011>>. Acesso em 10 jan. 2017.

BRADSHAW, Paul. **The online journalism handbook**. 2. ed. Abingdon e Nova York: Routledge, 2017.

BRADSHAW, Paul. What is data journalism? In: GRAY, Jonathan; BOUNEGRU, Liliana; CHAMBERS, Lucy (Org.). **The data journalism handbook: how journalists can use data to improve the news**. Sebastopol: O'Reilly, 2012. p. 2-3.

BRASIL. Decreto nº 8.777, de 11 de maio de 2016. Institui a Política de Dados Abertos do Poder Executivo federal. Brasília, DF, 11 maio. 2017. Disponível em:

<http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/D8777.htm>. Acesso em 25 maio. 2017.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações. Brasília, DF, 18 nov. 2011. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm>. Acesso em: 1 jun. 2017.

BRIGGS, Mark. **Journalism 2.0**: how to survive and thrive - a digital literact guide for the information age. [S.l.]: J-Lab and the Knight Citizen News Network, 2007.

BROESMA, Marcel; DEN HERDER, Bas; SCHOHAUS, Birte. A question of power: the changing dynamics between journalists and sources. **Journalism Practice**. [S.l.]: Routledge, 2013.

BUI, Quoc Trung. Actually, Income in Rural America Is Growing, Too. Nova York: **The New York Times**, 2016a. Disponível em: <<https://www.nytimes.com/2016/09/17/upshot/actually-income-in-rural-america-is-growing-too.html?rref=collection%2Fsectioncollection%2Fupshot>>. Acesso em 20 out. 2017.

BUI, Quoc Trung. Most Detailed Map of Gay Marriage in America. Nova York: **The New York Times**, 2016b. Disponível em: <<https://www.nytimes.com/2016/09/13/upshot/the-most-detailed-map-of-gay-marriage-in-america.html?rref=collection%2Fsectioncollection%2Fupshot>>. Acesso em 20 out. 2017.

BUI, Quoc Trung. A Secret of Many Urban 20-Somethings: Their Parents Help With the Rent. Nova York: **The New York Times**, 2017. Disponível em: <<https://www.nytimes.com/2017/02/09/upshot/a-secret-of-many-urban-20-somethings-their-parents-help-with-the-rent.html?rref=collection%2Fsectioncollection%2Fupshot>>. Acesso em 20 out. 2017.

CANAVILHAS, João; IVARS-NICOLÁS, Begoña. Uso y credibilidad de fuentes periodísticas 2.0 en Portugal y España. [S.l.]: **El profesional de la información**, 2012, jan-fev, N. 1, V. 21. p. 63-69.

CARLSON, Matt; USHER, Nikki. News startups as agents of innovation. **Digital Journalism**. Nova York: Routledge – Taylor & Francis Group, 2015.

CARROLL, Aaron. How Restricting Food Stamp Choices Can Fight Obesity. Nova York: **The New York Times**, 2016. Disponível em: <<https://www.nytimes.com/2016/09/23/upshot/restrictions-on-food-stamp-purchases-can-improve-diets.html?rref=collection%2Fsectioncollection%2Fupshot>>. Acesso em 20 out. 2017.

CASTEEL, Kathryn. Data On Drug Use Is Disappearing Just When We Need It Most. Nova York: **FiveThirtyEight**, 2017. Disponível em: <<https://fivethirtyeight.com/features/data-on-drug-use-is-disappearing-just-when-we-need-it-most/>>. Acesso em 20 out. 2017.

CASTEEL, Kathryn; CHIN, Kara. How Much Progress Have We Made Since Women Went On Strike In 1970? Nova York: **FiveThirtyEight**, 2017. Disponível em: <<https://fivethirtyeight.com/features/how-much-progress-have-we-made-since-women-went-on-strike-in-1970/>>.

on-strike-in-1970/>. Acesso em 20 out. 2017.

CASTELLS, Manuel. **A sociedade em rede**. 17 ed. São Paulo: Paz e Terra, 2016.

CASTELLS, Manuel. **O poder da comunicação**. São Paulo/Rio de Janeiro: Paz e Terra, 2015.

CASTRO, José Roberto. Qual o quadro do investimento estrangeiro direto no Brasil. São Paulo: **Nexo**, 2017. Disponível em: <<https://www.nexojournal.com.br/expresso/2017/06/09/Qual-o-quadro-do-investimento-estrangeiro-direto-no-Brasil>>. Acesso em: 15 jun. 2017.

CHAPARRO, Manuel Carlos. Aula 16: caracterização dos tipos de fonte. [S.l.]: **YouTube**, 2009.

CHAPARRO, Manuel Carlos. **Pragmática do jornalismo**: buscas práticas para uma teoria da ação jornalística. São Paulo: Summus, 2007.

CHEQUEADO. Acercza de Chequeado. Argentina: **Chequeado**: la verificación del discurso público, 2017. Disponível em: <<http://chequeado.com/quienes-somos/>>. Acesso em 20 jun. 2017.

CODDINGTON, Mark. Clarifying journalism's quantitative turn: a typology for evaluating data journalism, computational journalism, and computer-assisted reporting. **Digital Journalism**, 2014. Disponível em: <<http://dx.doi.org/10.1080/21670811.2014.976400>>. Acesso em: 5 jul. 2016.

COHEN, Sarah; HAMILTON, James T.; TURNER, Fred. Computational journalism: how computer scientists can empower journalists, democracy's watchdogs, in the production of news in the public interest. New York: **Communications of the ACM**, Vol. 54, N. 10, Oct. 2011. Disponível em: <<http://cacm.acm.org/magazines/2011/10/131400-computational-journalism/fulltext>>. Acesso em 3 jan. 2017.

COSTA, José María. Originales: unos 29 mil argentinos tienen un nombre único que no se repite en todo el país. Buenos Aires: **La Nación**, 2017. Disponível em: <<http://www.lanacion.com.ar/2056655-argentinos-unicos-fijate-si-solo-vos-te-llamas-de-esa-manera>>. Acesso em 20 out. 2017.

COUTO, Marlen; JULIÃO, Luis Guilherme; VASCONCELLOS, Fábio. O mapa de países onde 'Despacito' faz mais sucesso. **Blog Na Base dos Dados**. Disponível em: <<http://blogs.oglobo.globo.com/na-base-dados/post/onde-despacito-faz-mais-sucesso-isso-mesmo-na-noruega.html>>. Acesso em: 15 jun. 2017.

COUTO, Marlen. Despesa com vices ultrapassa R\$ 64 milhões em 21 estados. **Blog Na Base dos Dados**. Disponível em: <<http://blogs.oglobo.globo.com/na-base-dados/post/despesa-com-vices-ultrapassa-r-64-milhoes-em-21-estados.html>>. Acesso em: 15 jun. 2017.

COX, Melisma. The development of computer-assisted reporting. In: **Southeast Colloquium for the Association for Education in Journalism and Mass Communication**. Chapel Hill, 2000.

CRUCIANELLI, Sandra. **Ferramentas digitais para jornalistas 2.0**. Moçambique: IREX, 2013.

DADER, José Luis. **Periodismo de precisión**: vía socioinformática de descubrir noticias. Madrid: Editorial Sintesis, 2002.

DI SANTI, Matías. ¿Qué muestran los datos sobre los juicios laborales?. Buenos Aires: **Chequeado**, 2017. Disponível em: <<http://chequeado.com/el-explicador/que-muestran-los-datos-sobre-los-juicios-laborales/>>. Acesso em 30 out. 2017.

DI SANTI, Matías; SLIPCZUK, Martín. Crisis del sector lechero: datos para entender los problemas de la industria. Buenos Aires: **Chequeado**, 2017. Disponível em: <<http://chequeado.com/el-explicador/crisis-del-sector-lechero-datos-para-entender-los-problemas-de-la-industria/>>. Acesso em 30 out. 2017.

DIAS, Tatiana. 7 dados que mostram como está a educação brasileira hoje. São Paulo: **Nexo**, 2016. Disponível em: <<https://www.nexojornal.com.br/expresso/2016/04/04/7-dados-que-mostram-como-est%C3%A1-a-educac%C3%A7%C3%A3o-brasileira- hoje>>. Acesso em 15 jun. 2017.

ECHEVERRÍA, Javier. **Los señores del aire**: Telépolis y e el tercer entorno. Barcelona: Ediciones Destino, 1999.

ECHT, Iván. El caso de Chequeado en la Argentina. In: ZOMMER, Laura (Ed.). **El boom del fact checking en América Latina**: aprendizajes y desafíos del caso de Chequeado. Argentina: Chequeado, 2014a. p. 17-18.

EL PAÍS. **Libro de estilo**. Madrid: Ediciones El País, 2008.

ERBOLATO, Mario. **Técnicas de codificação em jornalismo**: redação, captação e edição no jornal diário. São Paulo: Ática, 1991.

ERICSON, Richard V.; BARANEK, Patricia M.; CHAN, Janet B.L. **Visualizing deviance**: a study of news organization. Canadá: University of Toronto Press, 1987.

ESTADÃO DADOS. **Blog**. São Paulo: O Estado de São Paulo, 2017. Disponível em: <<http://blog.estadaodados.com/>>. Acesso em: 15 jun. 2017.

FÁBIO, André Cabette. Por que o Brasil não dá uma resposta efetiva ao aumento persistente dos homicídios? São Paulo: **Nexo**, 2017. Disponível em: <<https://www.nexojornal.com.br/expresso/2017/06/11/Por-que-o-Brasil-n%C3%A3o-d%C3%A1-uma-resposta-efetiva-ao-aumento-persistente-dos-homic%C3%ADdios>>. Acesso em 15 jun. 2017.

FINK, Katherine; ANDERSON, C.W. **Data journalism in the United States**: beyond the “usual suspects”. Journalism studies. Londres: Routledge - Taylor & Francis Group, 2014

FOLHA DE SÃO PAULO. **Manual da redação**. São Paulo: Publifolha, 2010.

FOLHASPADOS. **Blog**. São Paulo: Folha de São Paulo, 2017. Disponível em: <<http://folhaspdados.blogfolha.uol.com.br/>>. Acesso em: 15 jun. 2017.

FONSECA, Virginia Pradelina da Silveira. **Indústria de notícias: capitalismo e novas tecnologias no jornalismo contemporâneo**. Porto Alegre: Editora da UFRGS, 2008.

GANDOUR, Ricardo. Orelha de livro. In: DANTAS, Humberto; TOLEDO, José Roberto de.; TEIXEIRA, Marco Antonio Carvalho (Org.). **Análise política e jornalismo de dados: ensaios a partir do Basômetro**. Rio de Janeiro: Editora FGV, 2014.

GANS, Herbert J. **Deciding what's News: a study of CBS evening news, NBC nightly news, Newsweek and Time**. Evanston: Northwestern University press, 2004.

GRAVES, Lucas; NYHAN, Brendan; REIFLER, Jason. The diffusion of fact-checking: understanding the growth of a journalistic innovation. Arlington: **American Press Institute**, 2015. Disponível em: <<http://www.americanpressinstitute.org/wp-content/uploads/2015/04/The-Growth-of-Fact-Checking.pdf>>. Acesso em 10 jul. 2017.

GEHRKE, Marília; MIELNICZUK, Luciana. Philip Meyer, the outsider who created Precision Journalism. Porto Alegre: **Intexto**, N. 39, 2017. p. 4-13.

GOMIS, Lorenzo. Os interessados produzem e fornecem os fatos. Florianópolis: **Estudos em jornalismo e mídia**, V. 1, N. 1, 2004.

HALL, Stuart et. al. A produção social das notícias: o mugging nos media. In: TRAQUINA, Nelson (Org.). **Jornalismo: questões, teorias e 'estórias'**. Florianópolis: Insular, 2016. p. 309-341.

HE, Lucia. In One Of Latin America's Poorest Countries, Boosting The Economy Is Women's Work. Nova York: **FiveThirtyEight**, 2017. Disponível em: <<https://fivethirtyeight.com/features/in-one-of-latin-americas-poorest-countries-boosting-the-economy-is-womens-work/>>. Acesso em 20 out. 2017.

HEINRICH, Ansgard. Network journalism: moving towards a global journalism culture. Public service media for communication and partnership. Mainz: **RIPE conference**, 2008.

HEINRICH, Ansgard. **Network journalism: journalistic practice in interactive spheres**. Nova York e Abingdon: Routledge, 2011.

HERAVI, Bahareh R. The state of data journalism globally. In: PETRICECK, Tomas; HERAVI, Bahareh R.; STARK, Jennifer A. et al. **Proceedings of the European Data and Computational Journalism Conference**. Dublin: University College Dublin, 2017. p. 5-7.

HICKEY, Walt. The Ultimate Halloween Candy Power Ranking. Nova York: **FiveThirtyEight**, 2017. Disponível em: <<http://fivethirtyeight.com/features/the-ultimate-halloween-candy-power-ranking/>>. Acesso em 20 out. 2017.

HOLOVATY, Adrian. A fundamental way newspapers sites need to change. [S.l.], 2006. Disponível em: <<http://www.holovaty.com/writing/fundamental-change/>>. Acesso em: 4 nov. 2016.

HOWE, Jeff. The rise of crowdsourcing. São Francisco (EUA): **Wired**, 1 jun. 2006. Disponível em: <<http://www.wired.com/2006/06/crowds/>>. Acesso em: 13 ago. 2016.

JORGE, Thaís de Mendonça. **Manual do foca**: guia de sobrevivência para jornalistas. São Paulo: Contexto, 2008.

KISCHINHEVSKY, Marcelo; CHAGAS, Luã José Vaz. Diversidade e pluralidade de fontes no jornalismo da Bandnews Fluminense FM. **Anais**. São Paulo: Compós, 2017.

KOLKO, Jed. 'Normal America' Is Not A Small Town Of White People. Nova York: **FiveThirtyEight**, 2016. Disponível em: <<https://fivethirtyeight.com/features/normal-america-is-not-a-small-town-of-white-people/>>. Acesso em 20 out. 2017.

KOVACH, Bill; ROSENSTIEL, Tom. **Blur**: how to know what's true in the age of information overload. Nova York, Berlim e Londres: Bloomsbury, 2011.

LA NACIÓN. **La Nación**: manual de estilo y ética periodística. Buenos Aires: Espasa Calpe, 1997. Disponível em: <<http://www.mdp.edu.ar/humanidades/documentacion/licad/archivos/modulos/inicial/archivos/bibliografia/inicial/MI019.pdf>>. Acesso em 22 jun. 2017.

LAGE, Nilson. **A reportagem**: teoria e técnica de entrevista e pesquisa jornalística. Rio de Janeiro: Record, 2009.

LEONHARDT, David. Navigate News with the upshot. Nova York: **The New York Times**, 2014. Disponível em: <<https://www.nytimes.com/2014/04/23/upshot/navigate-news-with-the-upshot.html?smid=tw-upshotnyt>>. Acesso e 13 nov. 2017.

LEWIS, Norman P.; WATERS, Stephenson. Data journalism and the challenge of show-leather epistemologies. **Digital Journalism**. Nova York: Routledge – Taylor & Francis Group, 2017.

LEWIS, Peter H. The New York Times introduces a web site. Nova York: **The New York Times**, 1996. Disponível em: <<http://www.nytimes.com/1996/01/22/business/the-new-york-times-introduces-a-web-site.html>>. Acesso em 13 nov. 2017.

LEWIS, Seth; WESTLUND, Oscar. Big data and journalism: epistemology, expertise, economics, and ethics. **Digital Journalism**. [S.l.]: Routledge – Taylor and Francis Group, 2014.

MACHADO, Elias. **O ciberespaço como fonte para os jornalistas**. Salvador: Calandra, 2003.

MACHADO, Elias. **O jornalismo digital em base de dados**. Florianópolis: Calandra, 2006.

MANNING, Paul. **News and news sources**: a critical introduction. Londres: Sage, 2001.

MARIANI, Daniel; ALMEIDA, Rodolfo. O horário das refeições no Brasil e no mundo, segundo o Twitter. São Paulo: **Nexo**, 2017. Disponível em: <<https://www.nexojornal.com.br/grafico/2017/06/04/O-hor%C3%A1rio-das->

refei%C3%A7%C3%B5es-no-Brasil-e-no-mundo-segundo-o-Twitter>. Acesso em 15 jun. 2017.

MARIANI, Daniel; LUPION, Bruno; ALMEIDA, Rodolfo. 10 índices econômicos e sociais nos 13 anos de governo PT no Brasil. São Paulo: **Nexo**, 2016. Disponível em: <<https://www.nexojornal.com.br/especial/2016/09/02/10-%C3%ADndices-econ%C3%B4micos-e-sociais-nos-13-anos-de-governo-PT-no-Brasil>>. Acesso em 15 jun. 2017.

MARIANI, Daniel; LUPION, Bruno; ALMEIDA, Rodolfo. Qual é o grau de discordância e concordância entre os ministros do Supremo. São Paulo: **Nexo**, 2017. Disponível em: <<https://www.nexojornal.com.br/especial/2017/03/21/Como-os-ministros-do-Supremo-se-aproximam-ou-se-distanciam-entre-si-de-acordo-com-suas-decis%C3%B5es>>. Acesso em 15 jun. 2017.

MARTÍNEZ, Lucía. Índice de Datos Abiertos: en qué mejoró y en qué no la Argentina para subir en el ranking. Buenos Aires: **Chequeado**, 2017. Disponível em: <<http://chequeado.com/el-explicador/indice-de-datos-abiertos-en-que-mejoro-y-en-que-no-la-argentina-para-subir-en-el-ranking/>>. Acesso em 20 out. 2017.

McDULING, John. “The Upshot” is the New York Times’ replacement for Nate Silver’s FiveThirtyEight. Nova York: **Quartz**, 2014. Disponível em: <<https://qz.com/185922/the-upshot-is-the-new-york-times-replacement-for-nate-silvers-fivethirtyeight/>>. Acesso em: 24 jun. 2017.

MEYER, Philip. **Paper route**: finding my way to precision journalism. Bloomington: iUniverse: 2012.

_____. **Precision Journalism**. Lanham: Rowman & Littlefield Publishers, 2002.

_____. The future of CAR: declare victory and get out. In: PAUL, Nora (Org.). **When nerds and words collide**: reflections on the development of Computer Assisted Reporting. São Petersburgo: Poynter Institute, 1999. p. 4-5.

_____. **The new precision journalism**. Bloomington: Indiana University Press, 1991.

MIELNICZUK, Luciana; TRÄSEL, Marcelo. Jornalismo guiado por dados como inovação profissional e seus desafios para a educação. **Contemporânea Revista de Comunicação e Cultura**, Salvador, V. 14., N. 2, p. 609-629, 2017.

MILLER, Claire Cain; BUI, Quoctrung. Switching Careers Doesn’t Have to Be Hard: Charting Jobs That Are Similar to Yours. Nova York: **New York Times**, 2017. Disponível em: <<https://www.nytimes.com/2017/07/27/upshot/switching-careers-is-hard-it-doesnt-have-to-be.html?rref=collection%2Fsectioncollection%2Fupshot>>. Acesso em 20 out. 2017.

MILLER, Claire Cain. The Gender Pay Gap Is Largely Because of Motherhood. Nova York: **The New York Times**, 2017. Disponível em: <<https://www.nytimes.com/2017/05/13/upshot/the-gender-pay-gap-is-largely-because-of-motherhood.html?rref=collection%2Fsectioncollection%2Fupshot>>. Acesso em 30 out. 2017.

MIRANDE, Ayelén. A nueve años de su sanción, ¿se cumple la llamada “ley de bosques”? Buenos Aires: **Chequeado**, 2016. Disponível em: <<http://chequeado.com/el-explicador/a-nueve-anos-de-su-sancion-se-cumple-la-llamada-ley-de-bosques/>>. Acesso em 20 out. 2017.

MOLOTCH, Harvey; LESTER, Marilyn. As notícias como procedimento intencional: acerca do uso estratégico de acontecimentos de rotina, acidentes e escândalos. In: TRAQUINA, Nelson (Org.). **Jornalismo: questões, teorias e ‘estórias’**. Florianópolis: Insular, 2016. p. 61-81.

MUSSE, Valeria. La telefonía móvil y los bancos se mantienen al tope de los reclamos desde hace cinco años. Buenos Aires: **La Nación**, 2017. Disponível em: <<http://www.lanacion.com.ar/2059400-la-telefonía-movil-y-los-bancos-se-mantienen-al-tope-de-los-reclamos-desde-hace-cinco-anos>>. Acesso em 20 out. 2017.

NA BASE DOS DADOS. **Blog**. Rio de Janeiro: O Globo, 2017. Disponível em: <<http://blogs.oglobo.globo.com/na-base-dos-dados/>>. Acesso em: 15 jun. 2017.

NACIÓN DATA. Acerca de Nación Data. Argentina: **La Nación**, 2017. Disponível em: <<http://blogs.lanacion.com.ar/data/about/>>. Acesso em 20 jun. 2017.

NEXO. Apresentação. **Documento eletrônico**. Recebido por <mariliagehrke@gmail.com> em 12 jun. 2017. São Paulo, 2017.

NEXO JORNAL. **Sobre o Nexo**. São Paulo, 2017. Disponível em <<https://www.nexojournal.com.br/>>. Acesso em: 15 jun. 2017.

NIKLAS, Jan. Pará tem 30% das mortes em conflitos por terra no Brasil. **Blog Na base dos Dados**. Rio de Janeiro: O Globo, 2017. Disponível em: <<http://blogs.oglobo.globo.com/na-base-dos-dados/post/estado-mais-violento-para-tem-30-das-mortes-em-conflitos-por-terra-no-brasil.html>>.

O ESTADO DE S. PAULO. Manual de redação. São Paulo: O Estado de S. Paulo, 2017. Disponível em: <<http://www.estadao.com.br/manualredacao/>>. Acesso em 30 out. 2017.

O GLOBO. Manual de Redação e Estilo. 28 ed. São Paulo: Globo, 2001.

O GLOBO MEMÓRIA. Rio de Janeiro: O Globo, 2017. Disponível em: <<http://memoria.oglobo.globo.com/>>. Acesso em 20 jun. 2017.

ORENSTEIN, José. Por que o Japão tem as menores taxas de obesidade do mundo. São Paulo: **Nexo**, 2017. Disponível em: <<https://www.nexojournal.com.br/expresso/2017/06/14/Por-que-o-Jap%C3%A3o-tem-as-menores-taxas-de-obesidade-do-mundo>>. Acesso em 15 jun. 2017.

PAINÉ, Neil. Who hates The Patriots the most? Nova York: **FiveThirtyEight**, 2017. Disponível em: <<https://fivethirtyeight.com/features/who-hates-the-patriots-the-most/>>. Acesso em 30 out. 2017.

PALLARO, Bianca. El 93% de las personas que esperan un trasplante necesita un riñón o hígado. Buenos Aires: **La Nación**, 2017. Disponível em: <<http://www.lanacion.com.ar/2047506-el-93-de-las-personas-que-esperan-un-trasplante-necesita-un-rinon-o-higado>>. Acesso em 30 out. 2017.

PARASIE, Sylvain. Data-driven revelation? Epistemological tensions in investigative journalism in the age of “big data”. **Digital Journalism**. [S.l.]: Routledge – Taylor and Francis Group, 2014.

PEUCER, Tobias. Os relatos jornalísticos. Trad. Paulo da Rocha Dias. Florianópolis: **Estudos em Jornalismo e Mídia**, V. 1, N. 2, 2004.

PIMENTEL, Matheus. Campeonato Brasileiro: história conturbada, disputas por dinheiro e paixão nacional. São Paulo: **Nexo**, 2017. Disponível em: <<https://www.nexojournal.com.br/explicado/2017/05/14/Campeonato-Brasileiro-hist%C3%B3ria-conturbada-disputas-por-dinheiro-e-paix%C3%A3o-nacional>>. Acesso em: 15 jun. 2017.

PINTO, Ana Estela de Sousa. **Jornalismo diário**: reflexões, recomendações, dicas e exercícios. São Paulo: Publifolha, 2009.

PINTO, Manuel. Fontes jornalísticas: contributo para o mapeamento do campo. [S.l.]: **Comunicação e Sociedade 2**, Cadernos do Noroeste, Série Comunicação, V. 14, 2000, p. 277-294.

PRESS, Gil. A very short history of big data. S.l.: **Forbes**, 2013. Disponível em: <<https://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/#662438bc65a1>>. Acesso em 10 maio. 2017.

PRIMO, Alex Fernando Teixeira. Transformações no jornalismo em rede: sobre pessoas comuns, jornalistas e organizações; blogs, Twitter, Facebook e Flipboard. Porto Alegre: **Intexto** (UFRGS), V. 2, N. 25, 2011. p. 130-146.

RODRIGO ALSINA, Miquel. **A construção da notícia**. Petrópolis: Vozes, 2009.

ROGERS, Simon. **Facts are sacred**. Faber & Faber, 2013a. E-book.

ROGERS, Simon. John Snow's data journalism: the cholera map that changed the world. Londres: **The Guardian**, 2013b. Disponível em: <<https://www.theguardian.com/news/datablog/2013/mar/15/john-snow-cholera-map>>. Acesso em: 18 nov. 2017.

ROGERS, Simon. Wikileaks data journalism: how we handled the data. Londres: **The Guardian**, 2011. Disponível em: <<https://www.theguardian.com/news/datablog/2011/jan/31/wikileaks-data-journalism>>. Acesso em 14 nov. 2017.

RUIZ, Iván; JASTREBLANSKY, Maia; MON, Hugo A. Los argentinos que figuran com sociedades offshore em paraísos fiscales. Buenos Aires: **La Nación**, 2016. Disponível em: <<http://www.lanacion.com.ar/1885775-los-argentinos-que-figuran-con-sociedades-offshore-en-paraisos-fiscales>>. Acesso em 30 out. 2017.

SANGER-KATZ, Margot. Is Terrorism Getting Worse? In the West, Yes. In the World, No. Nova York: **The New York Times**, 2016. Disponível em: <<https://www.nytimes.com/2016/08/16/upshot/is-terrorism-getting-worse-in-the-west-yes-in-the-world-no.html?rref=collection%2Fsectioncollection%2Fupshot>>. Acesso em 20 out. 2017.

SANTOS, Rogério. **A negociação entre jornalistas e fontes**. Coimbra: Minerva, 1997.

SCHMITZ, Aldo Antonio. **Fontes de notícias: ações e estratégias das fontes no jornalismo**. Florianópolis: Combook, 2011.

SCHUDSON, Michael. **The sociology of news**. Nova York: Norton & Company, 2003.

SIEGAL, Allan; CONNOLLY, William. **The New York Times Manual of Style and Usage: The Official Style Guide Used by the Writers and Editors of the World's Most Authoritative Newspaper**. Nova York: The New York Times Company, 2015.

SIGAL, Leon V. Who? Sources make the news. In: MANOFF, Robert Karl; SCHUDSON, Michael (Ed.). **Reading the news: a pantheon guide to popular culture**. New York: Pantheon Books, 1987. p. 9-37.

SILVA, Gislene; MAIA, Flávia. O método Análise de Cobertura Jornalística na compreensão do crack como acontecimento noticioso. In: LEAL, Bruno Souza; ANTUNES, Elton; VAZ, Paulo Bernardo. **Jornalismo e acontecimento: percursos metodológicos**. Florianópolis: Insular, 2011. p. 37-54.

SLIPCZUK, Martín. Aprender 2016: ¿qué relación hay entre el nivel socioeconómico del alumno y que vaya a una escuela pública o privada?. Buenos Aires: **Chequeado**, 2017. Disponível em: <<http://chequeado.com/el-explicador/aprender-2016-que-relacion-hay-entre-el-nivel-socioeconomico-del-alumno-y-que-vaya-a-una-escuela-publica-o-privada/>>. Acesso em 30 out. 2017.

SOARES, Marcelo. **Afinal de contas**. Blog. São Paulo: Folha de São Paulo, 2016. Disponível em: <<http://afinaldecontas.blogfolha.uol.com.br/>>. Acesso em: 15 jun. 2017.

SOHR, Olivia. La Argentina es el cuarto país de la región con más mujeres en el Congreso. Buenos Aires: **Chequeado**, 2017. Disponível em: <<http://chequeado.com/el-explicador/la-argentina-es-el-cuarto-pais-de-la-region-con-mas-mujeres-en-el-congreso/>>. Acesso em 30 out. 2017.

SOUZA, Rose Mara Vidal de. Lei de acesso à informação: um canal à fonte de notícia. Tese (Doutorado em Comunicação Social), **Unesp**, São Paulo, 2016.

STAVELIN, Eirik. Computational journalism: when journalism meets programming. Tese de doutorado, **Universidade de Bergen**, Bergen, 2013.

TAMBOSI, Orlando. Tobias Peucer e as origens do jornalismo. Florianópolis: **Estudos em Jornalismo e Mídia**, V. 1, N. 2, 2004.

THE NEW YORK TIMES. Our history. Nova York: **The New York Times**, 2017.
Disponível em: <<https://www.nytc.com/who-we-are/culture/our-history/#1835-1880>>.
Acesso em 13 nov. 2017.

TRÄSEL, Marcelo. Entrevistando planilhas: estudo das crenças e do ethos de um grupo de profissionais de jornalismo guiado por dados no Brasil. Tese (Doutorado em Comunicação Social), **PUCRS**, Porto Alegre, 2014.

VASCONCELLOS, Fábio. Denúncias de violência doméstica e familiar contra a mulher crescem 133%. Blog **Na base dos Dados**. Rio de Janeiro: O Globo, 2017a. Disponível em: <<http://blogs.oglobo.globo.com/na-base-dos-dados/post/denuncias-de-violencia-domestica-e-familiar-contramulher-crescem-133.html>>. Acesso em: 15 jun. 2017.

VASCONCELLOS, Fábio. Mudança de faixa populacional salva receita do FPM de 13 municípios. **Blog na Base dos Dados**. Rio de Janeiro: O Globo, 2016a. Disponível em: <<http://blogs.oglobo.globo.com/na-base-dos-dados/post/mudanca-de-faixa-populacional-salva-receita-do-fpm-de-13-municipios.html>>. Acesso em: 15 jun. 2017.

VILLAFANE, Justo; BUSTAMANTE, Enrique, PRADO, Emilio. **Fabricar noticias: las rutinas productivas en radio y television**. Barcelona: Editorial Mitre, 1987.

ZAGO, Gabriela da Silva. Circulação e recirculação de narrativas do acontecimento no jornalismo em rede: a copa do mundo de 2014 no twitter. Tese (Doutorado em Comunicação e Informação), **Ufrgs**, Porto Alegre, 2014.

APÊNDICE A – FICHAS DE LEITURA

O Apêndice A é um arquivo digital composto por todas as fichas de leitura utilizadas na descrição e classificação das fontes encontradas nas 60 notícias analisadas no *corpus* deste trabalho. Este arquivo pode ser conferido pelo link goo.gl/e9AC3c.

APÊNDICE B – CATEGORIAS PRELIMINARES

Em um primeiro momento, antes do processo de união das categorias por similaridade, as 315 fontes deste levantamento foram classificadas em 19 tipos. São eles:

- a) **Análise** – empregada principalmente por especialistas que visam dar contexto às notícias.
- b) **Assessoria** – traz informações oficiais e posicionamentos que são utilizados como fonte.
- c) **Base de dados** – é o nome inicialmente dado a bases de dados estruturadas e consultadas para fins de produção noticiosa.
- d) **Base de dados própria** – dá nome às informações reunidas e utilizadas pelo mesmo veículo de comunicação.
- e) **Comunicado** – termo associado principalmente às notas enviadas por assessorias de imprensa e cujo conteúdo foi utilizado nas notícias, tornando-se fonte.
- f) **Declaração** – em geral torna-se fonte para complementar as informações divulgadas por meio de dados.
- g) **Decreto** – torna-se fonte para contextualizar o tema da notícia, geralmente indicando a origem do documento.
- h) **Deputado** – costuma ser fonte declaratória a partir de entrevista.
- i) **Documento** – pode adquirir vários formatos, mais ou menos especificados nas notícias.
- j) **Especialista** – categoria de fonte identificada tanto em se tratando de pessoas quanto documentos.
- k) **Governo** – declarações ou documentos complementares aos dados e que são usados como fontes.
- l) **Nota** – geralmente tem caráter oficial e é utilizada desta forma pelos veículos de comunicação.
- m) **Nota técnica** – em geral é utilizada por órgãos oficiais e técnicos e reproduzida nas notícias como fonte.
- n) **Pesquisa** – geralmente é denominada desta forma quando tem a ver com estudos científicos.
- o) **Ranking** – geralmente feito por algum veículo de comunicação ou instituição confiável e reproduzido por outros.

- p) **Representante** – a fonte inicialmente chamada de representante está vinculada a uma instituição ou órgão de governo.
- q) **Relatório** – geralmente está ligada à produção de dados públicos sobre um tema.
- r) **Resolução** – vinculada a algum órgão oficial para explicar uma medida ou um posicionamento.
- s) **Site/blog** – podem ser utilizados como fonte quando apresentam informações necessárias e específicas para a construção de uma notícia.

APÊNDICE C – PLANILHA DE FONTES

O Apêndice C corresponde à planilha digital utilizada para organização estruturada das 315 fontes (pessoais e documentais) encontradas nas 60 notícias analisadas. O arquivo pode ser consultado pelo link goo.gl/BUiF Xu.