

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

MARÍLIA GEHRKE

**OS ELEMENTOS DE TRANSPARÊNCIA NO JORNALISMO GUIADO POR DADOS**

PORTO ALEGRE

2021

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

MARÍLIA GEHRKE

**OS ELEMENTOS DE TRANSPARÊNCIA NO JORNALISMO GUIADO POR DADOS**

Tese de doutorado apresentada como requisito parcial para obtenção do título de Doutora pelo Programa de Pós-Graduação em Comunicação da Universidade Federal do Rio Grande do Sul (PPGCOM/UFRGS).

Orientadora: Prof.<sup>a</sup> Dra. Marcia Benetti

PORTO ALEGRE

2021

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

MARÍLIA GEHRKE

**OS ELEMENTOS DE TRANSPARÊNCIA NO JORNALISMO GUIADO POR DADOS**

BANCA EXAMINADORA

---

Prof.<sup>a</sup> Dra. Marcia Benetti (orientadora) – UFRGS

---

Prof. Dr. Marcelo Träsel – UFRGS

---

Prof. Dr. Josenildo Luiz Guerra – UFS

---

Prof.<sup>a</sup> Dra. Suzana Oliveira Barbosa – UFBA

---

Prof. Dr. Carlos Eduardo Franciscato – UFS

PORTO ALEGRE

2021

CIP - Catalogação na Publicação

Gehrke, Marília  
Os elementos de transparência no Jornalismo Guiado  
por Dados / Marília Gehrke. -- 2021.  
230 f.  
Orientadora: Marcia Benetti.

Tese (Doutorado) -- Universidade Federal do Rio  
Grande do Sul, Faculdade de Biblioteconomia e  
Comunicação, Programa de Pós-Graduação em Comunicação  
e Informação, Porto Alegre, BR-RS, 2021.

1. jornalismo digital. 2. jornalismo guiado por  
dados. 3. jornalismo de dados. 4. transparência. 5.  
covid-19. I. Benetti, Marcia, orient. II. Título.

*À Luciana Mielniczuk (in memoriam)*

## AGRADECIMENTOS

Em 9 de novembro de 2015, durante a seleção para o mestrado no PPGCOM, eu estava nervosa e assustada. Morava no interior do Estado, havia visitado Porto Alegre algumas vezes, e uma vida na capital ainda era algo pouco tangível. Havia passado pela etapa de provas no PPGCOM/UFRGS, e chegava a hora da entrevista. Minha banca era composta por duas referências na pesquisa em jornalismo e jornalismo digital: Dra. Marcia Benetti e Dra. Luciana Mielniczuk. Em certo momento da conversa, quando eu pensava que meu projeto sobre a cobertura jornalística relacionada ao tabaco não se encaixava nas áreas de estudo de nenhuma delas, lancei mão do meu último recurso: perguntei se podia mostrar o *corpus* empírico que havia coletado. Imprimi e organizei os materiais em uma pasta composta por envelopes plásticos, para facilitar o manuseio e a apresentação. No ano seguinte, já como aluna do mestrado, a Luti me contou que esse movimento foi decisivo para ela ter um *insight* e me escolher como orientanda: naquele momento, quando tirei a pasta da mochila, ela associou os meus materiais aos *bancos de dados*, e perguntou se eu tinha interesse em estudar o tema. Naquele momento, nascia uma pesquisadora em Jornalismo Guiado por Dados. A cobertura relacionada ao tabaco e o uso de estatística nesse contexto, aliás, foi o primeiro artigo que escrevi no mestrado, em parceria com a Luti e com o Francisco Amorim, hoje doutor em Sociologia, entusiasta das planilhas e dos métodos quantitativos.

Agradeço à Luti, a quem dedico esta tese, por ver em mim algum potencial; e por ser uma pesquisadora visionária e orientadora dedicada. Começar a carreira acadêmica com a Luti me proporcionou aprendizados os mais diversos, incluindo estratégias de ensino que ela mesma criava, como a “carta ao Vítor” – uma referência a Vítor Prado, à época aluno de iniciação científica, para descrever detalhadamente o método científico adotado na pesquisa; e a “caixinha das ideias”, que servia para amenizar minha ansiedade por definir recortes de pesquisa. A Luti era assim: tinha uma solução para tudo. Hoje, quando me pego em uma situação difícil, tento imaginar um diálogo com ela.

Agradeço à minha orientadora, Marcia Benetti, referência nos estudos de jornalismo, pelo acolhimento e pelas trocas significativas ao longo do doutorado, não apenas nesta tese. Profissional ética, generosa, empática e inspiradora, em quem me espelho na docência, na pesquisa e quando me aventuro a escrever crônicas. Marcia tem um texto impecável e uma postura determinada e resiliente que muito admiro. Durante a pandemia

de Covid-19, não mediu esforços para preparar materiais didáticos e de apoio para os alunos da Fabico, levando em conta o (limitado) acesso à internet e aos equipamentos eletrônicos. Sorte a minha de começar e encerrar esta jornada no PPGCOM com a preciosa banca de 2015, composta pela Luti e pela Marcia.

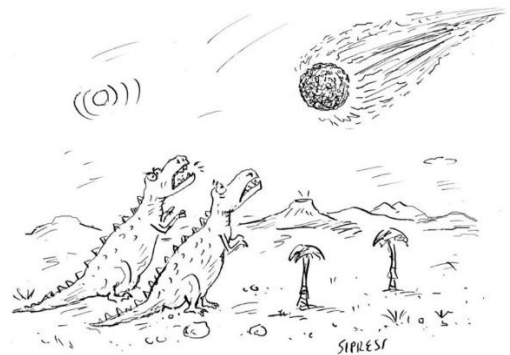
Nesta seara, incluo meu agradecimento à Universidade Federal do Rio Grande do Sul (UFRGS) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), respectivamente, pelo ensino de qualidade e pela oportunidade de cursar o mestrado e o doutorado com dedicação exclusiva. Espero que a universidade e o apoio à pesquisa resistam aos tempos sombrios pelos quais a ciência brasileira atravessa.

Agradeço, ainda, aos membros da banca, professores Marcelo Träsel, Josenildo Guerra, Carlos Franciscato e Suzana Barbosa, grandes nomes da pesquisa em jornalismo no Brasil, pela leitura atenta desta tese. Aos dois primeiros, pela interlocução na defesa do projeto de qualificação, crucial para os rumos desta pesquisa. Ao professor Träsel, meu agradecimento pela presença, pelo apoio e pelo diálogo ao longo de todo o meu período como pós-graduanda na UFRGS. Foi uma honra trabalhar muito próxima da minha principal referência em Jornalismo Guiado por Dados.

Sou grata por conhecer pessoas incríveis nesses seis anos, em especial Luiza Santos, Fabiana Freitas, Lucas Durr Missau, Ludimila Matos, Taís Seibt, Anelise Dias e Samyr Paz, e cultivar amizades de longa data (obrigada, Érica Mallmann e Juliana Spilimbergo). Aos colegas do Núcleo de Pesquisa em Jornalismo (Nupejor), agradeço pelas tardes de quinta-feira com leituras e desabafos, mesmo que à distância. Estendo a gratidão aos colegas do Grupo de Pesquisa em Jornalismo Digital (JorDi) e aos meus alunos do MBA em Jornalismo de Dados do Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP), onde nasceu o Núcleo de Estudos em Jornalismo de Dados e Computacional (DataJor).

Aos meus pais, Marlise e Armênio, agradeço pela construção de bases familiares sólidas pela via da educação e do trabalho; reconheço principalmente o papel da minha mãe neste contexto, e sou grata pela coragem, garra e fibra que teve para sair do interior e se especializar em uma época de poucas oportunidades para as mulheres; à minha irmã, Milena, pelo apoio e pelas ligações telefônicas nos momentos difíceis.

Ao meu esposo, Gabriel, minha fortaleza, agradeço a compreensão e cumplicidade. *Danke, Lieblingsmensch.*



*"I blame the media."*

*"Without facts, you can't have truth. Without truth, you can't have trust. Without trust, we have no shared reality, no democracy, and it becomes impossible to deal with our world's existential problems: climate, coronavirus, the battle for truth."*

Maria Ressa  
Nobel Peace Prize 2021



## RESUMO

Esta tese investiga a transparência no Jornalismo Guiado por Dados (JGD), conjunto de saberes e práticas vinculado ao jornalismo digital e em rede, e inserido em um cenário de excesso de informação e desinformação. Está assentada sobre as teorias do jornalismo como forma (PARK, 1940; MEDITSCH, 1992; 1997) e formas (NIELSEN, 2017) de conhecimento. Imbricada com valores fundamentais como a verdade e a objetividade, a transparência (KARLSSON, 2010; 2020; 2022) diz respeito à abertura de processos e decisões editoriais, permitindo que a audiência tenha acesso a informações normalmente assimétricas na relação entre jornal e público. Dessa forma, apresento a transparência em duas amplas vertentes: como **operadora de uma prática ética** e como **recurso discursivo de legitimação**. Para examinar como este valor aparece no produto jornalístico, combino o estudo de casos múltiplos (STAKE, 2006) e o estudo de caso como ilustração (MACHADO; PALACIOS, 2006), observando a cobertura de 11 jornais brasileiros de referência e nativos digitais em relação à pandemia de Covid-19. Como resultado, sistematizo 15 elementos de transparência, distribuídos nos níveis **organizacional e operacional**, detectados na cobertura feita dentro do JGD e aqui dispostos em ordem alfabética: **assinatura, perfil e expertise; contato; correção de erros; data de publicação; documentos; fontes de dados; fontes de financiamento; graus de incerteza; marcas de apuração; manual da redação e princípios editoriais; metodologia; política de diversidade; política de privacidade; tipo de conteúdo; reprodutibilidade e código aberto**. Esses parâmetros servem como guia para o que chamo de **Protocolo de Transparência no Jornalismo Guiado por Dados**, destinado à análise de produtos jornalísticos já publicados. Por fim, detectando uma lacuna teórico-experimental, produzo a proposta de **Método para o Jornalismo Guiado por Dados com foco na transparência e na reprodutibilidade** para implementação em iniciativas jornalísticas orientadas por dados. Entendo que a pandemia acelerou o processo de incorporação da transparência como valor fundamental para o JGD, demonstrada anteriormente no discurso de profissionais e agora verificada na esfera do produto. A tendência é de que essas mudanças perdurem e se aprimorem, reafirmando o compromisso do JGD de se aproximar da ciência pela via do método para gerar um tipo de conhecimento específico, mais contextual e duradouro.

**Palavras-chave:** jornalismo digital; Jornalismo Guiado por Dados; Jornalismo de Dados; transparência; Covid-19.

## ABSTRACT

This doctoral dissertation investigates transparency in data journalism. Data journalism, or data-driven journalism, is a set of practices and knowledge connected to digital and network journalism. It is based on journalism theories that discuss the news as a form (PARK, 1940; MEDITSCH, 1992; 1997) or forms (NIELSEN, 2017) of knowledge. Transparency (KARLSSON, 2010; 2020; 2022) is also connected to journalism fundamentals such as truth and objectivity, so it comprehends openness of processes and editorial decisions, giving the audience access to information typically only used by journalists. In this regard, I present transparency in two axes: as a **guide to ethical practices** and a **resource to legitimize news organizations**. To examine how this value appears in terms of journalistic product, I combine the methodologies of multiple case studies (STAKE, 2006) and case studies as illustration (MACHADO; PALACIOS, 2006) to observe the news coverage conducted by 11 news outlets (legacy media and native digital) during the Covid-19 pandemic. As a result, I systematize 15 elements of transparency in two levels: **organizational** and **operational**, here alphabetically organized: **byline, profile, and expertise; contact; content type; correction; data sources; diversity policy; documents; financing sources; manual of style and editorial principles; methodology; open-source code and reproducibility; privacy policy; publication date; uncertainty degrees; verification tactics**. These are the parameters for the **Data Journalism' Transparency Protocol**, which is destined to analyze journalistic products that were already published. In addition, observing a theoretical and experimental gap, I propose the **Data Journalism Method focused on transparency and reproducibility** to be implemented by news organizations in investigations using data. I argue that the pandemic has accelerated the presence of transparency as an essential value for data journalism. It was previously demonstrated by studies that included professionals' discourses, and now it is possible to recognize it in the published content. The tendency is that those changes will be maintained and improved, consolidating the commitment of data journalism to be practiced towards science through the method and, thus, generate a specific type of knowledge, which is contextual and has a longer duration.

**Keywords:** digital Journalism, Data-Driven Journalism, Data journalism, transparency, Covid-19.

## LISTA DE FIGURAS

Figura 1 – Argumentos e opiniões .....	37
Figura 2 – Painel Coronavírus, vinculado ao Ministério da Saúde .....	42
Figura 3 – Linha do tempo sobre a abertura e o apagão de dados.....	45
Figura 4 – Pirâmide invertida do Jornalismo Guiado por Dados .....	82
Figura 5 – Fluxograma de trabalho do <i>The Pudding</i> .....	83
Figura 6 – Fluxograma de trabalho da <i>Deutsche Welle</i> .....	85
Figura 7 – Modelo iceberg de transparência.....	88
Figura 8 – Conhecimento prévio do repórter é listado em notícia no <i>Jota</i> .....	132
Figura 9 – Nome dos repórteres funciona como indexador na <i>Folha</i> .....	133
Figura 10 – Contato com o ombudsman é um dos destaques na <i>Folha</i> .....	134
Figura 11 – <i>Nexo</i> apresenta errata ao final das notícias.....	136
Figura 12 – <i>Erramos</i> é destacado pela <i>Folha</i> junto à data e ao final do texto .....	137
Figura 13 – Botão para comunicação de erros aparece na <i>Folha</i> .....	138
Figura 14 – <i>Agência Pública</i> apresenta gráficos sobre despesas e receitas .....	139
Figura 15 – <i>Fiquem Sabendo</i> possui seção específica de prestação de contas.....	140
Figura 16 – <i>Nexo</i> apresenta aos leitores sua política editorial.....	141
Figura 17 – Política de diversidade da <i>Folha</i> .....	143
Figura 18 – <i>Núcleo</i> abre política de privacidade aos leitores .....	144
Figura 19 – <i>Núcleo</i> traz dados de tráfego de seu site.....	145
Figura 20 – <i>GZH</i> criou série chamada <i>Fluência em Notícias</i> .....	146
Figura 21 – Banner que relembra data de publicação é realçado no <i>Núcleo</i> .....	148
Figura 22 – Ofício acompanha reportagem da <i>Agência Pública</i> .....	149
Figura 23 – Hiperlinks remetem às fontes de dados no jornal <i>El País</i> .....	150
Figura 24 – Gráficos apresentam opção para download dos dados em <i>El País</i> .....	151
Figura 25 – Limitações em base de dados foram levantadas pela <i>Folha</i> .....	153
Figura 26 – <i>Piauí</i> registra o dia e horário da análise.....	154
Figura 27 – Metodologia é esclarecida pelo <i>Estadão</i> .....	156
Figura 28 – <i>Núcleo</i> usa <i>Como fizemos isso</i> para detalhar método .....	157
Figura 29 – Reportagem do <i>Estadão</i> divulga código-fonte no <i>GitHub</i> .....	159
Figura 30 – Projeto <i>No Epicentro</i> apresenta mapa para inserção de dados .....	160
Figura 31 – Projeto <i>No Epicentro</i> , da <i>Agência Lupa</i> , divulga código-fonte .....	161
Figura 32 – Proposta metodológica com ênfase na transparência .....	177
Figura 33 – Exemplo de repositório e código-fonte utilizado em <i>No Epicentro</i> .....	190

## LISTA DE QUADROS

Quadro 1 – Resumo dos princípios do Jornalismo Guiado por Dados.....	75
Quadro 2 – Os sete princípios do feminismo de dados .....	77
Quadro 3 – Lista dos textos informativos que compõem o corpus .....	162
Quadro 4 – Os elementos de transparência no Jornalismo Guiado por Dados.....	164
Quadro 5 – Indicadores do Protocolo de Transparência no Jornalismo Guiado por Dados .....	172

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>14</b>
<b>2 INFORMAÇÃO E DESINFORMAÇÃO NA COBERTURA DA COVID-19 .....</b>	<b>22</b>
2.1 O complexo cenário da desinformação e suas variações .....	29
2.2 Do topo para baixo: a influência de políticos e figuras públicas.....	34
2.3 A efetividade da desinformação e estratégias para combatê-la .....	37
2.4 Transparência pública e abertura de dados sobre a Covid-19.....	42
2.5 A pandemia de Covid-19 como acontecimento jornalístico .....	46
<b>3 DIGITAL E EM REDE: O LUGAR DO JORNALISMO GUIADO POR DADOS.....</b>	<b>53</b>
3.1 O potencial para geração de conhecimento específico.....	58
3.2 A base científica do Jornalismo de Precisão .....	66
3.3 Características do Jornalismo Guiado por Dados na atualidade.....	72
3.3.1 Modos de fazer: método e fluxos de trabalho .....	79
<b>4 TRANSPARÊNCIA COMO ABERTURA DO MÉTODO E DE DECISÕES EDITORIAIS .....</b>	<b>87</b>
4.1 Verdade .....	89
4.2 Objetividade .....	94
4.3 Transparência .....	102
4.3.1 Operadora de uma prática ética .....	109
4.3.2 Recurso discursivo de legitimação .....	112
4.3.3 A transparência no Jornalismo Guiado por Dados.....	117
<b>5 ELEMENTOS DE TRANSPARÊNCIA: PERCURSO METODOLÓGICO E RESULTADOS.....</b>	<b>122</b>
5.1 Estudo de casos múltiplos e como ilustração .....	124
5.2 Transparência e seus níveis: organizacional e operacional.....	129
5.2.1 Nível organizacional: o jornal .....	130
5.2.1.1 Assinatura, perfil e expertise.....	131
5.2.1.2 Contato .....	134
5.2.1.3 Correção de erros.....	135
5.2.1.4 Fontes de financiamento .....	138
5.2.1.5 Manual da redação e princípios editoriais.....	141
5.2.1.6 Política de diversidade .....	142
5.2.1.7 Política de privacidade .....	144
5.2.1.8 Tipo de conteúdo .....	146
5.2.2 Nível operacional: o método .....	147
5.2.2.1 Data de publicação .....	147
5.2.2.2 Documentos.....	148

5.2.2.3 Fontes de dados .....	150
5.2.2.4 Graus de incerteza .....	152
5.2.2.5 Marcas de apuração .....	154
5.2.2.6 Metodologia .....	155
5.2.2.7 Reprodutibilidade e código aberto.....	158
5.2.3 Resumo dos resultados.....	162
<b>6 CONTRIBUIÇÕES .....</b>	<b>166</b>
6.1 Protocolo de Transparência no Jornalismo Guiado por Dados .....	166
6.2 Proposta de método com foco na transparência e na reprodutibilidade.....	175
6.2.1 Problema.....	180
6.2.2 Hipótese .....	181
6.2.3 Pesquisa.....	183
6.2.4 Contexto .....	184
6.2.5 Coleta .....	187
6.2.6 Análise.....	188
6.2.7 Código aberto e reprodutibilidade .....	189
6.2.8 Produto .....	192
6.2.9 Replicabilidade .....	192
6.2.10 Documentação suplementar.....	193
6.3 Limites e perspectivas para a reprodução .....	194
<b>7 CONSIDERAÇÕES FINAIS.....</b>	<b>199</b>
<b>REFERÊNCIAS.....</b>	<b>208</b>
<b>APÊNDICE A – LEVANTAMENTO DO ESTADO DA ARTE .....</b>	<b>226</b>

## 1 INTRODUÇÃO

Em março de 2020, quando a pandemia de Covid-19 começou a se espalhar pelo mundo, ativistas pelo uso de dados abertos de pronto detectaram um problema: a precariedade nos registros de casos e mortes em decorrência da doença. No Brasil, onde as informações são inicialmente registradas em âmbito municipal e estadual, cabe ao Ministério da Saúde sistematizar e reunir os números. Por uma série de fatores que serão explicados ao longo da tese, incluindo a inépcia do governo federal, percebeu-se que não se poderia contar com o poder público para ter registros precisos e localizados sobre a pandemia, visto que esse divulgava apenas informações pouco granulares por estados da federação.

As limitações nas informações e a ausência de perspectiva de melhora provocaram uma contagem parcialmente independente de casos e mortes decorrentes da Covid-19 – parcialmente porque ainda existia um conhecimento de base que estava atrelado aos órgãos estaduais de Saúde – por parte de iniciativas da sociedade civil, como o *Brasil.io*<sup>1</sup>, ainda em março de 2020; e do consórcio de veículos da imprensa brasileira<sup>2</sup>, formado no início de junho daquele ano por conta de um “apagão de dados” registrado no Ministério da Saúde. Outras iniciativas pontuais, como a *Open Knowledge Brasil (OKBr)*, criaram formas de monitorar a qualidade da informação pública. A OKBr, por exemplo, fundou o *Índice de Transparência da Covid-19*<sup>3</sup>, cuja proposta era classificar o nível de abertura dos dados de estados e posteriormente de capitais.

Em um país de dimensões continentais como o Brasil, em que há heterogeneidade entre municípios e regiões, desde o começo era fundamental que houvesse dados específicos e detalhados sobre a situação da doença, de modo que orientasse o poder público a tomar decisões conscientes e embasadas em evidências. Como se pôde perceber no país, parte dos prefeitos e governadores levaram a situação a sério, sugerindo medidas de contenção do espalhamento do vírus, como o isolamento social e o uso de máscara. O presidente Jair Bolsonaro, por sua vez, preferiu atacar a ciência e o jornalismo, o qual repetidas vezes acusou de propagar “*fake news*”, optando por estimular o uso de medicamentos sem eficácia. Como percebemos em um levantamento desenvolvido no começo da pandemia

---

<sup>1</sup> Disponível em: <https://brasil.io/covid19/>

<sup>2</sup> Disponível em: <https://g1.globo.com/politica/noticia/2020/06/08/veiculos-de-comunicacao-formam-parceria-para-dar-transparencia-a-dados-de-covid-19.ghtml>

<sup>3</sup> Disponível em: <https://transparenciacovid19.ok.org.br/>

(GEHRKE; BENETTI, 2021), a política e a desinformação sobre a Covid-19 possuem relações intimamente imbricadas – do corpus formado por 407 unidades de conteúdo falso que analisamos naquele trabalho, um quarto (25,5%) pertencia a esse campo. Além disso, uma análise qualitativa dos materiais mostrou que o presidente, seus familiares e apoiadores se esforçaram para jogar a culpa em prefeitos e governadores pelo impacto econômico negativo gerado pelo fechamento dos estabelecimentos comerciais para conter o avanço da doença.

Desde o começo da pandemia, estive muito próxima dos números relacionados à Covid-19. Além de atuar localmente como integrante da Rede de Pessoas Embaixadoras da *Open Knowledge Brasil* desde 2019, organizando eventos anuais como o *Open Data Day*, fui uma das voluntárias no estado do Rio Grande do Sul pelo *Brasil.io*. Durante seis meses, eu e outras duas colegas coletamos diariamente os dados de casos e mortes divulgados pela Secretaria Estadual de Saúde (SES/RS); para obter os números em nível nacional, cerca de 40 voluntários trabalham todos os dias durante a pandemia. No caso do RS, percebemos, inclusive, avanços na abertura dos dados: enquanto os primeiros boletins eram divulgados em formato de arquivo fechado (.pdf) e posteriormente em formato online, exigindo técnicas de raspagem e organização, no segundo semestre de 2020 apresentava a possibilidade de fazer o download de planilhas em formato aberto e não proprietário (.csv). Até aquele momento, pode-se dizer que os dados estavam *disponíveis*, mas não eram *acessíveis* porque não eram oferecidos em formato aberto, que permitia o reuso das informações e a produção de novas análises. O *Brasil.io*, por sua vez, preocupou-se com a abertura desde o início, explicando, adicionalmente, a metodologia<sup>4</sup> de coleta nos bancos de dados públicos.

Uma das primeiras e principais lições que tirei de minha experiência com os dados abertos e a pandemia de Covid-19 foi justamente a precariedade na coleta e na abertura das informações – no primeiro ano, os problemas se resumiam aos números de casos e mortes; no segundo, as questões se estenderam aos registros de vacinação. Enquanto mediador qualificado da realidade, caberia ao jornalismo explicar às pessoas o que estava acontecendo e como interpretar toda a situação. Ou seja, era preciso que os jornalistas olhassem com ceticismo para as estatísticas, compreendendo seus contextos de produção e disponibilidade. Mais do que explicar conceitos como a média móvel, da qual muito se

---

<sup>4</sup> Disponível em: <https://github.com/turicas/covid19-br/blob/master/api.md>



ouviu falar, colocar os números em contexto exigia que os jornalistas se aproximassem da cobertura de ciência. Ciência, neste caso, em dois sentidos: de divulgação científica, visto que a pandemia está ligada à área da saúde, e de método, uma vez que os jornalistas precisaram compreender a provisoriedade da ciência e o caráter cumulativo deste tipo de conhecimento, sem que haja uma resposta definitiva e repentina para tudo.

Diante da provisoriedade da situação, comecei a observar que, quando acionavam os bancos de dados públicos para fazer reportagens, os jornalistas precisavam, por exemplo, expor as limitações daqueles registros. Precisavam explicar que os números não eram detalhados ou que poderia haver atrasos na inserção das informações em sistemas públicos, o que por tabela geraria uma distorção temporária nas análises. Em alguns casos, quando não havia condições de usar algumas estatísticas, os jornalistas de dados indicavam, nas notícias, por que resolveram excluir uma ou outra unidade da federação da análise, ou por que tais informações deveriam ser observadas com cautela pelo leitor. Estavam, então, explicando seu método de trabalho, informando quais eram as fontes consultadas, o quão confiáveis eram tais fontes e como se chegou a um ou outro resultado.

Dessa forma, ofertavam ao público uma espécie de bula sobre como ler o seu trabalho – dito de outra forma, estavam oferecendo **transparência**. O jornalismo, que muitas vezes se restringe à cobrança por abertura direcionada ao poder público, estava finalmente divulgando informações – também de interesse público – que tinham a ver com o seu método. É provável que a pandemia tenha acelerado o processo de transparência metodológica no jornalismo, ou mesmo que isso tenha acontecido especificamente com o Jornalismo Guiado por Dados porque, como veremos a seguir, este tipo de conjunto de práticas e saberes é historicamente constituído a partir de uma abordagem científica.

Posta a complexidade da pandemia, que incluem o excesso de informação e desinformação, há uma impressão de sobrecarga de conteúdo no ambiente digital (BOCZKOWSKI, 2021), o que por vezes torna difícil saber onde a confiança deve ser depositada. Conforme o *Digital News Report* (NEWMAN *et al.* 2021), a confiança nas notícias subiu três pontos percentuais no Brasil em 2021, segundo ano de pandemia, em comparação a 2020. Significa que 36% das pessoas de uma amostra representativa da população brasileira declaram confiar no jornalismo – o índice mais alto havia sido registrado em 2018, antes das eleições presidenciais, com 41%. Cabe lembrar que, durante a posse presidencial, em janeiro de 2019, apoiadores de Bolsonaro entoaram gritos de “Facebook” e

“WhatsApp” como provocação aos jornalistas que faziam a cobertura no local<sup>5</sup>, como se as plataformas fossem neutras e o recebimento de informações ganhasse mais qualidade sem a mediação do jornalismo.

Jornalistas e pesquisadores, então, buscam olhar para a transparência como um valor fundamental do jornalismo que pode fortificar ou até mesmo restabelecer o elo entre o público e os profissionais do campo. Repórter experiente e premiada da *Folha de S.Paulo*, Patrícia Campos Mello (2020, s. p.) compara a transparência a uma versão jornalística do “visite a nossa cozinha”, em que os leitores são convidados a entender os meandros da produção.

Karlsson (2022) detecta três ondas de transparência no jornalismo, sendo que a atual e terceira – a partir de 2016 – possui relação com o escrutínio e o questionamento às instituições sociais estabelecidas, o que se enquadra no cenário de desinformação e desconfiança do público. Assim, a transparência no jornalismo seria uma resposta a essa hesitação. O autor, contudo, demonstra certo ceticismo quanto à capacidade do jornalismo de abrir novas informações aos leitores, visto que o cenário de produção e distribuição noticiosa é atualmente permeado por plataformas digitais e constante interferência algorítmica, inviabilizando a abertura integral de um processo mais amplo.

Para contornar essa questão, autores falam sobre a transparência algorítmica (DIAKOPOULOS, 2019), ou seja, a possibilidade de verificação de informações estruturadas computacionalmente. Como parte do escopo do jornalismo digital, o Jornalismo Guiado por Dados apresenta características que o permitem alocar recursos humanos e tecnológicos em direção à transparência, a exemplo da arquitetura da Web, estruturada por meio de hiperlinks (MIELNICZUK, 2003). Assim, soluções pontuais, como a disponibilidade de fontes originais, poderiam ser implementadas sem grande dificuldade técnica.

De forma resumida, a transparência no jornalismo tem o sentido de abertura, de “visibilidade gerenciada estrategicamente” (KARLSSON, 2022, p. 78) e de redução de assimetria das informações (KARLSSON, 2010; 2020; 2022), visto que o jornalista tem acesso, em primeira mão, a dados que a audiência não possui; desse modo, autores como Kovach e Rosenstiel (2011; 2014) preconizam que sejam encontradas formas de o jornalismo compensar esse aspecto, revelando informações sobre fontes e métodos de apuração. No Jornalismo Guiado por Dados, isso inclui, por exemplo, oferecer hiperlink para

---

<sup>5</sup> Disponível em: <https://noticias.uol.com.br/politica/ultimas-noticias/2019/01/01/para-reporter-da-globo-apoiadores-de-bolsonaro-gritam-whatsapp-e-facebook.htm>

bancos de dados originais e, quando empregada alguma linguagem de programação, o código-fonte que conduziu a coleta, a análise e os resultados. Nesta tese, defendo que os produtos jornalísticos precisam demonstrar não só aspectos operacionais, mas também organizacionais, revelando formas de financiamento e políticas editoriais que guiam a tomada de decisões, como demonstrado na metáfora do iceberg (**Figura 7**).

Esta tese traz elementos de transparência encontrados em conteúdo informativo de Jornalismo Guiado por Dados (JGD). Para tanto, parte de uma base teórica que entende o jornalismo como forma (PARK, 1940; MEDITSCH, 1992; 1997) ou formas (NIELSEN, 2017) de conhecimento e o JGD como conjunto de saberes e práticas que se aproxima da ciência pela via do método (GEHRKE, 2017; TRÄSEL; GEHRKE, 2021), premissa básica de seu mais antigo predecessor, o Jornalismo de Precisão (MEYER, 2002).

Minha **hipótese** é de que, conectada com a verdade e a objetividade, a transparência torna-se um valor fundamental para a crescente prática do Jornalismo Guiado por Dados, o que se materializa em elementos cuja demonstração se acentua em períodos de acontecimentos significativos para a humanidade, como a pandemia de Covid-19. Na esteira dessa hipótese, tenho como **objetivo principal** identificar e analisar os elementos de transparência acionados em conteúdo informativo (notícias, reportagens e projetos) de JGD publicado por jornais de referência<sup>6</sup> e nativos digitais<sup>7</sup> brasileiros. Os **objetivos específicos** são: a) discutir o Jornalismo Guiado por Dados como forma capaz de gerar um conhecimento específico, ligado à ciência; b) compreender o modo como a transparência se apresenta neste tipo de jornalismo, relacionando-a com a ética jornalística e como estratégia de legitimação; c) sistematizar os parâmetros de transparência verificados em nível organizacional e em nível operacional; e d) elaborar um protocolo de identificação da transparência e uma proposta metodológica de aplicação para o JGD. Para testar minha hipótese e cumprir os objetivos relacionados a ela, emprego uma combinação de metodologias úteis para o campo do jornalismo digital e suas características inovadoras: o estudo de casos múltiplos (STAKE, 2006) e o estudo de caso como ilustração (MACHADO; PALACIOS, 2007).

---

<sup>6</sup> A imprensa de referência é formada por jornais consolidados no cenário brasileiro há décadas. Em geral, é composta por empresas que publicaram inicialmente um produto impresso, migrando posteriormente para a internet, com uma versão digital – o que ocorreu a partir de 1995. São representantes de um período em que o conteúdo era apenas transferido de sua versão impressa para a online (MIELNICZUK, 2003).

<sup>7</sup> Nativos digitais são os veículos jornalísticos cujo surgimento ocorreu na Web, por meio da criação de um site de notícias, sem experimentação prévia no jornal impresso. Desta forma, sua linguagem já nasce adaptada para o uso de hipertexto. No Brasil, a criação de empresas duradouras neste segmento se acentua a partir dos anos de 2010.

Nomeado desta forma na primeira década de 2000 e praticado internacionalmente em especial a partir de 2010, com o aumento das políticas de transparência pública e disponibilidade de bancos de dados, o Jornalismo Guiado por Dados tem recebido atenção crescente na academia e nas redações (DE LIMA SANTOS, 2019), mas a exploração da transparência neste tipo de jornalismo se resume a poucos trabalhos internacionais, como o de Zamith (2019), e a um limitado número de teses e dissertações no Brasil, que em geral discutem a transparência sob o ponto de vista da legislação, como ocorre por meio da Lei de Acesso à Informação (LAI)<sup>8</sup> e é demonstrado no estado da arte previamente apresentado na qualificação de tese (APÊNDICE A).

Esta tese está estruturada da seguinte forma: apresenta uma introdução, seguida de três capítulos teóricos, um capítulo metodológico e um capítulo de contribuição, encerrando com as considerações finais, referências e materiais complementares. A seguir, apresento, resumidamente, cada uma dessas partes, de modo que a introdução sirva como guia para o leitor.

No **capítulo 2**, discuto o cenário contemporâneo do excesso de informação (BOCZKOWSKI, 2021) e desinformação (WARDLE; DERAKHSHAN, 2017), acentuado por conta da pandemia de Covid-19 e de políticos negacionistas como Donald Trump, nos Estados Unidos, e Jair Bolsonaro, no Brasil, que acusam a imprensa de fabricar “*fake news*”. Como prática social qualificada que busca fazer a intermediação entre os acontecimentos e o público, o jornalismo profissional tem de lidar com esse contexto, e uma das formas de fazê-lo é oferecer informações rastreáveis ou replicáveis, como ocorre por meio de práticas como o *fact-checking* e o Jornalismo Guiado por Dados. Um dos pontos mais críticos

---

<sup>8</sup> A Lei de Acesso à Informação (LAI) – Lei 12.527 foi sancionada em maio de 2011 e vigora desde 2012 no Brasil. É inspirada em leis semelhantes ao redor do mundo, como a *Freedom of Information Act* (FOIA), dos Estados Unidos. A lei regulamenta o direito de informação garantido pela Constituição Federal brasileira de 1988 em seu artigo quinto, reforçando a publicidade de informações como regra e o sigilo como exceção. Além disso, está estruturada em duas vertentes: ativa e passiva. Em sua vertente ativa, determina que o poder público em diferentes âmbitos (Executivo, Legislativo e Judiciário) e esferas (federal, estadual e municipal) tornem abertas informações de interesse público. Em sua vertente passiva, possibilita que qualquer pessoa cadastre uma solicitação para obter informações ainda não divulgadas via Serviço de Informação ao Cidadão (SIC) – ferramenta disponível na sede do órgão público em questão, e digital (e-SIC), por meio dos portais de transparência. Isso ocorre via Serviço de Informação ao Cidadão (SIC) – ferramenta normalmente disponível em estrutura física e digital (e-SIC), por meio dos portais de transparência. Quando um pedido é cadastrado, há o prazo de 20 dias para retorno sobre a demanda, período prorrogável por mais dez dias. Havendo negativa de abertura das informações, é possível recorrer. Embora possa ser utilizada por qualquer pessoa, a vertente passiva da LAI é bastante acionada pelos jornalistas de dados brasileiros para desenvolver uma investigação inédita sobre um tema.

sob este aspecto tem a ver com a distribuição das notícias, que ocorre principalmente por meio das redes sociais digitais e aplicativos de mensagens, mesmos espaços por onde acontece a circulação de desinformação.

No **capítulo 3**, situo o Jornalismo Guiado por Dados no contexto moderno da sociedade em rede (CASTELLS, 2016) e do jornalismo em rede (HEINRICH, 2011), como parte integrante do jornalismo digital (WAISBORD, 2019; ZELIZER, 2019) em um cenário de transformação (DEUZE; WITSCHGE, 2020) e de uso das plataformas digitais (BELL; OWEN, 2017). Visto que o desenvolvimento de práticas sociais não é uniforme no mundo, trago contribuições de estudiosos para o desenvolvimento do campo do jornalismo digital no Brasil, como Mielniczuk (2003) e Barbosa (2007). Ciente de que o Jornalismo Guiado por Dados é classificado como o uso dos saberes sociais associados à computação desde a coleta até a visualização de dados (TRÄSEL, 2014), ainda destaco a aproximação deste jornalismo com o método científico a partir da formulação de hipóteses, passando pela busca e análise de dados, finalizando com a abertura metodológica pela via da transparência.

O **capítulo 4**, por sua vez, serve de espaço para discussão do que considero três princípios norteadores do jornalismo: a verdade, a objetividade e a transparência, e como esses pilares estão conectados. Discuto a verdade funcional, provisória e verificável (CORNU, 1994) característica do jornalismo e tomada como base para práticas que tenham a objetividade no método (KOVACH; ROSENSTIEL, 2014; SPONHOLZ, 2010) como ponto de partida, aproximando o fato real (GUERRA, 2008) do relato jornalístico (FRANCISCATO, 2005; LISBOA; BENETTI, 2015). A transparência (KARLSSON, 2010; 2020; 2022; MEIER, 2009), por sua vez, está associada à abertura de métodos e decisões editoriais no jornalismo. Por isso, classificamos e discutimos a transparência em duas principais vertentes: como **operadora de uma prática ética** e como **recurso discursivo de legitimação**. É perceptível, ainda, o papel proeminente da transparência entre os jornalistas de dados brasileiros (GEHRKE, 2020) e de outras nacionalidades (BEILER; IRMER; BRENDA, 2020), para quem este princípio assegura maior proximidade com o método científico e inspira confiança e credibilidade junto à audiência, também sendo capaz de combater a desinformação.

Para examinar como os elementos de transparência aparecem no Jornalismo Guiado por Dados a partir da cobertura de jornais brasileiros sobre a Covid-19, dedico o **capítulo 5** à metodologia, aplicando o estudo de casos múltiplos (STAKE, 2006) e o estudo

de caso como ilustração (MACHADO; PALACIOS, 2007) para formar um corpus empírico inédito. Dessa forma, mapeio 15 elementos de transparência em nível organizacional e em nível operacional ao examinar 11 veículos de comunicação brasileiros (de referência e nativos digitais), exemplificando cada um desses achados.

Os resultados que emergem do corpus são fundamentais para a elaboração do **Protocolo de Transparência no Jornalismo Guiado por Dados**, destinado à análise de produtos jornalísticos já publicados. Esse protocolo se origina no mapeamento de elementos efetuados no percurso metodológico, e congregam uma série de perguntas que servem como guia para pesquisas futuras, em que parâmetros de transparência podem ser avaliados individualmente e até mesmo quantificados. O Protocolo é a primeira de duas amplas contribuições que realizo no **capítulo 6** da tese, que culmina com a proposta de **Método para o Jornalismo Guiado por Dados com foco na transparência e na reprodutibilidade**, de modo que possa ser implementado em iniciativas jornalísticas orientadas por dados.

Mesmo ciente das limitações da transparência no jornalismo, entendo que a detecção e a proposta de 15 elementos no JGD operam em consonância com o que aponta Karlsson (2022): para que as expectativas do público e o cumprimento do que essa audiência espera saber estejam alinhados, é necessário que sejam estabelecidos guias que possam ser testados e cumpridos. Nas **considerações finais** desta tese, busco pontuar que o trabalho serve como fundamento e ponto de partida para novos estudos junto ao público, de modo que se possa entender o que funciona e o que não funciona nos nichos quando o tema é transparência.

A transparência, acredito, torna-se ainda mais necessária e evidente em momentos de crise e de grandes acontecimentos, como a pandemia de Covid-19. Por isso, tem potencial para levar adiante, e para outras circunstâncias e cenários, os aprendizados desta cobertura, que poderá servir como divisor de águas no modo como o jornalismo comunica o seu método ao leitor. A tendência é que essas práticas sejam apreendidas e incorporadas no dia a dia. Neste sentido, a construção de um protocolo busca munir futuros pesquisadores que pretendem acompanhar o andamento da transparência metodológica e organizacional nos produtos jornalísticos, bem como testar novos métodos em suas atividades.

## 2 INFORMAÇÃO E DESINFORMAÇÃO NA COBERTURA DA COVID-19

A rápida disseminação geográfica do vírus Sars-CoV-2 fez com que, em 11 de março de 2020, a Organização Mundial de Saúde (OMS) declarasse a Covid-19 uma pandemia. Desde então, rumores, dados pseudocientíficos e conselhos (WAISBORD, 2020) circulam particularmente no ambiente virtual, disputando espaço com informações confiáveis difundidas por entidades científicas e reproduzidas pela imprensa. Tão sintomático quanto o fato de a OMS ainda ter apontado a Covid-19 como uma “massiva infodemia”, expressão ajustada para “desinfodemia”<sup>9</sup> por Posetti e Bontcheva (2020, p. 2), é a indicação da Rede Internacional de Fact-checking (IFCN, na sigla em inglês) para receber o Prêmio Nobel da Paz devido aos serviços prestados no combate à desinformação nesse período. Fundada em 2015, a IFCN teve como uma de suas principais ações a aliança entre 99 veículos de 70 países para monitorar conteúdos falsos relacionados à Covid-19<sup>10</sup>.

Neste capítulo, trago elementos que auxiliam a compreender o cenário de desinformação no qual estamos imersos, além de discutir as dimensões conceituais dos termos normalmente associados à difusão de conteúdos mentirosos ou de baixa qualidade. Antes do aprofundamento, reúno dados sobre consumo de informações pelas redes sociais digitais. Apresento, ainda, uma rápida contextualização histórica sobre o uso insidioso da ciência para espalhar informações contrárias aos seus princípios. Embora a Covid-19 pareça um caso sem precedentes, estratégias para ameaçar a credibilidade de estudiosos e de pesquisas científicas foram registradas anteriormente na história, em especial por defensores do tabaco, que questionavam a relação entre o fumo e a incidência de câncer. Por fim, abordo a Covid-19 como um acontecimento jornalístico de nosso tempo – e que, por isso, guarda particularidades no modo como ocorre sua cobertura – e sistematizo as

---

<sup>9</sup> O termo infodemia (*infodemic*), segundo a OMS, está relacionado à rápida profusão de diferentes tipos de conteúdo acerca de um tema de interesse público, a exemplo da pandemia. Este cenário gera confusão e comportamentos de risco, visto que o público apresenta dúvidas na confiabilidade das informações. Posteriormente ao uso da expressão “massiva infodemia” pela OMS, Posetti e Bontcheva (2020, p. 2) passaram a empregar o termo “desinfodemia” (*disinfodemic*), que combina infodemia e desinformação. Para as autoras, a desinformação relacionada à Covid-19 é particularmente perigosa porque apresenta impacto imediato sobre todas as pessoas, sendo mais tóxica e mortal do que outros temas. Embora a OMS tenha utilizado *infodemia* para resumir em uma metáfora o rápido espalhamento da desinformação relacionada à pandemia, autores como Simon e Camargo (2021) criticam a utilização acrítica do termo. Entre as justificativas, os autores argumentam que a informação não se espalha da mesma forma que um vírus, gerando o risco de simplificação desse processo e do modo como a desinformação de fato é gerada e se espalha. Assim, dizem que metáforas de simples adoção têm potencial para causar mais confusão, e nem sempre são úteis para descrever um fenômeno.

<sup>10</sup> Disponível em: <https://piaui.folha.uol.com.br/lupa/2021/01/21/ifcn-nobel-da-paz/>

principais limitações e os desafios desta época, que apresenta o *fact-checking* e seu princípio de transparência metodológica, similar ao que ocorre no Jornalismo Guiado por Dados, como uma das principais respostas do jornalismo à desinformação.

Ao longo da tese, emprego o termo **desinformação** (GEHRKE; BENETTI, 2021; POSETTI; BONTCHEVA, 2020; TANDOC JR; LIM; LING, 2018; WARDLE; DERAKHSHAN, 2017) para me referir a um tipo de fraude informativa que distorce a realidade ou remove o seu contexto. Em geral, este tipo de conteúdo enganoso é criado e distribuído para atingir a reputação de pessoas e de instituições e costuma ser motivado por questões ideológicas e econômicas (TANDOC JR; LIM; LING, 2018). Pode ser viabilizado por meio de diferentes formatos (texto, áudio, foto, vídeo e outros) e circula principalmente através da internet, especialmente por meio das redes sociais digitais e dos aplicativos de mensagens.

Como veremos adiante, este fenômeno toma forma com um tipo de conteúdo falso – as chamadas “*fake news*” –, que pode ser considerado um gênero dentro do amplo escopo da desinformação, e é frequentemente associado à esfera política (DOURADO, 2020). Visto que as notícias produzidas no jornalismo estão ancoradas em uma realidade factual, o termo *fake news* torna-se um oxímoro – a rigor, não existem *notícias falsas* – incapaz de descrever um fenômeno amplo e complexo como a desinformação. Soma-se a isso o esvaziamento do termo, que além de munição para questionar a credibilidade da imprensa tornou-se sinônimo de mentira (EGELHOFER *et al.*, 2020) e passou ser adotado em circunstâncias cotidianas quaisquer. Reforço, ainda, que a desinformação é diferente de erro jornalístico (não intencional, presume-se) e do jornalismo de baixa qualidade, que inclui o emprego de métodos de apuração pouco claros e reprodução acrítica de discursos oficiais. Do jornalismo, espera-se o comprometimento com a verdade (KOVACH; ROSENSTIEL, 2014) e com a informação qualificada (REGINATO, 2019) e contextualizada<sup>11</sup> (LÜCKMAN, 2020).

No passado, o ruído informacional fez com que se espalhassem boatos sobre a suposta existência de alimentos com contornos de animais. Na Idade Média, os relatos de viagem do cavaleiro Sir John Mandeville surpreenderam pelo ineditismo – que

---

<sup>11</sup> Tomo como base a definição de Lückman (2020, p. 193), para quem a contextualização “[...] busca situar o acontecimento jornalístico dentro do recorte da realidade à qual pertence, com o estabelecimento do maior número possível de conexões entre esse acontecimento e os elementos relevantes a ele relacionados, partindo dos aspectos singulares e identificando informações conexas, pertinentes e consistentes que contribuam para ampliar a compreensão crítica do tema, possibilitando, assim, a produção de conhecimento.”



posteriormente mostrou-se não factível: ao passar pela Índia, supostamente encontrara árvores cujos frutos assemelhavam-se a pequenos carneiros, como se estivessem cheios de carne e sangue. Conhecido como o Cordeiro Vegetal da Tartária, este tipo de animal-planta (que na verdade nunca existiu) não foi o único a se espalhar pela Idade Média nem o único a ser tomado como verdadeiro inclusive por botânicos e outros estudiosos até ser desmentido no Século XVII (O'CONNOR; WEATHERALL, 2019).

Um dos motivos pelos quais o mito de que animais cresciam em árvores durou quatro séculos era a confiança atribuída às pessoas que o espalharam. Afinal, parecia um tanto óbvio crer em um viajante que observava tudo com seus próprios olhos e relatava seus registros. Para O'Connor e Weatherall (2019), o que explica a adesão das pessoas ao boato do cordeiro não tem relação com a botânica ou o mundo natural, mas é puramente um fenômeno social relacionado ao modo como as crenças são obtidas e se espalham. Situações como essa geram questões sobre como as crenças são formadas, espalhadas e até mesmo mantidas sem evidência científica, e que isso não necessariamente depende da tecnologia, ainda que esta tenha potencial para ampliar os canais de divulgação.

Embora pareça absurda nos dias de hoje, a adesão ao cordeiro vegetal não é muito diferente do caso do *boimate* registrado no Brasil dos anos de 1980, com o agravante de que o erro foi propagado inclusive pela imprensa, ainda que de forma não intencional. Ao final de abril de 1983, a revista *Veja* publicou um texto intitulado “Fruto da carne”, em que divulgava a suposta descoberta de dois biólogos alemães que teriam fundido as células de um boi com as de um tomateiro, gerando um produto híbrido, o *boimate* – um fruto que teria o gosto de filé com tomate. A reportagem na revista brasileira tomava como base um texto publicado pela *New Scientist* como piada de 1º de abril – fazer essas brincadeiras em alusão ao Dia da Mentira era algo tradicional na imprensa britânica, mas foi tomado como verdadeiro pela equipe da *Veja*. Quase um ano depois, a revista publicou uma errata.

No contexto brasileiro atual, a desinformação faz parte de um cenário político amplo que inclui a ascensão de Jair Bolsonaro à presidência. Em outubro de 2018, a *Folha de S.Paulo* revelou que empresários pró-Bolsonaro planejavam bancar, uma semana antes do segundo turno das eleições, uma operação de disparo massivo de mensagens via *WhatsApp*<sup>12</sup> para atacar seus opositores políticos e beneficiar o então candidato. Tratava-se de uma ação irregular porque, no país, não é permitido o financiamento de campanhas por

---

<sup>12</sup> Disponível em: <https://www1.folha.uol.com.br/poder/2018/10/empresarios-bancam-campanha-contra-o-pt-pelo-whatsapp.shtml>

empresas, apenas pessoas físicas. Após a publicação, a jornalista Patrícia Campos Mello, autora da reportagem na *Folha*, foi alvo de ataques misóginos e de ameaças à sua integridade e a de sua família. Os pormenores deste caso foram narrados por Mello (2020) no livro *A Máquina do Ódio*.

A tese de Dourado (2020) também indica que Bolsonaro foi o principal beneficiado por narrativas de desinformação nas eleições presidenciais de 2018, visto que maior parte das mensagens fraudulentas demonstravam um sentimento de antipetismo e pró-Bolsonaro. Para chegar a essa conclusão, a autora inicialmente mapeou 346 textos classificados como falsos por cinco projetos de checagem, selecionando 57 para aprofundamento. Dentro desse recorte, percebeu que 1.073 contas em redes sociais lideravam o processo de espalhamento de desinformação, que atingiu 4 milhões de compartilhamentos nos meses anteriores às eleições de 2018. Entre os resultados da pesquisa, percebeu que humanos, e não robôs de contas automatizadas, estão entre os principais propagadores de conteúdo falso, e que tal conteúdo transita de forma orgânica por sites de rede social e outras plataformas, principalmente *WhatsApp* e *Facebook*.

Ainda que em outro contexto, situação semelhante foi diagnosticada por estudo nosso (GEHRKE; BENETTI, 2021), que analisou 407 textos ligados à Covid-19 e classificados como falsos por agências de *fact-checking*. Esse material circulou por diferentes espaços simultaneamente, e por isso observamos um total de 493 ocorrências nas plataformas de redes sociais. Mais da metade (65,11%) desses episódios ocorreu no *Facebook*, sendo que o *WhatsApp* aparece na sequência, com 15,82% dos casos. Mesmo com a ressalva de que existe um projeto entre as empresas de tecnologia e as agências de checagem para a verificação de conteúdo, o *Facebook* se destaca como um dos espaços por onde circula significativa parcela da desinformação existente.

Também nessa linha, estudo de Rossini *et al.* (2020) utilizando o *Facebook* e o *WhatsApp* sugere que as pessoas mais engajadas em discussões políticas tendem a compartilhar desinformação com mais frequência, ainda que de forma acidental. Na pesquisa, percebeu-se que o risco de compartilhar conteúdo falso sem intenção é minimizado pelo nível de educação – quanto mais alto o grau de estudo, menor a propensão.

São também o *Facebook* e o *WhatsApp* os maiores aliados dos brasileiros na obtenção de informações, segundo dados do *Digital News Report* (NEWMAN *et al.*, 2021). A rede social tem 141,4 milhões de usuários no Brasil, e é utilizada como fonte por 47% dos brasileiros. O aplicativo de mensagens, por sua vez, faz parte da rotina de 120 milhões de

pessoas no país, servindo de fonte para 43% da população. Além disso, o relatório apontou que o celular é o dispositivo mais empregado para a busca de informações, sendo apontado por 77% dos entrevistados. Na sequência aparece o computador, com declaração de 36% das pessoas sobre o acesso por meio desse dispositivo.

A significativa adesão aos celulares nos hábitos de consumo dos brasileiros apresenta uma situação bastante particular no Brasil: a internet móvel. Um dos aspectos que pode contribuir para a alta busca por notícias em redes sociais digitais e aplicativos de mensagens é o fato de que as operadoras de telefonia móvel costumam comercializar pacotes de acesso gratuito a aplicativos de redes sociais, fazendo com que não sejam incluídos nos planos de dados de internet. Assim, há menos custos financeiros independentemente do tempo que se passa nas plataformas digitais.

Soma-se a isso o descrédito atribuído por figuras públicas, como o presidente brasileiro, à imprensa. Políticos de extrema-direita como ele – e Donald Trump<sup>13</sup> nos Estados Unidos, até o ano de 2020 – incentivam que as pessoas se afastem do jornalismo profissional, a quem denominam equivocadamente de *fake news*, e consomem conteúdo apenas nas plataformas, onde possuem controle sobre o conteúdo veiculado e comumente disparam mentiras para inflamar seus seguidores mais fiéis. Dessa forma, desviam do papel de mediador qualificado que deve ser desempenhado pela imprensa. A posse presidencial de Jair Bolsonaro, em 1º de janeiro de 2019, foi um termômetro significativo. Naquela data, apoiadores do presidente se concentraram na Praça dos Três Poderes, em Brasília, e, provocando jornalistas da Rede Globo, entoaram gritos de “WhatsApp” e “Facebook”<sup>14</sup>. Os ataques verbais de Bolsonaro à imprensa profissional também são um caso à parte: em 2020, a Federação Nacional dos Jornalistas (Fenaj) contabilizou 299 ocorrências<sup>15</sup>.

Além do Facebook e do WhatsApp, estão no radar de pesquisadores o YouTube, o Instagram e o Twitter, que possuem, respectivamente, 39%, 30% e 12% da preferência dos brasileiros no consumo de informações pela internet (NEWMAN *et al.*, 2021). Estudos como o de Machado *et al.* (2020a) mostram que o YouTube tem se tornado um ambiente propício para o espalhamento da desinformação. Com aproximadamente 120 milhões de

---

<sup>13</sup> Em janeiro de 2021, o Twitter banuiu a conta de Donald Trump após acusações infundadas sobre fraude nas eleições americanas em 2020, que culminou na vitória de Joe Biden, e após incitar a invasão ao Capitólio. A rede social entendeu que a postura do então presidente feria as regras da comunidade. O comunicado completo está disponível em: [https://blog.twitter.com/en\\_us/topics/company/2020/suspension.html](https://blog.twitter.com/en_us/topics/company/2020/suspension.html)

<sup>14</sup> Disponível em: <https://noticias.uol.com.br/politica/ultimas-noticias/2019/01/01/para-reporter-da-globo-apoiadores-de-bolsonaro-gritam-whatsapp-e-facebook.htm>

<sup>15</sup> Disponível em: <https://fenaj.org.br/nove-meses-bolsonaro-299-ataques/>

usuários no Brasil, cerca de 70% do tráfego da plataforma é resultado de recomendações. Dessa forma, os algoritmos operam de acordo com as preferências dos usuários, sugerindo outros vídeos para consumo. Empoli (2020) relata que, em 2015, vídeos conspiracionistas se espalharam no *YouTube* à época da difusão do Zika vírus, prejudicando médicos que tentavam distribuir vacinas e larvicidas para conter os mosquitos responsáveis pela transmissão. Pesquisa de Massarani *et al.* (2021) com jovens brasileiros de 15 a 24 anos de idade mostram que o *YouTube* vem sendo utilizado como mecanismo de busca quase tanto quanto o *Google*: 73% e 79% dos entrevistados, respectivamente, indicaram o seu uso.

A dependência do jornalismo em relação à distribuição de conteúdo por meio das redes sociais é um dos pontos discutidos por Bell e Owen (2017). Neste ambiente, são as plataformas que expõem as regras do jogo e condicionam a imprensa a seguir algumas premissas para ampliar o seu alcance. Assim, nota-se a utilização de títulos apelativos por parte do jornalismo para assegurar uma maior quantidade de acessos – os chamados *clickbait*s. Contudo, esta é uma estratégia perigosa porque, no ambiente digital, as pessoas estão sujeitas a consumir as notícias acidentalmente ou porque simplesmente passam os olhos rapidamente sobre seus títulos, fazendo um processo de escaneamento (COSTERA MEIJER; GROOT KORMELINK, 2015), eliminando o processo de aprofundamento. Torna-se um desafio para o jornalismo operar em um cenário onde grandes empresas de tecnologia apresentam papel proeminente, deslocando o papel do jornalista e do jornalismo em um ecossistema de produção e circulação de informações (ANDERSON; BELL; SHIRKY, 2013), sem que isso seja necessariamente algo benéfico para a sociedade como um todo.

O rastreamento de dados de usuários é um dos principais meios de monetização para as plataformas, visto que os dados coletados são utilizados para anúncios direcionados. Assim, essas empresas assumem um controle mais amplo de redes sociais e culturais, gerando também implicações para o jornalismo digital, segundo Steensen e Westlund (2021). Os autores mapeiam três principais mudanças: uma crise massiva nos modelos de negócio, uma vez que os anunciantes migraram para as plataformas digitais; ênfase no uso de dados da audiência e métricas de acesso aos sites; e alterações nos padrões de distribuição da notícia, espaço que empresas não-jornalísticas vêm dominando. O impacto dessas variáveis sobre o campo jornalístico pode ser percebido no contexto de estudos da área: tecnologia e plataformas foram os principais tópicos relacionados aos estudos de jornalismo digital em 343 publicações do periódico *Digital Journalism* de 2013 a 2019.

Segundo pesquisa de Rossini e Kalogeropoulos (2021), a maior parte dos brasileiros procurou informações sobre a pandemia de Covid-19 na televisão (82%), em redes sociais digitais como *Facebook* e *Twitter* (81%) e no *WhatsApp* (81%). Chama a atenção, neste estudo, que um quarto dos respondentes disse evitar ativamente notícias sobre a Covid-19. Essa atitude é mais recorrente entre pessoas identificadas como direita política (44%) e ocorre em menor proporção entre pessoas que se dizem de centro (33%) e de esquerda (23%). Entre os principais motivos para o afastamento deliberado das notícias está o excesso de repetição por parte da imprensa (87%), o efeito ruim sobre o humor das pessoas (66%), a falta de confiança nas notícias (57%) e a sensação de impotência diante dos fatos (53%). Embora seja difícil correlacionar de forma simplista quaisquer variáveis associadas à desinformação, é sintomático que o público de direita política identificado com o presidente – que, por sua vez, ataca a imprensa – evite consumir notícias.

Particularmente desde os anos de 2010, com a ascensão das redes sociais, o cenário da informação tornou-se complexo com a inclusão de novos atores (WAISBORD, 2020), incluindo os que trazem caos e buscam minar a confiança dos cidadãos na ciência. Veremos, a seguir, que não é a primeira vez que um caso desta natureza ocorre, e que a desinformação não é nova, mas que possui vias para se espalhar rapidamente – além das interações orgânicas, estamos falando da atuação de *bots* e *trolls* nas disputas pela verdade (KAKUTANI, 2018). O papel de atores políticos é fulcral para o entendimento das dinâmicas envolvendo a desinformação nas redes sociais digitais.

## 2.1 O complexo cenário da desinformação e suas variações

À época da disputa eleitoral para a presidência dos Estados Unidos, em 4 dezembro de 2016, Edgar Maddison Welch saiu de sua cidade, Salisbury, e dirigiu até Washington D.C. para verificar, com os seus próprios olhos, o local atribuído a uma suposta rede de pedofilia conectada à candidata Hillary Clinton. Com o intuito de salvar as crianças que supostamente eram vítimas de abuso no local – e que nunca foram encontradas –, abriu fogo contra a pizzaria *Comet Ping Pong*. Funcionários conseguiram fugir e ninguém se feriu. Entrevistado<sup>16</sup> por jornalista do *The New York Times* depois de se entregar à polícia, Welch, que é pai de duas meninas, disse que, a partir da leitura de uma sequência de artigos e uma “combinação de fontes”, encontrou evidências de que algo nefasto ocorria na pizzaria e decidiu ir até o local para conferir. Alguns dias antes, ele havia instalado serviço de internet em sua casa.

O encadeamento de desinformação que levou Welch a tomar uma medida drástica é conhecida como *Pizzagate*, uma teoria da conspiração criada a partir da correspondência de e-mail do proprietário da pizzaria acerca de um jantar para arrecadar doações para a campanha de Clinton. Tais mensagens, oriundas do chefe de campanha John Podesta, foram vazadas pelo *WikiLeaks*<sup>17</sup> e, em algum ponto, pessoas passaram a associar o uso de termos como *cheese pizza a child pornography*.

O caso *Pizzagate* ilustra com precisão o cenário caótico da atualidade, em que diversos tipos de informação e desinformação coexistem e podem ser acessados por meio de mecanismos de buscas e redes sociais digitais. As teorias da conspiração, também presentes nesse ambiente, adicionam um ingrediente perigoso a este cenário. Assim como outros tipos de desinformação, tomam como base temas que provocam revolta e ultraje. Welch declarou ter sido motivado pelo ímpeto de defesa às crianças.

As eleições estadunidenses de 2016 não foram marcadas apenas por esse caso, mas também pelo recorrente uso do termo “*fake news*” por parte de Donald Trump para atacar a credibilidade da imprensa. Depois dele, outros políticos de extrema-direita passaram a adotar discurso e estratégias semelhantes. Pesquisas mostram que conteúdo falso em geral tem mais potencial de propagação do que verdadeiro. Em cenário recente,

---

<sup>16</sup> Disponível em: <https://www.nytimes.com/2016/12/07/us/edgar-welch-comet-pizza-fake-news.html>

<sup>17</sup> Organização coordenada pelo jornalista e ativista Julian Assange, *WikiLeaks* é uma plataforma que publica informações e documentos vazados por fontes anônimas e em geral sobre temas sensíveis.

O'Connor e Weatherall (2019) mencionam que, à época da eleição de Trump, em 2016, as 20 peças de desinformação que mais circularam tiveram 8,7 milhões de likes e compartilhamentos três meses antes das eleições. As notícias compartilhadas por fontes confiáveis, por sua vez, atingiram a marca de 7,3 milhões curtidas ou compartilhamentos no Facebook. Ao que parece, a mentira tem uma narrativa convincente e é capaz de gerar engajamento.

Para Kakutani (2018, p. 11), a ascensão de Trump ao poder configura o triunfo do ataque à razão e à verdade, mas o processo começou a ser desenvolvido anteriormente: trata-se do resultado de um sentimento incubado pela extrema-direita dos Estados Unidos há anos.

Com Trump, a esfera pessoal é política e, em muitos sentidos, ele é menos uma anomalia caricata e mais um bizarro epítome de uma série de atitudes mais amplas e interligadas que corroem lentamente a verdade nos dias de hoje, desde a mistura do noticiário e da política com o entretenimento até a polarização tóxica que tomou conta da política norte-americana, passando pelo crescente desprezo populista em relação ao conhecimento especializado.

De todo modo, a candidatura de Trump só prosperou porque, conforme examinado pela autora, existe uma postura de indiferença das pessoas em relação aos fatos e à verdade factual. Ou, na pior das hipóteses, figuras como ele e outros políticos populistas são vistos como alguém que fala a verdade, como também observado por Empoli (2020).

Boatos, falsidades e manipulações foram registrados em diferentes momentos da história, como percebemos na Idade Média por meio do cordeiro vegetal da Tartária. O novo e perigoso ingrediente é o alto potencial de propagação e a velocidade com que isso ocorre por meio das redes sociais digitais e de aplicativos de mensagens. O ambiente digital também contribui para a confusão conceitual de fonte de informação, visto que receber conteúdo de pessoas próximas pode ajudar a legitimar a veracidade da informação que é compartilhada (TANDOC; LIM; LING, 2018).

Como dito anteriormente, nesta tese emprego o termo **desinformação** para designar conteúdo mentiroso que circula com o propósito de enganar e minar reputações. Entre outros fatores relativos à associação do jornalismo às falsidades que circulam na Web, concordo com Wardle e Derakhshan (2017), para quem o termo *fake news* é essencialmente raso para descrever um fenômeno complexo que abrange a poluição informacional. Assim, os autores entendem que três termos conseguem descrever com mais precisão o

que ocorre neste cenário. São eles: *misinformation* – quando conteúdo falso é compartilhado sem a intenção de enganar; *disinformation* – quando um conteúdo circula para causar danos; *malinformation* – quando uma informação verdadeira é compartilhada para causar dano, normalmente associado a algo que deveria ter permanecido na esfera privada. Em linhas gerais e de forma genérica, todos podem ser traduzidos como desinformação para o português.

Comumente acionado por políticos, o termo *fake news* pode ser entendido como um subtipo de fraude informativa que busca semelhança com as notícias (TANDOC JR; LIM; LING, 2018). Dourado (2020, p. 54) a define como “[...] um tipo de informação inverídica ou distorcida que simula uma notícia – ou novidade – para narrar fatos políticos e assim conquistar maior visibilidade no trânsito entre plataformas de mídias sociais”. Em sua percepção, o caráter urgente do jornalismo é fundamental neste caso, visto que as chamadas *fake news* imitam a ideia de novidade e de alegado interesse público. Assim, a desinformação mimetiza o formato de notícia – muitas vezes incluindo até mesmo supostas fontes –, com uso de manchetes, fotos e textos. Este tipo de conteúdo ainda apresenta, corriqueiramente, escritos como “urgente” e “compartilhe com seus contatos”. O processo de fabricação da desinformação deixa mais evidente o propósito de enganar. São indícios dessa confusão propositada, segundo Dourado (2020), o uso de *clickbait* nos títulos, montagens de vídeo ou de imagem e outras peças gráficas.

Autores como Egelhofer *et al.* (2020) observam um esvaziamento de sentido em relação ao termo *fake news*. Nota-se que há o emprego exacerbado do termo para normalmente referir-se a mentiras e incorreções de modo geral, o que ocorre, inclusive, em publicações divulgadas pela imprensa. Para as pesquisadoras, um dos problemas de o termo circular livremente no próprio jornalismo é a utilização por atores políticos, que se sentem à vontade para empregá-lo no ataque à legitimidade e credibilidade da imprensa enquanto instituição democrática.

Se, para os teóricos, o termo *fake news* é capaz de identificar um tipo de desinformação, entre a maioria das pessoas não há sinais claros. Por meio de questionários e grupos focais, Nielsen e Graves (2017) observaram que o público notadamente associa jornalismo de baixa qualidade (sensacionalismo, por exemplo), propaganda e alguns tipos de publicidade ao termo *fake news*, ainda que, em um primeiro momento, este tipo de conteúdo não seja produzido com o intuito de desinformar. Assim, os pesquisadores observaram que o público associa o termo *fake news* não só com conteúdo mentiroso, mas



com um cenário (des)informativo<sup>18</sup> mais amplo, que inclui a imprensa, os políticos e as plataformas.

Há, ainda, outras subdivisões dentro da amplitude da desinformação ou das *fake news*. Em linhas gerais, Wardle e Derakhshan (2017) e Wardle (2019) classificam a desinformação em: conteúdo fabricado, conteúdo manipulado, conteúdo impostor, falso contexto, conteúdo enganoso, falsa conexão e sátira ou paródia. Tandoc Jr., Lim e Ling (2018), por sua vez, observam a seguinte tipologia: sátira, paródia, fabricação, manipulação, publicidade e propaganda. A fabricação, afirmam os autores, é o que mais próximo poderia ser classificado como o que se conhece por *fake news*.

Em relação ao tipo de conteúdo que mais circula na pandemia, identificamos (GEHRKE; BENETTI, 2021) que, dos 300 casos analisados qualitativamente, 145 (48,43%) empregam falso contexto, ou seja, se utilizam de fato verdadeiro ou imagem genuína, mas tiram os materiais de contexto para criar uma mentira. Já o segundo tipo de conteúdo mais acionado é o conteúdo fabricado, com 86 (28,67%) ocorrências. Trata-se de conteúdo criado para enganar, incluindo o uso de instituições de pesquisa e nomes fictícios para tentar atribuir confiabilidade ao que foi escrito. Os resultados vão ao encontro do que foi detectado por Dourado (2020), que observou que o uso de fotos e vídeos funcionam como uma espécie de garantia ou evidência de que a história narrada é verossímil.

A Covid-19 não é o primeiro acontecimento científico que apresenta relação com a desinformação. Como observado por Machado *et al.* (2020b), não necessariamente existe um questionamento sobre a ciência ou os seus métodos, mas são utilizadas estratégias semelhantes para espalhar a desinformação, como a incerteza, que é inerente à ciência à medida que existe uma produção sistemática do conhecimento e da validade de teorias até que sejam substituídas por outras que deem conta de explicar melhor a realidade.

Na segunda metade do Século XX, ainda sem grande influência da tecnologia digital que conhecemos atualmente na distribuição da desinformação, o setor do tabaco descobriu na incerteza, inerente ao processo científico, uma possibilidade de explorar e

---

<sup>18</sup> Para dissociar o termo *fake news* do que se espera de um jornalismo confiável, Träsel (2019, s.p.) propõe o uso do termo pseudojornalismo para designar “textos, áudios, imagens ou vídeos que mimetizam formatos jornalísticos na tentativa de enganar o público”, ou seja, que dizem respeito aos conteúdos falsificados. O radical *pseudo* corresponde à artificialidade desse tipo de conteúdo, e o termo jornalismo, em português, abrange também formatos que, a rigor, não são somente notícia (mas também reportagem e outros formatos). Em inglês, o termo *news* é mais amplo e pode ser utilizado para se referir a outros formatos jornalísticos. Embora a argumentação do autor faça sentido, vejo o uso de pseudojornalismo com restrições por conta do emprego de jornalismo na formação da palavra, o que poderia reforçar crenças pré-existentes de que o jornalismo não opera a partir de uma verdade factual.

enfraquecer a afirmação de que fumar causa câncer. Como explicam Oreskes e Conway (2010), o objetivo era aproveitar-se das lacunas que naturalmente existem na ciência para chamar a atenção e questionar sobre os riscos causados pelo cigarro. Desde os anos de 1950, pesquisas comprovavam a relação entre o uso do cigarro e o desenvolvimento de câncer. Contudo, a indústria tabagista optou por patrocinar estudos que exploravam outros tipos de relação sem considerar o cigarro, associando o câncer à hereditariedade, por exemplo. Assim, buscou perpetuar a ideia de que não havia estudos suficientes para mostrar a relação entre fumar e desenvolver câncer.

Naturalmente, para difundir tais ideias, os empresários do setor do tabaco se aproveitaram do fato de a imprensa ter que mostrar os dois lados. “A posição da indústria era de que ‘não havia provas’ de que o tabaco era ruim, e eles sustentaram essa posição criando um ‘debate’, convencendo a imprensa de massa de que o jornalismo responsável tinha a obrigação de apresentar ‘os dois lados’ disso”<sup>19</sup> (ORESQUES; CONWAY, 2010, p. 16, tradução minha<sup>20</sup>). A cadeia do tabaco utilizou bem a estratégia de que o jornalismo deveria reportar os fatos com equilíbrio e, portanto, abrir espaço para que pudessem expor o seu lado da história. Os autores ponderam que, neste caso, não havia dois lados para mostrar; o equilíbrio era interpretado como fornecer pesos iguais aos dois lados, acima do que oferecer um peso justo aos dois lados.

Assim como qualquer outra rede social formada por seres humanos, os cientistas também são influenciados por fatores sociais e por seus pares, pois integram comunidades epistêmicas, ou seja, grupo de pessoas que estão tentando ganhar conhecimento sobre o mundo de alguma forma. Para compreender como a desinformação se perpetua, é necessário olhar para como as interações sociais mudaram ao longo do tempo, e como isso se dá na formação de crenças.

Considerando o contexto das redes sociais digitais, pessoas e algoritmos – instruções computacionais que visam atingir determinado objetivo – se misturam. Algoritmos são escritos por humanos e os sites de rede social não apresentam, com clareza, que tipos de instruções priorizam as interações. À medida que ocorre o compartilhamento de conteúdo nas plataformas, tais informações ganham relevância e se tornam mais visíveis,

---

<sup>19</sup> No original “The industry position was there was “no proof” that tobacco was bad, and they fostered that position by manufacturing a ‘debate’, convincing the mass media that responsible journalism had an obligation to present ‘both sides’ of it.”

<sup>20</sup> Todas os trechos foram traduzidos por mim. Desta forma, utilizo a expressão “tradução minha” somente neste primeiro caso. Todos os demais seguem a mesma lógica, com indicação do trecho original em nota de rodapé.

especialmente quando circulam pelas contas de influenciadores e atores políticos. Como veremos a seguir, algumas pessoas apresentam papel chave nas interações em rede.

## 2.2 Do topo para baixo: a influência de políticos e figuras públicas

Ao longo de 2020, segundo informações da agência de checagem *Aos Fatos*<sup>21</sup>, o presidente Jair Bolsonaro proferiu 915 afirmações falsas ou imprecisas sobre a pandemia de Covid-19. Em seu livro de memórias, o ex-ministro da Saúde Luiz Henrique Mandetta, à frente do cargo quando a doença começou a se espalhar pelo Brasil a partir de 25 de janeiro, disse que, ainda no início de março, o Palácio do Planalto vivia como se o novo coronavírus não existisse. Segundo Mandetta (2020), Bolsonaro nunca aceitou analisar gráficos e encarar a realidade da doença que o governo precisaria enfrentar. Isso se refletiu nas ações do presidente, que ao longo da pandemia minimizou o vírus, criticou o distanciamento social e incentivou que as pessoas tomassem medicamentos ineficazes.

Atores políticos apresentam papel fundamental nas redes e, por isso, na propagação de informação ou desinformação. Embora existam estratégias e mecanismos de propagação, estudos sugerem que o poder de influência de um indivíduo pode ser decisivo no espalhamento de conteúdo falso. Na pesquisa de Machado *et al.* (2020a), viabilizada por meio de análise de redes e análise lexical, pessoas com algum grau de autoridade social, ainda que distantes da ciência, se valeram de suas posições hierárquicas dentro de sua rede de atuação para questionar as orientações propostas por sistemas peritos como a imprensa, as universidades e organizações internacionais, respaldando teorias da conspiração.

Em sua tese, Dourado (2020) detectou que perfis pessoais são os mais relevantes na propagação de desinformação, mostrando que esse fenômeno se espalha muito mais por vias orgânicas do que automatizadas. Para embasar isso, o estudo mostrou a prevalência de perfis pessoais entre os principais disseminadores em geral e entre aqueles codificados como os propagadores mais relevantes. Ao mesmo tempo, também descobriu que, dentre os perfis pessoais mais relevantes, aqueles que não pareciam robôs (os *bots*, ou seja, contas automatizadas) foram quase cinco vezes mais frequentes do que os que pareciam robôs. Isso significa que, por mais que o clima de fraude informacional seja

---

<sup>21</sup> Disponível em: <https://www.aosfatos.org/todas-as-declara%C3%A7%C3%B5es-de-bolsonaro/>

inflamado pelo uso de robôs, *trolls* (pessoas que semeiam a dúvida para gerar discurso de ódio ou fazem provocações e espalham mentiras para se divertir) e perfis falsos, a distribuição de *fake news* sobre as eleições de 2018 no Brasil aconteceu por vias bem mais orgânicas do que automatizadas ou artificializadas.

Wardle e Derakhshan (2017) propõem que se examine os elementos da desinformação separadamente, a saber: o agente, a mensagem e os intérpretes. O agente, por exemplo, pode atuar em diferentes instâncias, como fabricação e distribuição. E este é justamente um dos principais desafios nos estudos de desinformação: compreender quais são as motivações das pessoas. Como observamos em nossa pesquisa (GEHRKE; BENETTI, 2021), não é possível afirmar que os propagadores de desinformação são também os criadores dos boatos; mas é fato que essas pessoas possuem alto grau de influência nas redes. Em relação à Covid-19, o deputado federal Osmar Terra e membros da família Bolsonaro, incluindo o próprio presidente, estão entre os que mais espalharam desinformação, conforme nosso levantamento.

Em pesquisa que examinou a desinformação sobre a Covid-19 no *YouTube*, Machado *et al.* (2020a) identificaram duas narrativas principais relacionadas à doença: uma de caráter geopolítico, culpando a China pela suposta produção do vírus como arma biológica e usando-a para fins econômicos; e a outra de caráter religioso, buscando enquadrar a pandemia em um contexto de simbologia bíblica. No caso da narrativa religiosa, uma das figuras proeminentes no espalhamento da desinformação é o pastor Silas Malafaia, líder da igreja neopentecostal Assembleia de Deus e apoiador do presidente Jair Bolsonaro. Nessa narrativa, ocorre o uso distorcido de dados e de aspectos científicos para associar a pandemia a uma praga divina. Outra pessoa que apareceu nessa pesquisa como um agente significativo na difusão e desinformação foi o jornalista Alexandre Garcia, ex-apresentador da Rede Globo. Em um dos vídeos publicados no *YouTube*, ele questiona a circulação e a permanência do vírus em ambientes tropicais como o Brasil. A marca comum entre as campanhas identificadas por Machado *et al.* (2020a) foi a minimização da pandemia. Os autores avaliam que, quando conveniente, o discurso da desinformação distorce a ciência e descontextualiza pontos de notas técnicas.

Outros estudos (MACHADO *et al.*, 2020b; SOARES; RECUERO, 2020) também demonstraram a influência de personagens como Bolsonaro no cenário de desinformação brasileiro. Ao analisar um conjunto de dados composto por 8,6 mil publicações verificadas por 70 agências de fact-checking em 134 países, Machado *et al.* (2020b) descobriram que

a desinformação sobre a Covid-19 segue padrões regionais, envolvendo questões culturais, geográficas e fatores políticos. Neste cenário, países como Brasil e Índia encontram-se mais isolados do que o resto do mundo. No caso do Brasil, o principal motivo é a insistência de Bolsonaro no uso de medidas sem comprovação científica, como os medicamentos hidroxicloroquina, azitromicina e ivermectina.

Em suas memórias, Mandetta (2020, p. 133) esclarece a insistência de Bolsonaro com a cloroquina: para o presidente, bastaria oferecer um remédio para que os trabalhadores voltassem à ativa e à movimentação da economia. “Para Bolsonaro a solução sempre foi simples: o projeto dele para o combate à pandemia é dizer que o governo tem o remédio e quem tomar o remédio vai ficar bem. Só vai morrer quem já ia morrer de qualquer maneira”.

Pesquisas como a de Rossini e Kalogeropoulos (2021) sugerem que a confiança no governo federal é maior entre os brasileiros que têm a direita como preferência política, ou seja, que estão alinhados ao governo. De forma geral, os respondentes disseram confiar mais nos cientistas para obter informações sobre a pandemia – os percentuais variaram de 70% (direita) a 90% (esquerda). Em segundo lugar, os participantes disseram confiar nas informações fornecidas pela Organização Mundial da Saúde (OMS), chegando a um percentual de 86% entre as pessoas de esquerda, com exceção dos participantes de direita, que afirmam confiar mais no Ministério da Saúde (63%) e no governo federal (63%).

Vimos, nesta seção, que os atores políticos e influenciadores apresentam papel importante na difusão de informação ou desinformação, uma vez que apresentam papel decisivo na forma como as redes estão organizadas e por conta de seu alcance na rede. A seguir, veremos quais são as principais medidas para combater a desinformação.

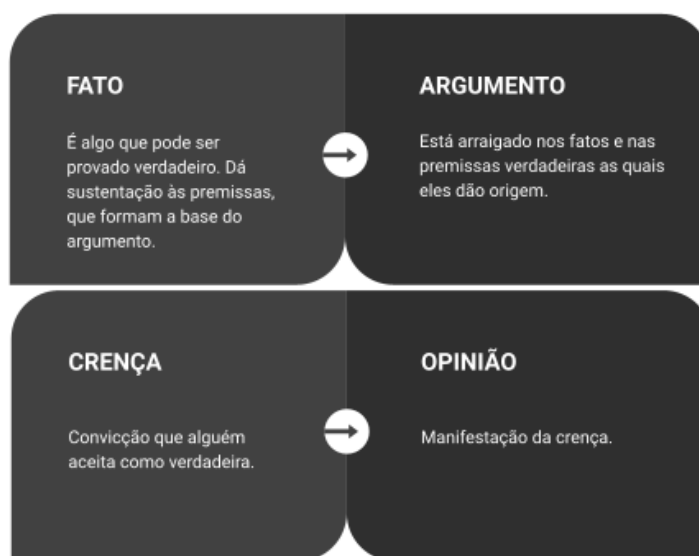
### 2.3 A efetividade da desinformação e estratégias para combatê-la

Pesquisadores de diferentes áreas do conhecimento buscam estratégias para combater a desinformação. Neste cenário, um dos maiores desafios é compreender quais são as intenções que motivam o agente criador ou disseminador do conteúdo falso, prática em geral motivada por questões ideológicas ou econômicas (TANDOC JR; LIM; LING; 2018). Além disso, este tipo de conteúdo altera a forma como a realidade é percebida e encarada pelas pessoas, visto que a imagem das instituições é afetada por esse tipo de discurso (CARNIELLI; EPSTEIN, 2019). Uma vez que existe confusão entre crenças, opiniões e argumentos, discursos que buscam desinformar estimulam que haja uma resposta emocional em torno dos acontecimentos em oposição à racionalidade.

Carnielli e Epstein (2019) definem da seguinte forma: a **crença** é uma convicção que alguém aceita como verdadeira, enquanto a **opinião** é uma manifestação dessa crença. Ambas não são argumentos. Um **fato**, por sua vez, é algo que pode ser comprovado (por exemplo, a partir de reiterados estudos científicos, como ocorre com o aquecimento global). Os fatos garantem a verdade das premissas, que formam a base do **argumento**. Assim, utilizar-se da máxima de que “esta é a minha opinião” é uma conduta desleal e completamente afastada dos fatos e da argumentação.

A **Figura 1** mostra que, enquanto os argumentos estão baseados em fatos que podem ser provados verdadeiros, as opiniões são oriundas das crenças.

**Figura 1 – Argumentos e opiniões**



Fonte: A autora (2021) com base em Carnielli e Epstein (2019).

As peças de desinformação costumam ser efetivas porque reproduzem ideias e valores que já estavam sendo compartilhados em nichos específicos (DOURADO, 2020). Além disso, nas plataformas, os algoritmos são adaptados para privilegiar a comunicação interpessoal, ampliando a interatividade entre os usuários e o reforço de crenças já estabelecidas em grupos menores. “Algoritmos se tornaram, portanto, sustentáculos de curadoria de conteúdo e fragmentação de público em mídias sociais a ponto de moldar como usuários são levados a interagir entre si e a consumir informações” (DOURADO, 2020, p. 63).

Existem diversos vieses cognitivos utilizados para perpetuar a desinformação, a exemplo do viés de confirmação, ou seja, o acolhimento de informações com as quais alguém concorda (CARNIELLI; EPSTEIN, 2019; COSENZA, 2016). Na tentativa de desvendar como as crenças são arraigadas e como se espalham na sociedade, os pesquisadores O’Connor e Weatherall (2019) buscam explicações examinando as redes sociais – não necessariamente digitais. Um dos modelos de redes trabalhado pelos autores diz respeito à predisposição de um grupo em confiar em alguém que tem as mesmas percepções que ele em algum aspecto. Por exemplo, se uma pessoa possui valores conservadores (a favor da família tradicional e contrário ao aborto, a título de ilustração) e vê que o presidente compartilha da mesma premissa, a tendência é que se confie a ele outras decisões.

Outros vieses cognitivos explorados por Carnielli e Epstein (2019) são o efeito caravana, em que as pessoas acreditam em algo porque muitas outras acreditam; a cascata repetitiva, em que o conteúdo se torna plausível por conta de repetições no discurso público; e efeito de terceira pessoa, a crença de que as mensagens têm mais efeito nos outros do que em si. Um dos exemplos mais comuns, explicam Capella, Ophir e Sutton (2018), tem a ver com o cigarro: um fumante sabe que fumar faz mal, mas tende a achar que pode fazer mal para a maior parte das pessoas que fuma, mas não para ele. Assim, os autores entendem que, para corrigir a desinformação de maneira efetiva, é preciso antes medir o conhecimento das pessoas acerca de um tema, estimulando abordagens diferentes para compreender, também, o quanto alguém se enxerga como parte do problema.

Uma das limitações, porém, é que aprendemos com os outros. Segundo O’Connor e Weatherall (2019, p. 8), a maior parte das pessoas obtêm crenças falsas nos mesmos locais onde obtêm crenças confiáveis, e por isso existe o risco de acabar apreendendo ambas. É isso que torna o *homo sapiens* o vencedor da evolução humana. “A habilidade de

compartilhar informações e influenciar as crenças dos outros é parte do que torna os seres humanos especiais.”<sup>22</sup> Não raro, como mostra Kakutani (2018), políticos populistas se utilizam do discurso de nós contra eles, atacando o que é externo a um pequeno grupo e reforçando sentimentos de nacionalismo, tribalismo e ódio aos estrangeiros.

Também há complexidade em relação aos motivos pelos quais as pessoas espalham desinformação – e não necessariamente acreditam nela. A persistência de circulação de algumas histórias, dizem O’Connor e Weatherall (2019), tem relação com o fato de algumas pessoas acreditarem que se trata de algo engraçado ou inacreditável, e por isso há o compartilhamento por conta de ironia ou sarcasmo. Além disso, há quem compartilhe um conteúdo porque o meme ou outra peça digital é capaz de refletir os seus sentimentos em relação a um tópico.

Especificamente sobre a Covid-19, Posetti e Bontcheva (2020) entendem que existem algumas estratégias para corrigir e frear a desinformação. Para as autoras, as atitudes frente a este problema de saúde pública podem ser estruturadas em quatro frentes: monitoramento e investigação (checagem e *debunking* – desmetir boatos publicamente –, por exemplo); políticas de Estado e legislação; curadoria, tecnologia e resposta econômica; e perspectiva ética, educacional e associada à credibilidade.

Para Carnielli e Epstein (2019), a avaliação de premissas, por meio do pensamento crítico, é uma forma de combate à desinformação. Nesse sentido, os autores estimulam que as pessoas identifiquem a quem interessa a notícia ou postagem, qual seu ponto central e suas conclusões, detecção de falácia no raciocínio, diferenciação entre os fatos e a interpretação dos fatos, qualidade das evidências, credibilidade da fonte, comparação com outras fontes, se a história é interessante ou impressionante demais, se alguém assina a matéria, entre outras medidas.

Em relação à Covid-19, dois estudos recentes apresentados por Pennycook *et al.* (2020) e que envolveram mais de 1,7 mil adultos nos Estados Unidos mostram que as pessoas compartilham desinformação em parte porque não refletiram suficientemente sobre a acurácia do conteúdo que leram e repassaram. Assim, os autores mostraram que, para instigar o pensamento crítico, uma medida efetiva pode ser a prática de *nudge*, ou seja, oferecer pequenos estímulos para que as pessoas reflitam antes de compartilhar um conteúdo. Ao analisar o *WhatsApp* e o *Facebook* como ambientes onde ocorre o

---

<sup>22</sup> No original “The ability to share information and influence one another’s beliefs is part of what makes humans special.”



espalhamento da desinformação, Rossini *et al.* (2020) observaram que as correções tendem a ser feitas no aplicativo de mensagens com mais frequência. Visto que o *WhatsApp* costuma ser utilizado entre contatos próximos, trata-se de um ambiente visto com mais segurança para fazer correções.

Uma das práticas mais conhecidas dentro do jornalismo como forma de combate à desinformação é o *fact-checking*, ou seja, a checagem de fatos. Esta atividade consiste na verificação de afirmações proferidas por pessoas públicas e ocorre principalmente na comparação de dados apresentados discursivamente com evidências presentes em bancos de dados públicos e fontes oficiais. A origem dessa prática é atribuída ao jornalista Brooks Jackson, da CNN, em 1991, nos Estados Unidos, em função da cobertura eleitoral (AMAZEEN, 2015; GRAVES, 2016). Pode ser considerada um gênero jornalístico que transita entre a área política e acadêmica, visto que envolve pessoas públicas e se trata de uma prática com método específico de verificação. Como veremos mais adiante, a transparência, associada à objetividade metodológica, é um dos princípios normativos desse tipo de jornalismo (SEIBT, 2019), semelhante ao que ocorre no Jornalismo Guiado por Dados (GEHRKE, 2018b).

Ao analisar as mudanças no jornalismo, Seibt (2019) identificou que o *fact-checking* e suas práticas representam um período pré-paradigmático de mudanças estruturais no jornalismo. Examinando o cenário brasileiro, a autora observou que, em 2016, esta prática estava de certa forma restrita aos nativos digitais. Atualmente, porém, empresas de referência também passaram a se envolver em iniciativas de verificação, em especial diante das circunstâncias especiais impostas pela pandemia. Em janeiro de 2021, a Rede Internacional de Fact-checking (*International Fact-checking Network – IFCN*) contava com as seguintes empresas brasileiras credenciadas e ativas: *Agência Lupa*, *Aos Fatos* e *Estadão Verifica*. A *Agence France-Presse* (AFP), que também opera no país, está vinculada à rede por meio de seu registro na França.

Embora seja cada vez mais recorrente inserir automatização de processos na verificação de fatos, o que é feito por iniciativas como *Radar Aos Fatos* e projetos ligados às universidades, como o detector de *fake news*<sup>23</sup> criado pela Universidade de São Paulo (USP) e a Universidade Federal de São Carlos (UFSCar), o olhar humano presente nas estratégias do *fact-checking* acaba se tornando uma medida mais efetiva quando a principal

---

<sup>23</sup> Disponível em: <http://nilc-fakenews.herokuapp.com/>

forma de causar desinformação está associada à exclusão ou alteração de informações conceituais ou mesmo recirculação de notícias antigas. Nem sempre ferramentas automatizadas conseguem executar a mesma tarefa de maneira precisa, visto que há nuances no processo.

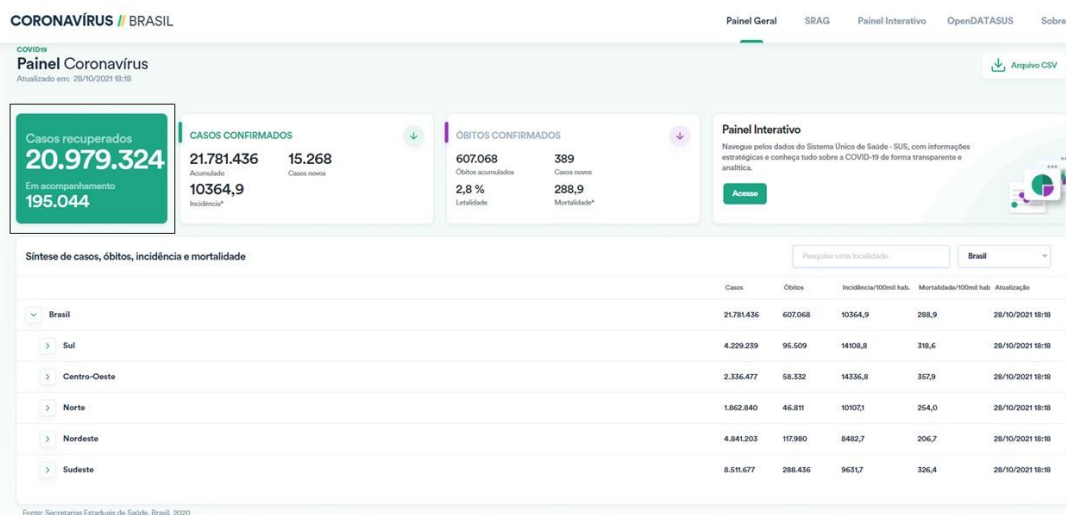
Na avaliação de Waisbord (2020), o *fact-checking* vem cumprindo um papel importante para corrigir informações, mas é preciso ir além de apontar os erros e efetivamente corrigir a desinformação. Para ele, é preciso que haja um esforço coletivo na desinformação relacionada à Covid-19, semelhante ao que aconteceu com o tabaco – depois de muitas discussões, finalmente percebeu-se que havia uma estratégia forte para tentar minar a credibilidade da ciência e até mesmo do jornalismo. Uma mobilização que, para o autor, envolve a comunidade médica, corporações de mídia e cidadãos em geral, compreendendo por que as pessoas acreditam em teorias da conspiração e corrigindo a desinformação.

Uma das soluções apresentadas por Green e Donahue (2018) no combate à desinformação é criar esforços para reduzir ou evitar a desinformação antes que ela atinja o público; neste caso, a postura pode levar, porém, à problematização de questões relacionadas à censura. De qualquer forma, o *fact-checking* tem capacidade para apresentar um efeito pedagógico profilático, fazendo com que, ao saber da existência de tais práticas, os políticos se desanimem a mentir na esfera pública. Até lá, cabe a iniciativas como a checagem de fatos e outras ações da sociedade civil a fiscalização do que vem sendo feito pelo poder público. No caso da Covid-19, veremos, a seguir, que existem problemas desde o início da pandemia em relação à qualidade das informações divulgadas pelo poder público.

## 2.4 Transparência pública e abertura de dados sobre a Covid-19

O Painel Coronavírus<sup>24</sup>, mantido pelo Ministério da Saúde, reúne, desde o início da pandemia, informações pouco detalhadas sobre a Covid-19. Apresenta número de casos e mortes de forma agregada por estados e regiões, o que prejudica análises aprofundadas e a reutilização de dados para implementar análises locais e regionais. Dessa forma, o conhecimento sobre o espalhamento do vírus em regiões periféricas acaba negligenciado, visto que não se tem uma fotografia precisa sobre o que está acontecendo. Além disso, o painel do governo federal prioriza, em seus dados, a quantidade de casos recuperados, como é possível visualizar no destaque a seguir:

**Figura 2 – Painel Coronavírus, vinculado ao Ministério da Saúde**



Fonte: *Ministério da Saúde*. Captura de tela feita em 29 out. 2021.

Embora a estatística destacada pelo governo federal seja a de recuperados (em destaque na **Figura 2**), essa informação numérica ignora qualquer tipo de cuidado necessário após o desenvolvimento da doença. O número de recuperados corresponde aos que não morreram, mas pouco ou nada se sabe sobre o estado de saúde dessas pessoas, incluindo sequelas. A narrativa descolada da realidade em relação ao impacto da Covid-19 foi

<sup>24</sup> Disponível em: <https://covid.saude.gov.br/>

recorrente no governo. Em agosto de 2020, quando o país batia a marca de 113 mil mortes, o governo federal realizou um evento chamado *Brasil vencendo a Covid-19*<sup>25</sup>.

A opacidade dos dados no Brasil, ainda em março de 2020, motivou a criação de iniciativas independentes para monitorar e mapear os casos de Covid-19, partindo principalmente de jornalistas de dados e desenvolvedores. Conhecida pelo viés colaborativo, a comunidade brasileira de Jornalismo Guiado por Dados articulou, em março de 2020, diferentes iniciativas para aprimorar os dados divulgados pelo poder público – geralmente em formatos pouco acessíveis – e organizar um banco de dados colaborativo sobre o tema. O *Brasil.io*<sup>26</sup>, site que reúne bancos de dados em formato aberto, mobilizou um time de 40 voluntários para extrair, diariamente, números apresentados em boletins de Covid-19 pelas secretarias estaduais de saúde. Como não há um sistema unificado entre as secretarias nem o acesso ao formato aberto em muitas delas, o trabalho dos voluntários é “libertar” os dados de um formato fechado e inacessível para um formato aberto, que pode ser lido por computador.

A ação conjunta permitiu que o *Brasil.io*, liderado pelo desenvolvedor e ativista Álvaro Justen, oferecesse planilhas atualizadas diariamente, com dados por estado e por município. Os voluntários trabalham com uma metodologia conjunta para seguir a mesma forma de coleta; além disso, há um constante processo de verificação das informações – incluindo uma espécie de verificação dupla independente antes de abastecer o sistema conjunto com novos dados –, de modo a evitar erros. Toda a metodologia e o código-fonte estão disponíveis no *GitHub*<sup>27</sup>. Como uma das integrantes da equipe de voluntários do *Brasil.io*, percebi as principais dificuldades em relação ao acesso e à obtenção de dados sobre Covid-19 no Brasil.

Outra iniciativa pontual partiu do jornalista de dados Marcelo Soares, com o abastecimento de informações municipais sobre Covid-19 apresentadas no *Lagom Data*<sup>28</sup>. O site criado por Soares realiza análises periódicas sobre o avanço da pandemia em capitais e cidades do interior. Tanto essa quanto a iniciativa do *Brasil.io* mostraram-se ainda mais relevantes em tempos de “apagões de dados oficiais” ocorridos em junho e novembro de 2020, tornando-se fonte para os jornalistas e oferecendo a série histórica (por município)

---

<sup>25</sup> Disponível em: <https://noticias.uol.com.br/saude/ultimas-noticias/redacao/2020/08/22/com-113-mil-mortes-bolsonaro-promove-evento-brasil-vencendo-a-covid-19.htm>

<sup>26</sup> Disponível em: <https://brasil.io/dataset/covid19/boletim>

<sup>27</sup> Disponível em: <https://github.com/turicas/covid19-br>

<sup>28</sup> Disponível em: <https://www.lagomdata.com.br/coronavirus>

que o poder público não foi capaz de fornecer. Embora o ideal fosse a liberação de microdados por parte dos gestores desde o começo da pandemia isso não ocorreu; o esforço de organização e compilação de dados por jornalistas e desenvolvedores auxiliou e ainda auxilia jornalistas e pesquisadores em seu dia a dia de trabalho, em que é necessário encontrar dados precisos de forma rápida.

Em abril de 2020, a *Open Knowledge Brasil*, organização que incentiva o uso de dados e conhecimento aberto, passou a monitorar<sup>29</sup> a transparência dos estados – e posteriormente de capitais – em relação à abertura de dados sobre casos e mortes por Covid-19. Em 4 de março de 2021, a organização também se engajou no monitoramento da quantidade de vacinas aplicadas, cobrando mais transparência nas variáveis divulgadas. A *Open Knowledge* e outras seis instituições divulgaram uma carta aberta<sup>30</sup> em que concluíram que 70% das informações sobre a vacinação estavam incompletas. O painel de vacinação do governo federal<sup>31</sup> apresentava limitações nos dados, com um nível amplo de agregação e pouca efetividade na separação das informações.

Ao longo de 2020, ocorreram pelo menos dois apagões de dados significativos, conforme mencionado anteriormente: em junho e novembro. Em 8 de junho de 2020, logo após o primeiro sumiço das informações, os veículos jornalísticos *G1*, *O Globo*, *Extra*, *Estadão*, *Folha* e *UOL* uniram esforços para coletar os dados sobre casos e mortes por Covid-19, semelhante ao que faz o *Brasil.io* e seu time de voluntários. A prática busca suprir uma demanda interna na produção de reportagens, e não culmina em um banco de dados que pode ser acessado publicamente. Em 5 de novembro de 2020, mais uma vez houve problemas<sup>32</sup> de abastecimento de dados no portal do Ministério da Saúde, prejudicando o conhecimento da população em relação à doença. Os acontecimentos ao longo de 2020 e do início de 2021 em relação ao monitoramento de dados podem ser assim resumidos:

---

<sup>29</sup> Disponível em: <https://transparenciacovid19.ok.org.br/>

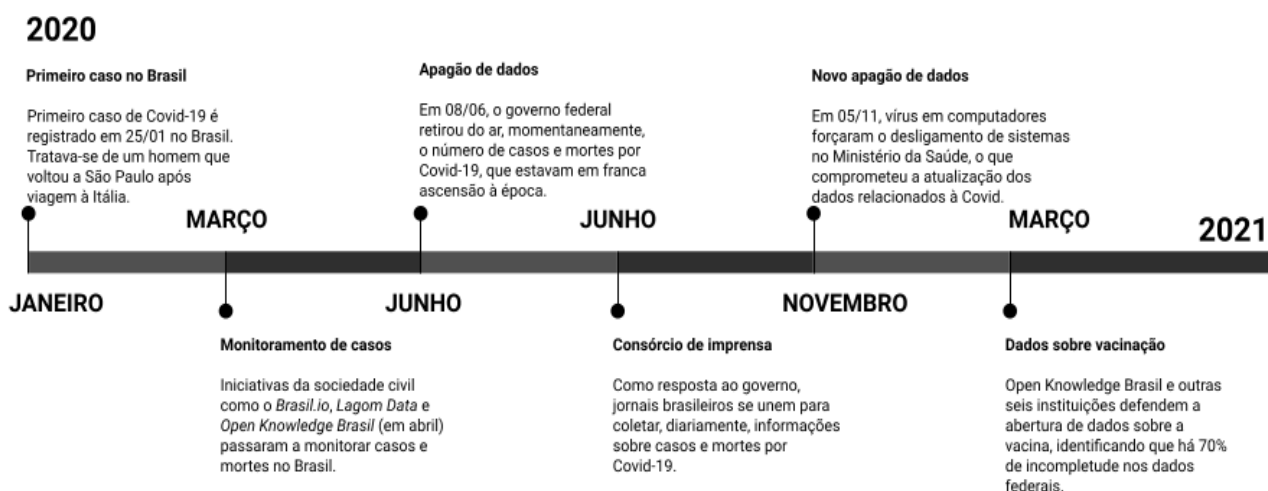
<sup>30</sup> Disponível em: <https://www.ok.org.br/noticia/okbr-e-mais-seis-organizacoes-denunciam-falta-de-transparencia-publica-sobre-a-vacinacao/>

<sup>31</sup> Disponível em: [https://viz.saude.gov.br/extensions/DEMAs\\_C19Vacina/DEMAs\\_C19Vacina.html](https://viz.saude.gov.br/extensions/DEMAs_C19Vacina/DEMAs_C19Vacina.html)

<sup>32</sup> Disponível em:

<https://g1.globo.com/ciencia-e-saude/noticia/2020/11/06/ministerio-enfrenta-incidente-em-sistemas-que-afeta-atualizacao-de-casos-e-mortes-da-covid-19.ghtml>

**Figura 3 – Linha do tempo sobre a abertura e o apagão de dados**



Fonte: A autora (2021)

A **Figura 3** envolve o período até março de 2021, referente ao primeiro ano da pandemia, pois foi nesse intervalo que as questões problemáticas relacionadas aos dados se acentuaram, embora não tenham se encerrado por completo mesmo no segundo ano da pandemia. Além dos problemas pontuais, a falta de segurança nos dados armazenados pelo Ministério da Saúde foi observada meses após o início da pandemia. Em 10 de dezembro de 2021, por exemplo, a pasta informou que foi vítima de um ataque cibernético, impossibilitando, entre outras questões, o registro de notificações de casos de Covid-19 pelos estados e prejudicando a emissão da carteira de vacinação dos cidadãos por meio do aplicativo ConecteSUS. A disponibilização dessas informações foi retomada dias depois - o aplicativo, por exemplo, voltou a funcionar após cerca de duas semanas.

A falta de clareza nos dados e os problemas na acessibilidade das informações por parte do público pode abrir espaço para o contraditório e, conseqüentemente, para a circulação de conteúdo falso. Smith, Cubbon e Wardle (2020) observam que o déficit nos dados sobre vacinação é terreno propício para o desenvolvimento da desinformação. Quando há incertezas e lacunas em relação a determinado tema, a tendência é que as pessoas tentem preencher esses espaços com as informações disponíveis no momento, que nem sempre são as melhores. A Covid-19 tem sido o principal tema abordado pela imprensa nacional e internacional desde o seu surgimento, e o fato de a pandemia configurar

um acontecimento jornalístico de grandes proporções tem relação com o modo como essa cobertura ocorre.

## 2.5 A pandemia de Covid-19 como acontecimento jornalístico

Monitorar o andamento da pandemia de Covid-19 e informar sobre seus desdobramentos tem sido o principal dever da imprensa desde que a doença passou a ser conhecida no mundo, ao final de 2019, e principalmente após a declaração de suas proporções globais pela OMS, em março do ano seguinte. Mais do que nunca, o jornalismo tem um papel social a cumprir, o que deve ser feito com a publicação de informações verdadeiras, precisas, qualificadas e verificadas.

Até o dia 26 de outubro de 2021, de acordo com o monitor da *Universidade Johns Hopkins*, havia 244,3 milhões<sup>33</sup> de casos de infecção no mundo, e um total de 4,9 milhões de mortes. Os Estados Unidos eram até então o país com o maior número. Sozinho, contabilizava cerca de 45,5 milhões de contaminados, além de 737,6 mil mortes. Desde o início da pandemia, o Brasil trilha caminho semelhante, com 21,7 milhões de casos e 605 mil mortes. O cenário começou a melhorar com o avanço da vacinação nos países. Segundo este mesmo monitor, 6,8 bilhões de doses foram administradas. É provável que a Covid-19 seja o principal fenômeno de saúde pública registrado neste século e, por consequência, o maior acontecimento jornalístico do período.

Imprevisível e improvável, o acontecimento é o que foge à normalidade da rotina. Em suma: “[...] o acontecimento é imprevisível, irrompe acidentalmente à superfície epidérmica dos corpos como reflexo inesperado, como efeito sem causa, como puro atributo” (RODRIGUES, A. 2016, p. 53). Também pertence a Rodrigues (2016, p. 51) a clássica definição de que “É acontecimento tudo aquilo que irrompe na superfície lisa da história [...]”

A perspectiva do tempo está intimamente atrelada ao pano de fundo do acontecimento. Para Quéré (2005, p. 61), o acontecimento é começo e ruptura, pois cria ações e apresenta caráter inaugural. “O acontecimento introduz uma descontinuidade, só perceptível num fundo de continuidade.” Neste caso, não estamos falando de acontecimentos artificiais e fabricados (pseudo-acontecimentos) com a ideia de novidade, como ocorre com as coletivas de imprensa, ou mesmo de acontecimentos midiáticos (tal como as

---

<sup>33</sup> Disponível em: <https://coronavirus.jhu.edu/map.html>

Olimpíadas e eventos semelhantes), que dependem da cobertura da imprensa para que tomem forma (PONTES; SILVA, 2010).

Para compreender onde o acontecimento se insere, é preciso retomar a ideia de um mapa comum de significados, ou seja, dos valores compartilhados em sociedade. O jornalismo tem papel fundamental na comunicação dos acontecimentos, visto que é capaz de legitimar quadros de referência já estabelecidos. Em texto que propõe pensar o jornalismo como acontecimento, Benetti (2010, p. 147) diz que, o que se pode questionar, é a suposta homogeneidade entre os mapas de significados pertencentes a diferentes grupos. “É um movimento circular que começa nos interesses da fala institucionalizada do poder e retorna a esses mesmos interesses, sempre mediado pelos procedimentos técnicos que legitimam a prática discursiva do jornalismo.” Ao acionar conceitos de forma reiterada, o jornalismo torna-se um acontecimento. “O jornalismo é, assim, *um acontecimento* quando posso tomá-lo como registro dos valores hegemônicos de uma sociedade em uma época” (BENETTI, 2010, p. 162).

É importante reconhecer a pandemia de Covid-19 como acontecimento porque sua cobertura escapa ao convencional e requer outro tipo de tratamento jornalístico. Por se tratar de um encadeamento serial (QUÉRÉ, 2005), demanda interpretação e explicação por parte do jornalismo, o que pode ser viabilizado pelo contexto dos fatos. Tudo isso ocorre em um cenário de ciberacontecimentos (HENN, 2013, p. 32), em que não é possível descolar os fatos da cultura contemporânea inerente às redes digitais, onde misturam-se esfera pública e privada. O autor compreende o ciberacontecimento como “[...] uma nova modalidade no processo de acontecer” em que o jornalismo perde a centralidade em detrimento das características da rede.

Rodrigues (2016) identifica três categorias principais de registro para determinar o acontecimento referencial: 1) o excesso – envolve massacres, casos de embriaguez e outros; 2) a falha – engloba acidentes naturais, inundações, terremotos e outros casos que demonstram a falha no funcionamento de alguma instituição; 3) a inversão – compreende o oposto do que é esperado em situações normais (exemplo: o ladrão que rouba um carro e depois o devolve ao proprietário). Além dos acontecimentos referenciais, o autor propõe a categoria dos meta-acontecimentos: aqueles que só existem se forem enunciados, que estão diretamente condicionados pelo discurso. Ao fazer o relato dos fatos, o jornalismo, enquanto gênero discursivo (BENETTI, 2008), produz um novo acontecimento. O meta-



acontecimento é o evento que permite ir além de suas próprias dimensões, transpondo a singularidade do fato relatado (BENETTI; STORCH; FINATTO, 2011).

A cobertura da pandemia de Covid-19 aciona o *acontecimento jornalístico referencial* marcado pela *falha* e pelo *excesso* e o *meta-acontecimento*. O espalhamento de um vírus inicialmente desconhecido em nível global é capaz de, em um primeiro momento, pôr em xeque instituições consolidadas como a própria ciência. De pronto surgem perguntas sobre a previsibilidade de eventos como esse, assim como uma possível imunização ao vírus por meio da vacina. Cobra-se que a ciência tenha todas as respostas: que vírus é este e como se espalhou? Quais são suas características e sintomas? Como ninguém foi capaz de rastreá-lo com antecedência? Qual é o tratamento? Existe cura? E, à medida que a quantidade de infectados e mortos avança, o funcionamento de outras instituições é igualmente colocado sob suspeita. A política, materializada por seus governantes e pelos serviços essenciais oferecidos à população, é alvo de questionamentos ao passo que a precariedade nas estruturas de atendimento à saúde pública é escancarada.

Com o avanço do acontecimento e as mudanças que causa no dia a dia, outro tipo de cobertura jornalística é exigido. Assim, a notícia já não é mais sobre o surgimento do vírus propriamente dito, mas, por exemplo, como os governos em diversas partes do mundo tratam o caso em seus sistemas de saúde e quanto ao isolamento social obrigatório. Deste modo, surge o meta-acontecimento: uma narratividade que ocorre no seu próprio tempo. Em casos como esse, o acontecimento narrado já não é o primeiro acontecimento, mas configura uma narrativa em cima de outra narrativa (PONTES; SILVA, 2010). O eixo de significação de um meta-acontecimento é longo (BENETTI; STORCH; FINATTO, 2011), o que lhe confere potencial para diversas abordagens temáticas. Como diz Genro Filho (2012), os acontecimentos devem ser compreendidos como processos incompletos e narrados pelos jornalistas com a abertura de sentido necessário para que sejam estabelecidas as relações dialéticas entre as categorias singular, particular e universal.

O acontecimento jornalístico ocorre à medida que esta prática profissional é capaz de tratar fenômenos que geram experiência compartilhada, organiza a experiência temporal do homem contemporâneo e produz supostos consensos (BENETTI, 2010). Por vezes desnorteadas pelo imprevisível, as pessoas recorrem ao jornalismo para buscar explicações e inserir o fato em uma série temporal. Desta forma, a ideia de contexto é fundamental para o jornalismo, de modo que se busca estabelecer uma explicação causal para os fatos. Assim, a cobertura jornalística, ainda que concentrada no presente, busca

elementos do passado para explicar o andamento das ocorrências. O acontecimento, portanto, faz surgir uma dimensão de passado. Ao noticiar a cobertura de Covid-19, por exemplo, o jornalismo recorre a outras epidemias ou pandemias enfrentadas no mundo. Olhar para trás, neste caso, dá a ideia de haver elementos para compreender o que ocorre no momento. “Há coisas que acontecem, e que julgávamos impossíveis de acontecer, porque excediam o pensável ou o nosso sentido do possível. Ao acontecerem, somos obrigados a reconhecer que havia possibilidades, potencialidades ou eventualidades” (QUÉRÉ, 2005, p. 63).

Acontecimentos com as dimensões desta pandemia trabalham, sobretudo, com valores caros à sociedade. Visto que nascimento e morte estão entre os acidentes-limites em relação às ocorrências (RODRIGUES, A. 2016), há um retorno aos processos ritualísticos (BENETTI, 2010, p. 155) nesse tipo de cobertura. “Os movimentos da vida nos trazem de volta a grandes questões existenciais: a morte, a violência, o amor, o acaso, o imponderável.”

No jornalismo, quanto menos previsível for, maior a probabilidade de um acontecimento se tornar notícia e integrar o discurso jornalístico (RODRIGUES, 2005). Neste caso, a crítica de Benetti (2010) à prática jornalística é a de naturalização de problemas persistentes, como a fome, as desigualdades e questões de injustiças sociais. De modo que são tratados como algo que simplesmente ocorre na sociedade, dificilmente são abordados pelos veículos de comunicação. “Dito de forma crua, são fatos clinicamente percebidos como ordinários e comuns e que, por isso, não alcançam os requisitos que lhes permitiriam ocupar o estatuto de acontecimento jornalístico” (BENETTI, 2010, p. 146).

É provável que um acontecimento das dimensões da Covid-19 seja especialmente impactante na ordem das experiências vividas. Se o acontecimento geralmente é elevado ao caráter de fato (QUÉRÉ, 2005), em situações de alto impacto opera na ordem da experiência. Visto que o acontecimento é um fenômeno que acontece a alguém, as pessoas normalmente se encontram bastante marcadas pelas ações que ele gera. Se, via de regra, as pessoas se lembram do que estavam fazendo no dia 11 de setembro de 2001, quando as torres gêmeas do *World Trade Center* foram alvo de um ataque terrorista em Nova York, é muito provável que, em alguns anos, as pessoas registrem suas próprias lembranças e experiências da época em que a pandemia acometeu a humanidade. A experiência humana permite que o acontecimento seja memorizado, rememorado e reconstruído. A

memória, por sua vez, depende de contextos sociais e como tais recordações são construídas e vivenciadas em grupos sociais (BABO-LANÇA, 2011).

Sob o meu ponto de vista, tratar a cobertura da pandemia de Covid-19 como acontecimento jornalístico é fundamental para compreender que papel esta prática profissional desempenha em situações extremas cujo desenrolar dos fatos introduz novas possibilidades interpretativas, que dialogam com presente, passado e futuro (QUÉRÉ, 2005). O jornalismo, por vezes, é acusado de ter memória curta e de não fazer revisões e coberturas sistemáticas sobre um mesmo tema, preocupando-se, apenas e essencialmente, com os pequenos acontecimentos que surgem a cada dia. A cobertura de Covid-19 tende a ser, assim, cercada de lições e aprendizados em todo o mundo. Uma pandemia é um exemplo preciso de uma ocorrência que transborda espacial e temporalmente, como diz Quéré.

O acontecimento ganha significado social sobretudo ao dialogar com outros acontecimentos (BABO-LANÇA, 2011). Além disso, a pandemia evoca um caso que põe em diálogo fatos jornalísticos e científicos (ARQUEMBOURG, 2011). Ao descrever estudos científicos – que provêm de observação, descrição, medição e comparação relacionada a um objeto –, os jornalistas organizam tais informações em um plano narrativo com a descrição dos fatos. Quando estudou o tremor de terra que atingiu e devastou Lisboa em 1755, Arquembourg (2011) observou o fato de os jornais e gazetas da época realizarem uma cobertura contínua do acontecimento, fazendo articulações entre jornalismo e ciência para comunicar os fatos a um público leigo e a um público erudito.

Para compreender e explicar ao público os objetos de sua apuração, jornalistas que cobriram a Covid-19 se viram obrigados a se aproximar da cobertura especializada que demandam a ciência e a saúde pública, aprendendo a ler e compreender estudos científicos, bem como interpretar dados de contaminação e mortes pelo vírus. No Jornalismo Guiado por Dados, percebeu-se a existência de um processo colaborativo na coleta e disponibilização dos números, que na cobertura se transformaram em infográficos e mapas, recursos que auxiliam a destrinchar os fatos.

No entanto, mesmo ancorado em uma verdade verificável e sob o ímpeto de informar de modo qualificado (REGINATO, 2019), o jornalismo pode gerar um desserviço à população à medida que reproduz discursos de forma acrítica ou que prioriza aspectos comerciais em detrimento da informação qualificada. Um ano depois da pandemia e batendo recorde de mortes, um grupo intitulado *Médicos pela Vida – Covid-19* publicou na *Folha de S.Paulo* e em outros jornais do país, em 23 de fevereiro de 2021, um anúncio de

meia página em que defendia o suposto tratamento precoce, conforme propagado pelo presidente, com a utilização de medicamentos como cloroquina, ivermectina, zinco e vitamina D. Mesmo tendo publicado diversos textos jornalísticos em que desmente a existência de tratamento precoce, a *Folha* aceitou estampar o anúncio. Como uma espécie de mea-culpa, publicou, também, uma reportagem<sup>34</sup> em que o desmente, informando que o suposto tratamento precoce não está amparado pela ciência. A publicação do anúncio<sup>35</sup> também foi tema da coluna *Ombudsman* da *Folha*, à época assinada por Flavia Lima, que criticou a postura de dualidade do jornal.

Em circunstâncias como essa, em que a saúde pública está em jogo e deve ser pensada sobre como é retratada pela imprensa, ocorre o que fala Henn (2013, p. 44) sobre o ciberacontecimento: “O jornalismo titubeia nos seus modos tradicionais e nos vínculos empresariais que fundamentam boa parte dos seus enquadramentos”. Embora à primeira vista possa parecer uma decisão simples – separar o conteúdo jornalístico da publicidade do jornal –, o jornalismo apresenta dificuldades em contemplar a complexidade exigida de uma narrativa como a Covid-19.

Estudo de Soares e Recuero (2021, s.p.) mostra que o jornalismo profissional também contribuiu em alguma medida para a construção de narrativas de desinformação durante a pandemia. Por utilizar títulos declaratórios, produções jornalísticas foram empregadas para corroborar conteúdos enganosos no Facebook. Isso ocorreu especialmente por meio da associação de URLs jornalísticas a títulos de baixa qualidade ou com problemas de acurácia informativa. “Consequentemente, os jornalistas precisam adotar cuidados extras na produção das notícias, especialmente as manchetes, que serão a parte mais visível das matérias nas redes sociais.”<sup>36</sup>

Jornalistas também precisam prestar atenção às falsas equivalências e com o tipo de fontes que utilizam. Assim como no caso do tabaco, explicado neste capítulo, há algumas circunstâncias em que não existem dois lados a serem contemplados – pense, por exemplo, sobre o aquecimento global, em que negacionistas das mudanças climáticas são corriqueiramente entrevistados para dar voz “ao outro lado”. Green e Donahue (2018) chamam a atenção para a dificuldade de corrigir informações já propagadas, inclusive

---

<sup>34</sup> Disponível em: <https://www1.folha.uol.com.br/equilibrioesaude/2021/02/grupo-de-medicos-defende-tratamento-precoce-sem-eficacia-contra-covid-19-em-jornais.shtml>

<sup>35</sup> Disponível em: <https://www1.folha.uol.com.br/colunas/flavia-lima-ombudsman/2021/02/mais-uma-vez-a-cloroquina.shtml>

<sup>36</sup> No original “Consequently, journalists ought to take extra care when producing news, especially headlines, which will be the most visible part of the stories on social media.”

quando os erros partem até mesmo dos estudos científicos – quando uma abordagem equivocada é espalhada, a correção acaba não sendo tão efetiva porque muitas vezes não alcança as mesmas proporções do primeiro fato noticiado.

Apesar de as decisões editoriais normalmente sequer passarem pelas equipes de reportagem, os jornalistas parecem estar de certa forma conscientes dos desafios que a era da informação, mas também da desinformação, interpõem à profissão. Em estudo com 405 jornalistas dos Estados Unidos, Vu e Saldaña (2021) concluíram que, desde 2016, época da eleição do presidente Donald Trump, os repórteres estão mais inclinados a adotar estratégias de transparência e *accountability* em suas reportagens para auxiliar no combate à desinformação, a exemplo de checagem cruzada de fontes, uso de tecnologias recentes no *fact-checking* e limite de uso de fontes anônimas. Jornalistas que acreditam que a desinformação é uma ameaça à democracia tendem a abrir mais os seus processos, apostando na transparência metodológica. Como veremos nos próximos capítulos, o Jornalismo Guiado por Dados, situado no escopo do jornalismo digital, é uma dessas formas.

### 3 DIGITAL E EM REDE: O LUGAR DO JORNALISMO GUIADO POR DADOS

O Jornalismo Guiado por Dados, que será estudado ao longo da tese e definido neste capítulo, é um conjunto de saberes e práticas aplicado à reportagem e inspirado nas ciências sociais, alicerçado na verdade, na objetividade e na transparência ao oferecer uma abordagem contextual e duradoura em suas produções. Por conta dessa referência, destaca-se pela aproximação com o método científico, o que se viabiliza por meio de uma série de procedimentos que se inicia com o teste de hipóteses, passando pela busca e análise de dados, e culminando com a abertura metodológica pela via da transparência. As bases epistemológicas do JGD foram desenvolvidas principalmente nos Estados Unidos, a partir da segunda metade dos anos de 1960, com o Jornalismo de Precisão e a Reportagem Assistida por Computador (RAC).

Nesta seção, procuro situar essa forma de jornalismo no contexto moderno da sociedade em rede (CASTELLS, 2016) e do jornalismo em rede (HEINRICH, 2011), no qual também se desenvolve o jornalismo digital. Compreendo o Jornalismo Guiado por Dados como parte do jornalismo digital, sendo constantemente afetado por fluxos e interações descentralizadas que ocorrem por meio das redes. Tais redes são mediadas pelas plataformas (BELL; OWEN, 2017) e caracterizam um cenário de jornalismo em transformação (DEUZE; WITSCHGE, 2020) e pós-industrial (ANDERSON; BELL; SHIRKY, 2013).

O jornalismo em rede dá nome ao fenômeno que identifica mudanças significativas na busca e disseminação de informação, mostrando que existem novos padrões no sistema de produção jornalística que se inserem não mais em um contexto físico determinado, mas em fluxos organizados por meio da internet. Heinrich (2011) entende que as transformações estão alicerçadas principalmente em dois processos: globalização – por ela entendida como o aumento de conexões em uma rede – e avanços tecnológicos por meio da digitalização. A horizontalidade que toma o lugar da verticalização nos fluxos de produção e distribuição de conteúdo gera o que Barbosa (2013, p. 33) denomina de “continuum multimídia de cariz dinâmico” para os produtos jornalísticos disponibilizados na Web. Dessa forma, novos arranjos se configuram, e os desafios profissionais incluem aprender a incorporar o conteúdo gerado pelos usuários na prática jornalística, ou mesmo medir a influência das plataformas no processo de produção e distribuição das notícias (BELL; OWEN, 2017).

Neste ponto, cabe ressaltar que o desenvolvimento do jornalismo digital não é estanque nem uniforme, por isso não acontece de maneira única e homogênea ao redor do globo. Países do Norte global, cujas sociedades geralmente têm acesso mais rápido à tecnologia em detrimento de outras comunidades, costumam ter acesso a uma ampla gama de tecnologias e possibilidades. Os demais países, por sua vez, adaptam e ressignificam as primeiras diretrizes a partir de suas próprias ferramentas e possibilidades. Ciente desse cenário, trago, nesse capítulo, aspectos sobre o desenvolvimento do jornalismo digital enquanto campo teórico e apresento brevemente seu desenvolvimento no Brasil para então passar ao Jornalismo Guiado por Dados. Além disso, é preciso considerar que o contexto informacional e em rede gera transformações anteriores à própria prática jornalística. Está relacionada, por exemplo, aos modos de trabalho dos profissionais – a redação jornalística já não é a única configuração viável (DEUZE; WITSCHGE, 2020) –, pois conformações de trabalho descentralizado – a exemplo de investigações em rede – também são estabelecidas.

Deuze e Witschge (2020) argumentam que, por se tratar de um processo, o jornalismo está em constante transformação. Para os autores, as práticas atuais da área carregam três principais marcadores: o modo como a tecnologia digital afeta o campo, incluindo o papel de plataformas, agregadores e distribuidores para além da imprensa de referência; a integração de áreas editoriais e de marketing, trazendo novas reflexões ao campo profissional e para a própria definição do papel do jornalista; e a adesão de profissionais não só às redações, mas sua inserção em entidades com fins comerciais e não comerciais. Similarmente, Waisbord (2019) identifica que há mudanças no funcionamento e na disposição das organizações jornalísticas. Embora a imprensa de referência ainda se mantenha relevante, existem outros atores nesta sociedade em rede, em fluxos multidirecionais de comunicação e em meio à abundância de informação nos meios digitais, alterando as rotinas e percepções da audiência (BOCZKOWSKI, 2021), naturalmente exausta pelo excesso de informação.

Nesta tese, tomo como base os estudos que situam o jornalismo como instituição social (FRANCISCATO, 2005; PICKARD, 2020), agente na construção social da realidade (TRAQUINA, 2005; TUCHMAN, 1978), sistema perito (MIGUEL, 1999) e forma ou formas de conhecimento, o que foi tratado inicialmente em texto seminal de Park (1940) e discutido posteriormente por autores como Genro Filho (2012 [1987]), Meditsch (1992; 1997), Lisboa e Benetti (2015) e, mais recentemente, atualizado por Nielsen (2017). A busca pela

geração de conhecimento se viabiliza pela prática de informar de modo qualificado (REGINATO, 2019) a partir da verdade verificável que tem os fatos como ponto de partida (KOVACH; ROSENSTIEL, 2014). Naturalmente, o desenvolvimento tecnológico atravessa o jornalismo e, por isso, é inegável que haja impacto em suas práticas. Mesmo assim, a tecnologia não substitui o jornalismo, que é o agente transformador. Concordo com Zelizer (2019, p. 343) quando diz que “[...] é o jornalismo que fornece propósito, forma, perspectiva, significado e significância à tecnologia, e não o contrário.”<sup>37</sup>.

Para nomear o cenário de produção jornalística que tem a Web como principal meio de produção e distribuição, utilizo o termo *jornalismo digital* e me alio a autores como Waisbord (2019, p. 352), para quem o digital se refere “[...] à ascensão e consolidação de formas de ação social em rede, facilitadas por inovações tecnológicas”<sup>38</sup>. O jornalismo e as notícias, neste caso, estão associados às transformações institucionais e estruturais, viabilizadas principalmente por meio da internet e seus atributos de informação e tecnologia. Associando os dois termos, a definição de jornalismo digital significa “[...] a produção em rede, distribuição e consumo de notícias e informação sobre temas de interesse público”<sup>39</sup> (WAISBORD, 2019, p. 352). Em português, nomenclaturas como ciberjornalismo e webjornalismo também são utilizadas, e possuem suas próprias especificidades (MIELNICZUK, 2003). Como essa não é a discussão principal da tese, entendo que não há prejuízo em utilizar o termo jornalismo digital como marco conceitual para situar o Jornalismo Guiado por Dados.

Waisbord (2019) entende que o jornalismo, ainda nos moldes industriais, já era caracterizado por um tipo de rede – mas a conexão, nesse caso, formada pelo jornalismo e outros atores, como governos, corporações, anunciantes, relações públicas e outros; e tudo isso influenciava características como conteúdo, formas de financiamento, estilo, formato, posições editoriais. No jornalismo digital, essas redes se complexificam e estão abertas, promovendo a relação de múltiplos interesses e diferentes formas de fazer reportagem.

---

<sup>37</sup> No original “it is journalism that gives technology purpose, shape, perspective, meaning and significance, not the other way around”.

<sup>38</sup> No original “[...] the rise and consolidation of networked social action, ‘digital’ foregrounds institutional and structural transformations in journalism and news.”

<sup>39</sup> No original “Digital Journalism is the networked production, distribution, and consumption of news and information about public affairs.”



Estudiosos como Mielniczuk (2003), Machado (2006) e Barbosa (2007) se debruçaram sobre o desenvolvimento do jornalismo feito para e por meio da internet e das bases de dados no Brasil, demonstrando sua centralidade. Em sua tese de doutorado, Mielniczuk (2003) identificou três gerações de desenvolvimento do jornalismo na Web. Resumidamente, a pesquisadora descreve um cenário de prática de jornalismo que começa com a fase de transposição de conteúdo offline para a Web, ainda muito parecido com um jornal impresso e com atualizações a cada 24 horas; a segunda fase, por sua vez, demonstra o início da utilização de hiperlinks e alguns recursos como forma de explorar as potencialidades oferecidas pela Web. A terceira geração se materializa com a chegada de negócios e editorias destinadas exclusivamente para a internet; e há, também, uso de recursos multimídia, como sons e animações na produção jornalística.

Somando-se às categorias de Mielniczuk, Barbosa (2007) identifica que o terceiro e o quarto estágios do jornalismo têm as bases de dados como parte fundamental, e propõe em sua tese o paradigma Jornalismo Digital em Base de Dados (JDBD), que reúne características verificadas em sites jornalísticos de perfil dinâmico. Neste modelo, as bases de dados<sup>40</sup> são definidoras da estrutura, da organização e da apresentação dos conteúdos jornalísticos, ou seja, perpassam todo o processo.

A quinta geração (BARBOSA, 2013, p. 42), por sua vez, é definida pelo desenvolvimento do jornalismo nas redes digitais e mídias móveis (como os *smartphones*), capazes de reconfigurar “a produção, a publicação, a distribuição, a circulação, a recirculação, o consumo e a recepção de conteúdos jornalísticos em multiplataformas”. O paradigma JDBD ainda apresenta extensões no jornalismo contemporâneo, como o próprio Jornalismo Guiado por Dados (BARBOSA; TORRES, 2013).

Em termos de produção jornalística e potencialidades, é possível refletir sobre as características inerentes ao jornalismo digital. Canavilhas (2014) identifica sete principais, a saber: hipertextualidade, multimedialidade, instantaneidade, interatividade, memória, personalização e ubiquidade. Duas dessas características foram atualizadas recentemente por Träsel (2020, p. 33-34) em atualização de texto do pesquisador Marcos Palacios. Träsel propõe a substituição do termo “personalização” por “adaptabilidade”,

---

<sup>40</sup> Nesta tese, empregamos o termo *base de dados* conforme Barbosa (2007), para quem a Web funciona desta forma, abrangendo sites que oferecem diferentes elementos (texto, fotos, infográficos, áudio e vídeo, por exemplo). As bases de dados, assim, servem de estrutura e suporte para o conteúdo hospedado nos sites da Web. Os *bancos de dados*, por sua vez, são normalmente temáticos, organizam informações e são as principais fontes do Jornalismo Guiado por Dados. Em alguns contextos e dependendo da área, os termos são usados como sinônimos.

considerando recursos incorporados na publicação de notícias digitais, e “memória” por “perenidade” para designar a disponibilidade cumulativa de material jornalístico, não havendo limites espaço-temporais.

O foco desta tese não é discutir as propriedades ou características do jornalismo digital ou mesmo do jornalismo em rede ou pós-industrial. Cabe pontuar, porém, que as possibilidades trazidas pelo hipertexto apresentam caráter transformador também no Jornalismo Guiado por Dados. Como veremos ao longo da tese, o hipertexto<sup>41</sup> é uma característica fundamental para viabilizar a transparência metodológica. Os links, parte fundamental do hipertexto, desempenham, segundo Baccin (2021, p. 124), “funções vitais para o jornalismo” – entre elas a contextualização e a produção de conteúdo em camadas de aprofundamento. O hipertexto também permite a hipernarrativa no JGD (BARBOSA; TORRES, 2013), visto que possibilita múltiplas direções viabilizadas pelas bases de dados, permitindo camadas de aprofundamento e novos direcionamentos na informação. Constrói-se, assim, um tipo de jornalismo que dialoga com a contextualização e apresenta potencial para gerar uma forma específica de conhecimento.

---

<sup>41</sup> Segundo Baccin (2021, p. 123): “O hipertexto é composto por blocos de informação que são denominados de *lexias* e pelas ligações que se estabelecem entre esses blocos de informação, que são os *links*. Não existe hipertexto sem *link*.”

### 3.1 O potencial para geração de conhecimento específico

O jornalismo<sup>42</sup> é uma instituição estruturada sob o princípio da verdade e que apresenta finalidades específicas a serem cumpridas em sociedades democráticas, a exemplo da fiscalização do poder público e da circulação de informações precisas, qualificadas e relevantes (REGINATO, 2019). A notícia, forma mais basilar do jornalismo informativo, tem potencial para gerar conhecimento à medida que apresenta fatos verificados e pressupõe a contextualização de fenômenos.

Na tentativa de avançar em um caminho teórico trilhado por diversos pesquisadores, busco situar o Jornalismo Guiado por Dados como um **conjunto de práticas e saberes cujos procedimentos permitem aproximá-lo do conhecimento científico pela via do método**. Para tanto, parto de reflexões anteriores (GEHRKE, 2016a; TRÄSEL; GEHRKE, 2021) e de estudos realizados por pesquisadores contemporâneos (NIELSEN, 2017; WIHBEY, 2019) que atualizam o pensamento de autores seminais como Robert Park. Assim, compreendo que o JGD apresenta três pilares que permitem aproximá-lo de um saber sistemático: **verdade, objetividade e transparência**, características exploradas nesta tese.

Ao situar o Jornalismo Guiado por Dados, não pretendo minimizar outras práticas jornalísticas ou formas de conhecimento; apenas busco contribuir apontando suas especificidades. O JGD é usualmente percebido como uma área essencialmente prática. No entanto, ao ter suas origens examinadas, revela-se uma forma assentada na objetividade metodológica, aproximando-se do que preconiza a ciência, como veremos a seguir. Trata-se de um lugar em que há conteúdo explicativo e duradouro, diferente de outras formas digitais existentes na atualidade (NIELSEN, 2017).

Por que, afinal, é importante situar a geração do conhecimento próxima ao JGD? Para Wihbey (2019), somente a partir desta condição o jornalismo pode adotar melhores práticas. Ao estabelecer processos baseados no conhecimento, os jornalistas evitam, por exemplo, matéria-prima de baixa qualidade na produção de conteúdo (como o uso de

---

<sup>42</sup> O jornalismo do qual falo se desenvolve em sociedades democráticas e tem como ponto de partida o século XIX, especialmente na Europa e nos Estados Unidos, com o objetivo de oferecer informação e não propaganda (TRAQUINA, 2005). Assim, este jornalismo se desenvolve à luz do paradigma da verdade, associado a valores como objetividade, independência, verdade e noção de serviço prestado ao público. Ligado aos valores democráticos, o jornalismo apresenta dois papéis principais a cumprir, segundo Traquina: vigiar o poder público e proteger os cidadãos de eventuais abusos de poder de seus governantes; e fornecer aos cidadãos as informações necessárias para a vida em sociedade.

simples declarações e cópias de releases), visto que consideram seu conteúdo uma composição qualificada e que pode orientar o público na tomada de decisões. Sponholz (2009), por sua vez, identifica o jornalismo como forma de conhecimento porque por meio dele é possível conhecer a realidade e processar o que é vivenciado em sociedade.

Em linhas gerais, o conhecimento pode ser compreendido como o resultado de informações sistemáticas a serem validadas cientificamente, independente do julgamento de indivíduos ou grupos sociais (WIHBEY, 2019). Sabe-se, ainda, que a maior parte dos avanços científicos ocorre a partir de uma lenta acumulação de conhecimento (O'CONNOR; WEATHERALL, 2019). No contexto do jornalismo, diz respeito à elaboração das notícias, que refletem padrões representativos do mundo. Sobre isso, é possível pensar em pelo menos três processos na cadeia de produção noticiosa: o **conhecimento do jornalista** (DONSBACH, 2014; GUERRA, 2008; KOVACH; ROSENSTIEL, 2014; PATTERSON, 2013); o **conhecimento social compartilhado** (BERGER; LUCKMANN, 2014 [1966]; TRAQUINA, 2005; TUCHMAN, 1978), que dialoga com a produção de conteúdo e os valores sociais em comum; e o **conhecimento a ser produzido** (LISBOA; BENETTI, 2015; GENRO FILHO, 2012 [1987]; MEDITSCH, 1992; 1997; PARK, 1940) pelo jornalismo.

Para autores como Donsbach (2014), Guerra (2008), Kovach e Rosenstiel (2014) e Patterson (2013), o **conhecimento do jornalista** serve de ponto de partida para o modo como os acontecimentos são apreendidos e comunicados à audiência. Donsbach (2014, p. 666) descreve o jornalista como indivíduo pertencente a uma profissão de conhecimento, visto que desempenha suas funções (KOVACH; ROSENSTIEL, 2014; REGINATO, 2019) em uma sociedade baseada nos mais diversos tipos de saberes, indicando como as coisas funcionam e compartilhando parâmetros sociais esperados do comportamento em sociedade. Ao coletar e traduzir informações à comunidade, o profissional está apto a gerar conhecimento. Para que isso de fato ocorra, o jornalista deve, segundo o autor: conhecer questões da história e da atualidade, assim como exercitar um pensamento analítico; ser especialista em tópicos de cobertura; apresentar conhecimento científico sobre os processos de comunicação; ter habilidades jornalísticas; e exercer a profissão de acordo com as normas éticas.

Para Donsbach, o jornalista, com o intuito de identificar a singularidade dos eventos e os elementos que se sobressaem em cada fato noticiável, deve apresentar características intelectuais que permitam enxergar o fenômeno como um todo. No percurso de apuração, os profissionais têm de aplicar o conhecimento analítico para testar a validade

das informações junto às fontes, por exemplo. O papel do jornalista, para Kovach e Rosenstiel (2014), está na geração de sentido – isto é, colocar os eventos em contexto e, assim, transformar a informação em conhecimento.

Pode-se dizer, então, que o conhecimento do jornalista é fundamental porque é ele quem realiza o processo de interpretação dos acontecimentos. Para fazer isso, os jornalistas são guiados por “técnicas cognitivas”, ou seja, seu aporte de conhecimento, para apurar e analisar um fato dentro de um “percurso interpretativo” (GUERRA, 2008, p. 189). Essa interpretação pressupõe uma retomada dos conceitos previamente acionados, e que são socialmente compartilhados, como veremos a seguir. Guerra (2008) estrutura o percurso interpretativo em quatro etapas: a pergunta, o recorte, a seleção e a sistematização.

Contudo, para alcançar a geração de conhecimento, a narrativa jornalística não pode ser apenas uma interpretação. Embora todo e qualquer fato seja mediado por uma, deve ter caráter explicativo (LISBOA; BENETTI, 2015), com uma narrativa clara e apurada (GANDOUR, 2020). Além do pensamento crítico, o Jornalismo Guiado por Dados aciona outros tipos de habilidades – essas competências por vezes dependem do estabelecimento de uma equipe multidisciplinar.

Wihbey (2019) lista seis competências profissionais para que seja efetivada a relação jornalismo/conhecimento, todas associadas fortemente ao JGD: 1) Compreender estatística básica e possuir habilidades quantitativas, incluindo como uma pesquisa é desenvolvida, as questões de amostragem etc.; 2) Entender como novas hipóteses são testadas; 3) Ter a habilidade de manipular dados em formas tabulares; 4) Mapear discursos e conseguir estabelecer o estado do conhecimento sobre algo; 5) Apresentar fluência com registros públicos, compreendendo procedimentos relativos à legislação; 6) Conhecer técnicas da Web para adquirir e apresentar informações.

Sob a vertente do jornalismo baseado em conhecimento, Patterson (2013) entende que a explicação dos fatos deva ser protagonista nos textos jornalísticos. O autor estabelece duas formas principais para que essa premissa seja viabilizada: o conhecimento de conteúdo (*content knowledge*), ligado a um tópico específico e que exige grau de especialização entre os profissionais que cobrem uma área; e o conhecimento do processo (*process knowledge*), relacionado à aplicação e à comunicação dos métodos de apuração – em outras palavras, a **transparência metodológica**.

Para que o conhecimento do jornalista possa ser útil à sociedade, é preciso que seus membros compartilhem valores em comum. Isso nos leva ao segundo tópico do

jornalismo como forma de conhecimento: o **conhecimento social compartilhado**. A realidade retratada nas notícias representa os fatos já apreendidos e objetivados por meio da linguagem (BERGER; LUCKMANN, 2014 [1966]) e com base nos processos de produção, seleção e transformação dos acontecimentos (TRAQUINA, 2005). O jornalismo trata de um mundo de significados compartilhados socialmente, ainda que cada indivíduo viva e interaja no seu próprio entorno.

Para produzir sentido socialmente, o jornalismo se utiliza de tipificações, isto é, a atribuição de características a indivíduos ou grupos específicos. Tais tipificações são acionadas nas notícias e contribuem para a produção de sentido, visto que os indivíduos partilham de um mesmo acervo social de conhecimento. “A estrutura social é a soma dessas tipificações e dos padrões recorrentes de interação estabelecidos por meio delas. Assim sendo, a estrutura social é um elemento essencial da realidade da vida cotidiana” (BERGER; LUCKMANN, 2014, p. 51). Nesta mesma linha, Tuchman (1978, p. 2), por sua vez, admite que o jornalismo não é o único meio capaz de promover o entendimento do mundo no dia a dia, mas se mostra fundamental para auxiliar na contextualização dos fenômenos. “Na busca por disseminar as informações que as pessoas querem, precisam e deveriam saber, os veículos de comunicação circulam e moldam o conhecimento.”<sup>43</sup>

Ao mesmo tempo em que o repertório compartilhado socialmente é essencial para a operacionalização das notícias e para a compreensão de sentidos, abre precedente para a reprodução de estereótipos e de lugar-comum. É preciso ponderar, assim, que as notícias devem ser observadas como uma construção social que reflete questões da realidade de um acontecimento, e incluem os constrangimentos organizacionais de onde são produzidas, as narrativas dos jornalistas, suas rotinas de trabalho, os valores-notícia e as fontes de informação que são consultadas na produção noticiosa (TRAQUINA, 2005). Somam-se a isso as diversas formas em que a comunicação mediada (COULDRY; HEPP, 2017) se apresenta atualmente, ou seja, dentre plataformas e dispositivos.

Sob o eixo do conhecimento social compartilhado, pode-se pensar no jornalismo também como instituição social ou bem público. Com base em teorias econômicas, Pickard (2020, p. 64) provoca que o jornalismo seja compreendido não como um produto de compra e venda, mas como um bem público que a sociedade necessita. “Então, o valor do

---

<sup>43</sup> No original “By seeking to disseminate information that people want, need, and should know, news organizations both circulate and shape knowledge.”

jornalismo para a sociedade transcende a receita que gera.”<sup>44</sup> Neste caso, o autor defende que não gera valor apenas para quem consome seus produtos diretamente, mas para toda a população, que se mantém informada.

Muitos bens públicos – ar limpo, espaço aberto, luz artificial, e conhecimento, para nomear alguns – produzem significativas externalidades positivas [...]. A sociedade demanda esses bens, mas os indivíduos geralmente os subestimam porque são incapazes ou não têm vontade de pagar por seus custos, fazendo com que o mercado os produza menos (PICKARD, 2020, p. 64)<sup>45</sup>.

Ao tratar o jornalismo e a produção noticiosa como um bem público, o autor alavanca a importância das notícias e as coloca em outro patamar, reconhecendo que desempenham um papel essencial. A principal questão, sob este ponto de vista, parece ser a monetização e a inclusão do jornalismo em espaços limitados ou pouco convencionais. Em uma sociedade democrática, o autor considera que o acesso universal ao conteúdo jornalístico deveria ser prioridade, mas admite que o mercado, em geral, falha ao dar suporte à infraestrutura adequada para o consumo de notícias. Deveria se pensar, portanto, sobre um sistema de mídia público, que assegure a inserção do jornalismo em comunidades locais e regionais.

A execução da ideia de Pickard dependeria de um trabalho conjunto entre as organizações, visto que o jornalismo é categorizado também como uma instituição social<sup>46</sup>. Tuchman (1978) apresenta três razões para tal classificação: 1) as notícias configuram um método institucional de disponibilizar informações para os consumidores; 2) o jornalismo é aliado das instituições legitimadas – em geral, políticos e outras figuras conseguem acesso à imprensa sendo fonte; 3) o jornalismo é produzido por profissionais que trabalham em organizações e, por isso, têm práticas institucionais.

Para Franciscato (2005), a instituição jornalística tem o tempo como principal marcador em relação a outras práticas. O tempo é uma das dimensões do mundo social

<sup>44</sup> No original “Thus, journalism’s value to society transcends the revenue that it generates.”

<sup>45</sup> No original “Many public goods – clean air, open space, artificial light, and knowledge, to name a few – all produce tremendous positive externalities. [...] Society requires these goods, but individuals typically undervalue them because they are unable or unwilling to pay for their full costs, which leads to the market under-producing them.”

<sup>46</sup> Segundo Berger e Luckmann (2014, p. 77), “A institucionalização ocorre sempre que há uma tipificação recíproca de ações habituais por tipos de atores. Dito de maneira diferente, qualquer uma dessas tipificações é uma instituição”. O conceito de instituição não será aprofundado neste momento, mas torna-se necessário para compreender o momento pelo qual passa o jornalismo.

(as três são tempo, espaço e dados) mencionadas por Couldry e Hepp (2017). Em outras palavras, o jornalismo busca atender a uma demanda do público por novidade – o jornalismo como forma de conhecimento está preocupado essencialmente com o presente. “Por princípio, o jornalismo leva o leitor a ver o evento jornalístico não como um fragmento de algo que pertence ao passado, mas como um fragmento do presente, mesmo que tenha ocorrido há poucos momentos.” (FRANCISCATO, 2005, p. 134).

Até agora, entendemos que considerar o jornalismo uma forma de conhecimento implica admitir que o jornalista detém saberes individuais que serão utilizados na prática profissional, assim como reconhecemos que existe um conhecimento social compartilhado que serve de ponto de partida para as notícias. Implica aceitar, também, que o jornalismo é uma instituição social balizada pelo tempo e identificada como credível e com autoridade suficiente para, ainda que dependente de uma série de processos, orientar a vida em sociedade.

Nesta seara, cabe mencionar a percepção de Miguel (1999), que classifica o jornalismo como sistema perito. Isto é, um sistema que apresenta competência especializada e que recebe a confiança do público. Tomemos como exemplos: viaja-se de avião porque existe confiança em um sistema aéreo (por vezes bastante abstrato) seguro, em que os acidentes são raros; vai-se ao médico porque o pressuposto é de que, tendo estudado para isso, o profissional tem conhecimento para diagnosticar o problema. Para que o jornalismo seja percebido como um sistema perito, a audiência precisa confiar nas escolhas dos jornalistas quanto à veracidade das informações, quanto à seleção dos elementos relativos ao fato e quanto à seleção e hierarquização de notícias relevantes; além disso, reverbera o conhecimento produzido por outras instituições sociais.

Partindo do pressuposto de que o jornalismo reúne profissionais qualificados e capazes de fazer jus à produção de conhecimento, que executam sua prática em um cenário compartilhado de saberes e valores, é preciso entender como funciona, por fim, o **conhecimento a ser produzido**. Lisboa e Benetti (2015, p. 16) nos ajudam a compreender, em suma, como isso acontece: “O jornalismo se torna conhecimento quando ele se constrói como um testemunho verossímil da realidade, baseando sua produção noticiosa em métodos e processos que tentam reduzir o erro e os relatos falsos”.

Um dos trabalhos pioneiros nos estudos de jornalismo justamente para problematizar o tipo de conhecimento gerado por esta prática social parte do sociólogo Robert Park (1940). Em estudo amplamente citado na área, estabelece um *continuum* entre o



*conhecimento sobre*, formal e sistemático (*knowledge about*), e a *familiaridade com* (*acquaintance with*), ligado ao senso comum. Para o autor, as notícias geram conversação na sociedade, têm caráter de documento público e são capazes de gerar mudanças imediatas.

A *familiaridade com*, dita por Park (1940, p. 674), caracteriza um saber intuitivo, produto da experiência pessoal e da experimentação. Trata-se de algo efêmero, com o qual temos contato, mas não nos aprofundamos. O *conhecimento sobre*, por sua vez, é formal, racional e sistemático – pressupõe fatos confirmados e verificados. Ao contrário da *familiaridade com*, que está na ordem da percepção, o *conhecimento sobre* pode ser testado, experimentado e comunicado, tal como ocorre na ciência. “O conhecimento sobre não é meramente experiência acumulada, mas o resultado de uma investigação sistemática da natureza.”<sup>47</sup>

A categorização de Park ainda captura os aspectos essenciais do jornalismo e da circulação de informações na atualidade, mas não abrange todos os tipos de conteúdo hoje publicados. Pensando em um cenário de notícias digitais, Nielsen (2017) atualiza as teorias de Park e cria três tipos ideais para abranger tipos de conteúdo digital. O que o autor denomina de formas de *notícias-como-impressões* correspondem aos fragmentos de informação que circulam em alertas de notícia, sites de rede social e aplicativos de mensagens. Corresponde ao que Park (1940) classificaria como *familiaridade com*. O segundo tipo ideal é o que o Nielsen chama de *notícias-como-itens*, que são artigos de jornal, site, ou histórias apresentadas no rádio e na televisão com a ideia de atualidade e referente às últimas 24 horas. São as formas tradicionais de notícias do final do século XX, com características clássicas do jornalismo. Por fim, o autor apresenta as *notícias-sobre-relações* – o equivalente ao conhecimento sobre – com conteúdo explicativo e contextual, a exemplo das reportagens *longform* e do Jornalismo Guiado por Dados.

As três categorias criadas por Nielsen compreendem percepções de um conteúdo efêmero – ideia de fluxo, de informações rápidas que circulam pelo feed – e de estoque, ou seja, um tipo de conteúdo durável e que pode ser consultado independente da data. As formas rápidas de notícias, que circulam em fluxo, integram o ecossistema de notícias do qual fala Wihbey (2019). Em um cenário de rede, descentralizado, o surgimento de fatos sociais<sup>48</sup> é recorrente e por vezes substitui os fatos verdadeiros, que dão origem e

---

<sup>47</sup> No original “Knowledge about is not merely accumulated experience but the result of systematic investigation of nature.”

<sup>48</sup> Fatos sociais, segundo Wihbey (2019), levam ruídos e desinformação para as mesmas plataformas por onde as notícias circulam, podendo gerar desinformação. É fato, por exemplo, que o ex-presidente Barack

sustentam as notícias. Na época em que Park publicou o texto inaugural sobre a notícia como forma de conhecimento, na primeira metade do século XX, o cenário dos veículos de imprensa era muito diferente: a maior parte do conteúdo jornalístico era distribuída por meio de jornais e rádios, e o jornalismo era essencialmente regional. Atualmente, se sobressaem os jornais de referência, as emissoras de televisão e a distribuição online de notícias, em um mercado de abrangência nacional e internacional (NIELSEN, 2017).

Há uma longa trajetória de arguição sobre o conhecimento gerado pelo jornalismo. Enquanto alguns teóricos ignoram a existência do saber produzido, outros pesquisadores o inserem em um espaço distinto, a exemplo do próprio Robert Park (1940), Genro Filho (2012 [1987]), Meditsch (1992; 1997) e Franciscato (2006). Para autores como Meditsch (1992; 1997) e Franciscato (2006), o jornalismo apresenta certa fragilidade metodológica – e, por isso, certo distanciamento da ciência – porque, ao trabalhar com pautas e não necessariamente com hipóteses, desenvolvem seus temas a partir de observações não controladas. Como veremos adiante, o Jornalismo Guiado por Dados busca contornar este aspecto.

O JGD, penso, tem potencial para produzir conhecimento a partir de conteúdos duráveis e explicativos, ou de estoque, como afirma Nielsen (2017). Em síntese, defendi, até o momento, que o jornalismo como um todo pode ser caracterizado como forma capaz de gerar conhecimento social específico, e que isso se estrutura a partir de três fatores principais: o **conhecimento do jornalista**, o **conhecimento social compartilhado** e o **conhecimento a ser produzido**. Para tanto, suas práticas devem estar ancoradas na **verdade**, na **objetividade** e na **transparência**. A exemplo do que ocorre na ciência, permite a reprodutibilidade do método e a revisão entre pares. Neste sentido, o JGD tem forte influência de um de seus precursores, o Jornalismo de Precisão, criado pelo jornalista Philip Meyer.

---

Obama nasceu nos Estados Unidos, e é um fato social algumas pessoas terem acreditado que ele nasceu em outro país, conforme propagado por outros políticos.

### 3.2 A base científica do Jornalismo de Precisão

O termo Jornalismo Guiado por Dados passou a ser adotado na primeira década dos anos 2000 e tem como marco o texto escrito pelo programador estadunidense Adrian Holovaty (2006) sobre a necessidade de estruturar e reaproveitar as informações coletadas no dia a dia da reportagem. Soma-se a esse texto a ampliação de políticas públicas de transparência ao redor do mundo e a consequente ampliação da disponibilidade de dados abertos, que preparou o terreno para a expansão do JGD.

O desenvolvimento do JGD e de práticas semelhantes, contudo, não ocorreu de maneira uniforme em nível global, visto que o panorama social, cultural e tecnológico varia de uma região para outra. Suas origens remetem principalmente aos Estados Unidos, onde foi registrado o surgimento de formas jornalísticas que viriam a dar origem a ele: o Jornalismo de Precisão, criado pelo repórter Philip Meyer na segunda metade dos anos de 1960, com livro de mesmo nome publicado pela primeira vez em 1973; e a Reportagem Assistida por Computador (RAC), do inglês *Computer-Assisted Reporting* (CAR), considerada por pesquisadores uma versão atualizada de seu precursor, porém com viés acentuado na tecnologia. Em um cenário recente, um dos principais marcadores do Jornalismo Guiado por Dados é o *The Guardian Data Blog*<sup>49</sup>, fundado em 2009 e liderado por Simon Rogers.

Nesta tese, o objetivo não é recapitular em profundidade os primeiros usos de dados na história do jornalismo moderno, que remetem ao século XIX e estão ligados à publicação de tabelas e infográficos (ROGERS, 2013). Por ora, considerando o tema deste trabalho, resgato brevemente a proximidade entre o Jornalismo Guiado por Dados e como essa prática possibilitou que o jornalismo tivesse mais conhecimento sobre as epidemias que acometeram sociedades ao longo da história.

Um dos pontos de partida para o uso de dados para mudanças em políticas públicas de saneamento básico, e inspiração para o JGD, é o mapa<sup>50</sup> desenhado pelo médico John Snow como solução para a epidemia de cólera que acometeu a metrópole de Londres em 1854. À época, a teoria predominante era de que a doença se espalhava pelo ar. Por meio de observação sistemática e análise microscópica, Snow foi quem desenvolveu a tese de que a transmissão ocorria por meio do consumo de água contaminada. Ao situar os óbitos em um mapa, identificou um padrão: as mortes por cólera ocorriam nas proximidades da

---

<sup>49</sup> Disponível em: <https://www.theguardian.com/news/datablog/2009/mar/10/blogpost1>

<sup>50</sup> Disponível em: <https://www.theguardian.com/news/datablog/2013/mar/15/john-snow-cholera-map>

bomba de água da *Broad Street*. Munido de informações estatísticas da cidade, então sistematizadas no mapa, Snow identificou que, entre 83 mortes, 73 eram de pessoas que moravam nas proximidades da bomba em questão (JOHNSON, 2008). Londres já havia vivenciado outras crises de contaminação pela bactéria causadora de cólera. Descobriu-se que a epidemia de 1854 foi originada pelo contato da água com as fezes de um bebê doente: a água (contaminada, descobriu-se mais tarde) usada para limpar as fraldas foi despejada no esgoto da residência, que entrou em contato com o sistema de abastecimento do poço e, por tabela, contaminou os moradores que consumiram a água do local. Naquele período, a capital da Inglaterra experimentava sua expansão como metrópole, tendo cerca de 2,5 milhões de habitantes, muitos dos quais viviam aglomerados em residências, contribuindo para o rápido espalhamento de doenças em uma época de precariedade no sistema de abastecimento de água e esgoto.

Retomando o olhar para os dias de hoje, é difícil não traçar um paralelo com a pandemia de Covid-19: os dados mais uma vez são capazes de estabelecer padrões em um cenário bastante abstrato de contaminação pelo vírus Sars-Cov-2. Com a mobilidade acentuada entre os países, uma epidemia ganha proporções de pandemia e se espalha rapidamente. Para observar o cenário de contaminação e preparar alternativas, é preciso coletar, analisar os dados e colocá-los em contexto. Em 2021, o projeto *No Epicentro*<sup>51</sup> (desenvolvido pela *Agência Lupa* e *Google News Lab*), que busca situar em um mapa interativo a quantidade de mortes por Covid-19, foi um dos vencedores na competição internacional intitulada *Sigma Awards*<sup>52</sup>.

O Jornalismo de Precisão, estabelecido por Meyer (2002) como a aplicação dos métodos científicos de pesquisas sociais e comportamentais para a prática do jornalismo, surgiu como uma resposta científica e é considerado um dos precursores do JGD. Ele acreditava que só assim seria possível diminuir o risco de manipulação sofrido pelos repórteres no diálogo com as fontes oficiais. A proposta de Meyer não foi imediatamente bem recebida e sequer compreendida em sua totalidade, e seu livro demorou alguns anos para ser publicado (MEYER, 2012). Em entrevista a Gehrke e Mielniczuk (2017), Meyer disse que seu principal aprendizado no processo foi compreender que a pesquisa científica era algo praticável no campo do jornalismo, podendo ir além de observar e reportar um fato. Tendo passado um ano sabático em Harvard, onde aprendeu sobre métodos das ciências

---

<sup>51</sup> Disponível em: <https://piaui.folha.uol.com.br/lupa/epicentro/>

<sup>52</sup> Disponível em: <https://sigmaawards.org/>

sociais e aproximou-se da estatística, Meyer abriu os olhos para o universo e o poder da análise de dados. Assim, acredita que sua principal contribuição tenha sido demonstrar o valor do método científico para a busca de fatos e a análise de dados no jornalismo.

O término do período de estudos de Meyer em Harvard, onde teve contato com computadores, coincidiu com uma série de distúrbios sociais ocorridos em Detroit, nos Estados Unidos, no ano de 1967. A primeira ação que levou ao colapso social na cidade foi a invasão da polícia a um bar clandestino em 23 de julho. O enfrentamento que lá ocorreu se espalhou e, cinco dias depois, havia 43 mortos, 7 mil detidos e quarteirões destruídos (TRÄSEL, 2014). A suspeita inicial era de que o colapso havia sido causado por integrantes de camadas socialmente excluídas. Designado para trabalhar junto à equipe do jornal *Detroit Free Press* e cobrir os acontecimentos, Meyer propôs uma pesquisa na região afetada pelos casos com o intuito de descobrir as motivações para o confronto com a polícia e as ações subsequentes, como a destruição da cidade. A partir do teste de hipóteses, como se costuma fazer na ciência, a equipe descobriu que o motim fora causado principalmente pela brutalidade policial, as más condições de moradia e a falta de empregos (TRÄSEL, 2014). Ao contrário do esperado, os imigrantes não eram os principais responsáveis pelas ações em Detroit; a maior parte era composta por moradores da própria cidade. A pesquisa de Meyer e seus colegas também permitiu uma generalização: em oposição ao que se pensava inicialmente, os negros que participaram dos manifestos eram minoria, e suas ações não representavam o sentimento da comunidade como um todo. A equipe confirmou, então, a teoria da privação relativa: o colapso social teria sido causado pela frustração. Quanto mais perto se chegava a um objetivo, maior era a frustração por não o alcançar. Era como se uma parcela da população fosse abandonada e não conseguisse avançar socialmente – e esse era o motivo-chave para os distúrbios.

Devido a esses movimentos, Meyer pôde colocar em prática os aprendizados obtidos em Harvard e, finalmente, testar a própria teoria ao aplicar métodos de ciência social junto ao trabalho de reportagem. Em 1968, a publicação de um relatório da pesquisa desenvolvida pela equipe do *Detroit Free Press* venceu na categoria de reportagem local do *Prêmio Pulitzer*, o mais importante prêmio de jornalismo no mundo. Meyer ficou conhecido à época como um “repórter de computador”, uma vez que a aplicação de métodos da ciência social parecia ser uma premissa complicada de compreender.

Em termos práticos, o repórter que adotava os métodos do Jornalismo de Precisão não esperaria, por exemplo, pelos resultados de uma pesquisa de intenção de votos – ele

mesmo sairia às ruas e executaria a pesquisa. Apesar de o Jornalismo de Precisão ser frequentemente associado à combinação das pesquisas de opinião com a prática da reportagem (MEYER, 1999), seu criador esclarece que esse modelo profissional previa diversas formas aplicáveis de método científico, incluindo outros tipos de experimentos de campo, desenvolvido não só por ele enquanto repórter, mas também por seus alunos de graduação na Universidade da Carolina do Norte, em Chapel Hill. Meyer, enquanto professor universitário, relata ter enviado estudantes para medir a cortesia dos motoristas, o preconceito racial de pedestres e a venda de álcool a menores de idade.

A delimitação do escopo ao qual cada termo ligado às origens do Jornalismo Guiado por Dados se enquadra não é exatamente consensual. Alguns pesquisadores consideram a RAC uma versão repaginada do Jornalismo de Precisão à época em que os computadores passaram a ser utilizados com mais frequência nas redações, o que nos Estados Unidos aconteceu em meados dos anos de 1980 – antes disso, no início dos anos de 1950, o computador havia sido usado pontualmente em cálculos e modelos preditivos. Enquanto LaFleur (1999, p. 27) define a RAC como uma iniciativa que visa “[...] analisar dados para elaborar textos que os repórteres não poderiam fazer de outra forma”<sup>53</sup>, Meyer afirma que, no Jornalismo de Precisão, o uso do computador não é essencial. Sobre isso, critica o uso da palavra “computador” na formação do termo RAC, defendendo a adoção de um novo nome, com ênfase na ciência, e admitindo que o Jornalismo de Precisão estava datado.

Qualquer que seja o novo termo, vamos desistir da RAC. Computadores são usados de tantas maneiras, até mesmo na redação, que isso já não nos define, se algum dia definiu. Chegou o tempo de declarar a RAC vitoriosa e seguir para um conceito novo, com um conceito mais ambicioso (MEYER, 1999, p. 5)<sup>54</sup>.

O Jornalismo Guiado por Dados parece ter assumido o papel conceitual do qual fala Meyer (1999), visto que o termo se apresenta amplamente difundido na atualidade. Ao redor do mundo, é comum perceber sua associação ao jornalismo computacional<sup>55</sup> – Jornalismo de Precisão, a RAC, o jornalismo de dados e computacional integram o que

---

<sup>53</sup> No original “[...] as analyzing data to do stories that reporters couldn’t do any other way”.

<sup>54</sup> No original “Whatever the new term, let’s give up on CAR. Computers are used in so many different ways, even in the newsroom, that it no longer defines us, if ever did. The time has come to declare CAR victorious and move on to a fresher, more ambitious concept.”

<sup>55</sup> O jornalismo computacional é definido por Thurman (2020) como o uso avançado de computação, algoritmos e automação no processo desde a coleta até a apresentação das notícias.

Coddington denomina de virada quantitativa no jornalismo. Em geral, o jornalismo se utiliza de nomes específicos para marcar determinado tempo ou atender uma demanda específica. Para se ter uma ideia, o projeto *X Journalism*, desenvolvido por Loosen *et al.* (2019), mapeou 130 termos utilizados para designar tipos de jornalismo.

Para Anderson (2018), o Jornalismo de Precisão concretiza a virada interpretativa e explicativa que o jornalismo experimentava desde os anos de 1930 nos Estados Unidos. À época, justamente quando o jornalismo buscava explicar o motivo dos eventos, pensando em um processo de contextualização, era alvo de críticas de cientistas, que o consideravam distante da ciência por pautar-se principalmente por questões de interesse público<sup>56</sup> em vez de formular modelos abstratos da vida social. Pode-se inferir, assim, que as origens do JGD têm como pano de fundo um jornalismo contextual, também conhecido como explicativo ou interpretativo<sup>57</sup>.

A exemplo do Jornalismo de Precisão e da RAC, o Jornalismo Guiado por Dados apresenta dimensões tecnológicas, políticas e culturais. Referente à sua própria época de operação, cada tipo apresenta determinada ênfase sobre como obter dados. O Jornalismo de Precisão e a RAC, desenvolvidos na segunda metade dos anos de 1960, nos Estados Unidos, puderam se beneficiar do *Freedom of Information Act*<sup>58</sup>, de 1967, que tornou

---

<sup>56</sup> Parto da definição de Sartor (2018, p.2) para dizer o que entendo como interesse público aplicado aos dados de transparência liberados pelas instituições públicas e utilizados no Jornalismo Guiado por Dados. Interesse público, assim, é um “[...] fundamento ético-epistêmico do jornalismo que diz respeito à relevância dos fatos noticiáveis, à construção da notícia visando ao esclarecimento dos cidadãos, ao papel de vigilância socialmente atribuído ao jornalismo e aos interesses do público no consumo de informações.”

<sup>57</sup> Há uma extensa bibliografia que aborda os jornalismo interpretativo e explicativo. A Sociedade Americana dos Editores de Jornais decidiu, em 1933, que deveria haver maior espaço e dedicação ao jornalismo interpretativo e explicativo com a apresentação de informações de bastidores para que os leitores pudessem entender a importância dos eventos da época (BARNHURST, 2014). Apesar do aumento nas notícias interpretativas de 1930 a 1950, Fink e Schudson (2014) mostram que o jornalismo estadunidense passou por mudanças especialmente após a segunda metade dos anos de 1960. Considerado por estudiosos o pioneiro, o jornal *The National Observer*, que operou de 1962 a 1977 com edições semanais, dedicava-se à publicação de tendências e análises sobre cultura, política, acontecimentos sociais e assuntos gerais (LANDERS, 2005). Durante 15 anos, o jornal publicou conteúdos interpretativos na capa e frequentemente na contracapa sem levar o rótulo de “análise de notícias”, como costumava ocorrer na época. Apesar de não ter prosperado em termos econômicos, o *The National Observer* teve o reconhecimento de outros jornalistas e, em 1974, venceu em uma categoria do Prêmio Pulitzer – comentário sobre temas públicos. E foi finalista em 1967 pela cobertura sobre direitos civis. Uma das possibilidades é que o termo “jornalismo explicativo” tenha se tornado mais conhecido quando virou parte da premiação. Isso ocorreu no ano de 1985, com a criação da categoria do *Exploratory Journalism* no prêmio Pulitzer. Apesar da mudança no nome para *Explanatory Reporting* anos depois, a ideia continuou a mesma: premiar os trabalhos – em geral séries ou conjunto de matérias com suítes – que de alguma maneira ajudaram o público a compreender a complexidade de um tema.

<sup>58</sup> *The Freedom of Information Act* (FOIA) é o nome atribuído à legislação estadunidense que permite a solicitação de informações de qualquer agência do governo federal por qualquer motivo. Entrou em vigor em 1967, um ano depois de ser aprovado no Congresso dos EUA. Embora o próprio FOIA tenha inspirado

possível maior obtenção de dados (BRADSHAW, 2017). Na atualidade, o JGD conta com um movimento global pela abertura de dados (*open data*). Enquanto o Jornalismo de Precisão e a RAC utilizavam-se principalmente de planilhas, aplicando métodos de pesquisa social em seu trabalho, o JGD inclui *Application Programming Interface* (API) e visualização da informação, seguindo, ainda, os preceitos de uma cultura hacker (BRADSHAW, 2017; TRÄSEL, 2014). Conforme Belisário et al. (2021, p. 43), “APIs são interfaces que facilitam a comunicação de dados entre computadores”, estando presentes em redes sociais digitais e portais de dados governamentais.

O surgimento do Jornalismo Guiado por Dados ocorre, ainda, em um contexto internacional de transparência pública e como resposta à crise econômica, identitária (TRÄSEL, 2014) e epistêmica (STEENSEN, 2019) do jornalismo atual. Para Anderson (2018), pode ser compreendido como uma extensão de conceitos e tecnologias desenvolvidas após a Segunda Guerra Mundial. Esse cenário inclui o aumento na digitalização de bases de dados (entre elas os dados de governo), a velocidade dos computadores, o prestígio do jornalismo investigativo, a importância atribuída à identificação de padrões nas reportagens (em vez de contemplar apenas fatos isolados) e, por fim, a premissa de que o jornalismo deve ir além das declarações emitidas pelas fontes, mas de fato comprometer-se com a objetividade.

Atualmente, o JGD se desenvolve em um contexto complexo, cercado pela falta de confiança nas instituições, incluindo a imprensa (PICKARD, 2020), pela disseminação de desinformação (WARDLE; DERAKHSHAN, 2017) e pela influência das plataformas em todos os pontos da cadeia noticiosa (BELL; OWEN, 2017), condicionando a produção e a disseminação de conteúdo. Dialoga, ainda, com a geração de conhecimento a partir das interações sociais mediadas por computador, adicionando, ao tempo e ao espaço, uma camada de dados nas dimensões do mundo social (COULDRY; HEPP, 2017). O uso de método de reportagem e a transparência na comunicação dos dados parecem atuar como diferenciais do JGD em relação a outras práticas.

---

legislações semelhantes no resto do mundo, o primeiro ato nesse sentido partiu do rei da Suécia dois séculos antes (POZEN; SCHUDSON, 2018).



### 3.3 Características do Jornalismo Guiado por Dados na atualidade

O Jornalismo Guiado por Dados é definido por Träsel (2014, p. 119) como “[...] 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”. Comumente associado ao pragmatismo da reportagem, reitero a definição de que se trata de um **conjunto de práticas e saberes cujas origens estão associadas ao conhecimento científico, pela via da verdade e objetividade no método, capaz de oferecer graus elevados de transparência e colaboração, produzindo um conteúdo duradouro**. Guardadas, é claro, as especificidades dos contextos sociais onde é empregado.

Trata-se, ainda, de um conjunto de práticas situado no contexto do jornalismo digital e em rede, e do jornalismo investigativo (HAHN; STALPH, 2019), porém com um viés colaborativo marcante desde o início do processo, em oposição ao estereótipo do lobo solitário do jornalista investigativo. Primeiro porque um jornalista dificilmente terá todas as habilidades necessárias para desenvolver, do começo, uma reportagem guiada por dados (BRADSHAW, 2013); ao mesmo tempo, existe uma proximidade do *ethos* do jornalista investigativo e de dados na identificação de problemas sistêmicos ou padrões e comumente o fato de lançar luz sobre informações escondidas; segundo, em um cenário de conexão via internet além-fronteiras, é possível expandir investigações locais para outras partes do planeta, viabilizando as investigações em rede que são aderidas por jornalistas e redações.

O Jornalismo Guiado por Dados se transformou em um campo relativamente estável no início do Século XXI, segundo Anderson (2018). Neste cenário, há abundância na disponibilidade de dados abertos em novos domínios digitais. É como se o sonho de Meyer e tantos outros estudiosos finalmente tivesse se tornado realidade: jornalistas usam bancos de dados e computadores para comunicar contexto social a suas audiências, eventualmente usando técnicas de ciências sociais para fazer isso. “O jornalismo agora pode implantar a certeza profissional e mobilizar o público para propósitos democráticos baseando-se nos padrões de evidências das ciências sociais<sup>59</sup>” (ANDERSON, 2018, p. 137). Ao avaliar a situação atual do Jornalismo Guiado por Dados, Meyer relatou, em entrevista

---

<sup>59</sup> No original “Journalism can now deploy professional certainty and mobilize the public for democratic purposes by drawing on the evidentiary standards of social science.”

a Gehrke e Mielniczuk (2017), que o conceito ainda parece estar em desenvolvimento. Para ele, o escopo do JGD pode incluir análise de dados e reportagem, mas também uma espécie de curadoria dos dados – isto é, oferecer uma interface amigável à audiência com pouca familiaridade com dados e planilhas.

A comunidade de jornalistas de dados vem ganhando importância no mundo. Ironicamente, é um desafio obter números sobre quantos são seus praticantes. Mesmo assim, a busca por determinados tipos de cursos nos dá pistas de como a prática vem se espalhando e gerando interesse global. Oferecido em 2019 pelo *Centro Knight para Jornalismo nas Américas*, o curso gratuito sobre Jornalismo de Dados e Visualização com Ferramentas Gratuitas (*Data Journalism and Visualization with Free Tools*) teve inscritos em diversos países, conforme mapa divulgado<sup>60</sup>. A maior parcela dos participantes está no Brasil (2.297 inscritos), seguido dos Estados Unidos (1.756) e México (775). Embora este centro priorize a capacitação de profissionais nas Américas, também houve o registro de participantes em países como Espanha (684 participantes), Índia (284) e Reino Unido (237).

Gradativamente, pesquisadores da América Latina conquistam espaço para descrever situações específicas de desenvolvimento do JGD em seus países, visto que é mais recorrente o acesso às práticas do Norte global. Em uma abordagem hiperlocal do Jornalismo Guiado por Dados – as redações nas favelas brasileiras –, De Lima Santos e Mesquita (2021) reconhecem características próprias desse tipo de prática no cenário do Brasil, incluindo as dificuldades na produção noticiosa e a participação da audiência nesse processo. Estima-se que, no país, o JGD seja praticado em 52 instituições (DE LIMA SANTOS, 2019) – ainda que não se saiba exatamente a quantidade de praticantes –, especialmente no Sudeste, que apresenta perto de 70% das organizações de imprensa do levantamento.

Ainda que de forma empírica e pouco sistemática, percebi, ao longo de quase seis anos de observação periódica dos produtos lançados por equipes de JGD, que a prática está muito centralizada nas pessoas – ou seja, em profissionais com alto conhecimento técnico –, e não necessariamente na identidade das redações. De maneira geral, são essas questões técnicas e da ordem do conhecimento que representam uma barreira para a maior difusão do JGD. Pesquisa realizada em 2015 com 875 jornalistas e mencionada por Wihbey (2019) mostra que os profissionais consideram importante, mas se sentem

---

<sup>60</sup> Disponível em: <https://public.flourish.studio/visualisation/807617/embed?fbclid=IwAR1AWG>

poucos seguros ao acionar fontes de dados e realizar análise estatística. Questionados sobre o quão bem equipados/preparados estão para realizar análise estatística, apenas 11% dos profissionais responderam “muito bem”, enquanto 46% disseram “um pouco” e 43% “mal”. Em relação ao conhecimento para interpretar estatística de outras fontes, um quarto (25%) dos respondentes afirma ter uma capacidade “muito boa” para tal, enquanto 58% dizem saber “um pouco” e outros 17% “nada”.

Percentuais semelhantes foram encontrados nas respostas sobre o quanto os profissionais sabem interpretar estudos científicos – 32% disseram saber “muito bem”, 54% “um pouco” e 14% admitiram não ter essa competência. O levantamento ainda mostrou que os profissionais consultados consideram mais importante saber interpretar a análise estatística feita pelas fontes (80% consideram “muito importante”), mas não tão fundamental realizar análise estatística por conta própria (neste caso, somente 39% dos entrevistados consideraram essa habilidade “muito importante”). Saber interpretar estudos científicos – o que é fundamental na cobertura de pandemias como a de Covid-19, por exemplo – é considerado “muito importante” por 77% dos participantes.

Apesar da evolução tecnológica registrada com o passar dos anos, as barreiras ao desenvolvimento da RAC ainda parecem se apresentar no presente. Uma delas tem a ver com a falta de transparência do poder público, que demora a divulgar certos dados; a segunda está ligada às limitações do jornalista (LaFLEUR, 1999). À época em que criou o Jornalismo de Precisão, Meyer alertou que o excesso de informações faria com que os profissionais precisassem passar a saber gerenciar, processar e analisar dados. Assim, a competência de escrever bem já não seria suficiente. “[...] o jornalista precisa ser um filtro, assim como um transmissor; um organizador e interpretador, assim como aquele que obtém e entrega os fatos”<sup>61</sup> (MEYER, 2002, p. 1). Trazendo a discussão para a atualidade, Wihbey (2019) considera que dificilmente os jornalistas se tornarão estatísticos, mas pondera que tais saberes são valorizados e deveriam ser considerados pilares no trabalho da imprensa, pois fazem parte do conhecimento esperado para a evolução desta especialidade dentro do jornalismo digital.

Particularmente, entendo que o jornalista de dados não necessariamente deva aprender uma linguagem de programação, ainda que o conhecimento técnico gere algumas facilidades e autonomia, incluindo a clara estruturação do método de trabalho. É

---

<sup>61</sup> No original “[...] the journalist needs to be a filter, as well as a transmitter; an organizer and interpreter, as well as one who gathers and delivers facts”.

fundamental, porém, que tenha pensamento crítico, segurança para acessar as fontes documentais normalmente utilizadas no JGD e conhecimento para identificar limites e possibilidades desse tipo de jornalismo, a exemplo do uso de códigos de programação, em consonância com suas práticas multidisciplinares. Justamente por ser um campo que congrega competências das mais diversas áreas, não se pode esperar que o jornalista de dados consiga executar, sozinho, todas as etapas do fluxo de trabalho. Assim como na ciência, é esperado que haja colaboração e troca entre pares.

O Jornalismo Guiado por Dados segue em transformação na atualidade, e algumas questões são esperadas para que o seu avanço ocorra a partir de sua segunda década de práticas. É o que Bradshaw (2017) chama de “segunda onda” do JGD. Para o autor, esta etapa é caracterizada pela ascensão do jornalismo robô – uso de script para analisar e gerar uma grande quantidade de dados – e do pensamento computacional, que envolve pensar sob a lógica de um computador e compreender seu poder de resolução de problemas. Inspirado nos elementos do jornalismo escritos por Kovach e Rosenstiel, Bradshaw (2017) estabeleceu dez princípios para o que se espera desta prática profissional.

### Quadro 1 – Resumo dos princípios do Jornalismo Guiado por Dados

Elemento	O que compreende
1. Interrogar os dados para orientar a tomada de decisões	Atenção à disponibilidade, à acurácia e ao emprego de dados no jornalismo, de modo que a publicação dessas informações seja útil à tomada de decisões por parte do poder público, a exemplo de pautas como imigração e segurança pública.
2. Defender a independência editorial e tecnológica	Prevenção aos vieses inerentes às ferramentas utilizadas para criar códigos de programação. Tais vieses podem afetar a verificação de fontes e documentos, além de representar um risco à proteção das fontes.
3. Demonstrar objetividade nas fontes, na linguagem e nas ferramentas utilizadas	Frequentemente, o Jornalismo Guiado por Dados elabora ferramentas interativas, que serão utilizadas pelo público final. Assim, é importante que o público conheça como a ferramenta foi elaborada e suas eventuais limitações.

4. Buscar dados que vão além do óbvio	A abundância de informação provoca o uso de fontes oficiais e de bancos de dados públicos de forma recorrente. Para aprimorar a qualidade do jornalismo, contudo, é preciso olhar para o que está escondido.
5. Dar voz às minorias	O desenvolvimento de projetos jornalísticos que apresentem dados sobre minorias para preencher a lacuna do “deserto de dados”, incluindo a digitalização de arquivos inexistentes na Web.
6. Considerar o impacto da personalização	Ao refletir sobre temas relevantes que desejam cobrir, jornalistas de dados precisam levar em conta o impacto da personalização e da interatividade para fornecer diferentes dimensões de uma mesma história, incluindo profundidade sobre o tema.
7. Contar a história a partir de personagens	O Jornalismo Guiado por Dados não é apenas sobre números, mas sobre pessoas. Sem números, o texto corre o risco de se tornar uma anedota. Sem personagens, apenas mera estatística.
8. Abrir o código dos algoritmos	A transparência algorítmica nos scripts é essencial para que o público possa fazer sua própria avaliação do tópico trabalhado. Submeter os códigos ao escrutínio público pode contribuir para a credibilidade.
9. Compartilhar o código de programação	A disponibilização do código de programação passível de reprodutibilidade e replicabilidade, recursos que dialogam com a transparência, são processos-chave para a compreensão da origem e da comunicação dos dados, além de gerar colaboração na comunidade.
10. Empoderar cidadãos	Auxílio ao público para conhecer seus direitos e exercitar a cidadania a partir da criação de ferramentas que dialogam com áreas como política e economia, de significativo impacto social.

Fonte: Bradshaw (2017) com adaptações da autora (2021).

No entendimento do autor, visível no **Quadro 1**, a utilização de dados deve ocorrer como base para os avanços sociais, incluindo a tomada de decisão em políticas públicas e a maior variedade de pautas, a exemplo de tópicos que são importantes para as minorias. A interatividade e o toque humano às produções, o que ocorre com a utilização de

personagens nas notícias, igualmente são apontados como princípios fundamentais para o avanço desta prática.

Destaco, ainda, os pontos *três, oito e nove*, que dialogam com as questões cruciais a esta tese: objetividade no método e transparência. Além da maior disponibilidade de dados e a criação de novas ferramentas para análise, abrir o próprio trabalho para o escrutínio público – permitindo que os processos sejam avaliados, reproduzidos e reutilizados, como preconizava Meyer desde o Jornalismo de Precisão – parecem ser o principal diferencial do JGD. Além disso, são um incentivo à confiança e à credibilidade das organizações jornalísticas.

D'Ignazio e Klein (2020, p. 19), por sua vez, propõem uma abordagem feminista em relação aos dados, evidenciando as lacunas na produção de informações e o quanto isso prejudica minorias, frequentemente desconsideradas na coleta de dados. A ausência dessas informações revela estruturas de poder que se perpetuam socialmente. É comum, por exemplo, que não haja dados públicos disponíveis em relação às áreas periféricas de uma cidade, ou então a grupos raciais ou pessoas não binárias. O feminismo de dados expõe que a geração de produtos é impactada pela coleta e disponibilidade de informações. As autoras elencam sete pontos de atenção:

### Quadro 2 – Os sete princípios do feminismo de dados

Elemento	O que compreende
1. Examinar o poder	Analisar como o poder se comporta no mundo.
2. Desafiar o poder	Desafiar as estruturas de poder e trabalhar na direção da justiça social.
3. Elevar a emoção e a corporalidade	Valorizar múltiplas formas de conhecimento, incluindo aquele que vem das pessoas e de sua vivência no mundo.

4. Repensar binários e hierarquias	Desafiar o tratamento de gênero como algo binário, além de outros sistemas restritivos de contagem e classificação que perpetuam a opressão.
5. Abraçar o pluralismo	Valorizar o conhecimento que transita entre diferentes comunidades.
6. Considerar o contexto	Admitir que os dados não são neutros, mas funcionam como resultado de estruturas de poder consolidadas e relações sociais desiguais.
7. Tornar o trabalho visível	Reconhecer e valorizar o trabalho de profissionais em toda a cadeia de trabalho, normalmente feita a muitas mãos.

Fonte: D'Ignazio; Klein (2020) com adaptações da autora (2021).

Dentre todos os pontos, destaco principalmente o *sexto* – visto que os dados não estão em uma condição de simples existência no mundo, mas são elaborados por uma pessoa ou instituição que deseja investigar um tema ou uma perspectiva em detrimento de outra. Nesse sentido, a expressão “dados brutos” é paradoxal (GITELMAN; JACKSON, 2013). Como lembram as autoras, os dados não existem por si só no mundo, mas precisam ser coletados, compilados, gerados e processados. Ao esclarecer que a origem do termo “dado” vem do Latim *datum*, Rosenberg (2013, p. 18) explica que “dado” deriva do verbo *dare*, ou dar, e, assim, está relacionado à produção de um argumento – ou seja, pressupõe uma camada interpretativa. Fato, por sua vez, deriva do Latim *facere*, que tem como base algo que de fato ocorreu ou existe. “Há importantes diferenças aqui: fatos são ontológicos, evidência é epistemológica, e dados são retóricos.”<sup>62</sup> Isso não significa dizer, é claro, que os dados não devam ser utilizados; apenas faz-se necessário esclarecer de que forma se dá a sua produção e o seu contexto de uso. A clareza no método, preconizada pela ciência e executada pelo JGD, é uma das formas de se fazer isso.

<sup>62</sup> No original “There are important distinctions here: facts are ontological, evidence is epistemological, data is rhetorical.”

### 3.3.1 Modos de fazer: método e fluxos de trabalho

O Jornalismo Guiado por Dados tem como um de seus diferenciais a proximidade com o método científico. Por isso aposta, por exemplo, no emprego de hipótese (HUNTER *et al.*, 2013) no lugar da convencional construção de pauta. Essa hipótese servirá de ponto de partida para a investigação, que apresenta método próprio. Nesta seção, esclareço alguns pontos de interseção entre jornalismo e ciência e apresento exemplos fluxos de trabalho utilizados como método para o JGD.

Na ciência, de modo geral, o método é compreendido como um “[...] conjunto de regras básicas para desenvolver uma investigação com vistas a produzir novos conhecimentos ou corrigir e integrar conhecimentos existentes” (GIL, 2019b, p. 9). Ao fazer uso de procedimentos metodológicos que podem ser descritos e testados pelos seus pares, o cientista – ou jornalista, no caso do JGD – utiliza recursos para que sua subjetividade não comprometa o passo a passo dos procedimentos adotados.

O método jornalístico é diferente do científico, embora também funcione como uma busca estruturada por respostas. E, assim como na ciência – que parte do acúmulo de conhecimento viabilizado pelo método –, o jornalismo atua como descobridor de fatos em constantes desdobramentos. Como o jornalismo se viabiliza pela descrição e argumentação, empenha-se para mostrar que a verdade não é apenas uma condição subjetiva sua, mas trabalha objetiva e metodicamente para afastar os riscos de engano ou erro (GOMES, 2009).

Seguindo as discussões deste capítulo, cabe pontuar, mais uma vez, que o conhecimento produzido pela ciência e pelo jornalismo são distintos, ainda que ambas sejam formas sociais de conhecimento (MEDITSCH, 1992). Em termos práticos, existem algumas características marcadamente diferentes entre esses campos sociais de produção de conhecimento. A primeira e talvez mais significativa seja o tempo: enquanto o jornalista apresenta prazos apertados e pouco tempo de apuração, o cientista tem potencial para testar suas hipóteses ao longo de um período maior. A esses fatores, Franciscato (2006) adiciona outros pontos significativos, incluindo o fato de o cientista e o jornalista pertencerem a comunidades diferentes, com suas próprias regras, normas e valores.



O autor ainda aponta algumas especificidades do método jornalístico:

No jornalismo, podemos nos referir à existência de um método como conjunto de regras procedimentais de apuração de um fato, baseadas em valores e concepções sobre o papel e as formas de atuação jornalística para atingir um determinado fim (produzir notícias) sobre fatos e opiniões considerados de interesse público ou diversionais. Portanto, um método de apuração jornalística significa um conjunto de práticas (habilidades e técnicas) executadas pelos jornalistas, como observação de eventos em desdobramento, entrevista com fontes de informações, leitura de documentos e outros dados de natureza técnica etc. (FRANCISCATO, 2006, p. 5)

Além disso, o trabalho do jornalista ocorre dentro de uma estrutura organizacional, e por isso está sujeito a enquadramentos. Outra característica diz respeito ao recorte do objeto: o jornalista olha para o fato social sem ser especialista no tema no qual está imerso; o cientista, por sua vez, trabalha com o acúmulo de conhecimento por meio de uma pequena contribuição a uma área bastante específica.

Na atualidade, com a facilidade de acesso aos dados públicos e linguagens de programação, o Jornalismo Guiado por Dados busca se consolidar como um braço do jornalismo que se propõe a levantar informações que possam ser contextualizadas a partir do teste de hipóteses, do cumprimento de método objetivo e da apresentação de resultados – neste caso, a metodologia ou o código-fonte aberto, como veremos no capítulo sobre os elementos de transparência – como forma de se aproximar do que é feito na ciência, a partir da revisão entre pares, como observado por Träsel (2014).

Pesquisadores e praticantes do Jornalismo Guiado por Dados apresentam diferentes perspectivas do que consideram o conjunto ideal de procedimentos. Em linhas gerais, veículos e autores concordam com uma estrutura básica: tudo começa com uma afirmação ou pergunta que precisa ser testada a partir de dados publicamente disponíveis acerca de um tema de interesse coletivo (BRADSHAW, 2013; BELISÁRIO *et al.*, 2020). Suas raízes científicas empurram o Jornalismo Guiado por Dados na direção da investigação a partir de hipóteses – o primeiro passo é estabelecer uma pergunta norteadora para então testá-la com fontes abertas, compreendendo se existe possibilidade de explorar determinado tema. Em trabalhos anteriores (GEHRKE; FREITAS, 2021), defendi que o uso de hipótese também auxilia na busca pelo contexto de uma informação, permitindo o aprofundamento de um tema.

Nas próximas páginas, apresento algumas das principais propostas de métodos utilizados na área e, no capítulo final da tese, elaboro uma proposta de abordagem para aproximar o método do Jornalismo Guiado por Dados da ciência, com foco na transparência e na reprodutibilidade. Ressalto que a sistematização das informações apresentadas nesta seção tem como base o modelo do jornalismo digital e, portanto, algumas etapas não poderão ser integralmente cumpridas em outros meios, como o rádio e a televisão – ainda que boa parte do processo de apuração se repita independentemente da plataforma final de veiculação da notícia.

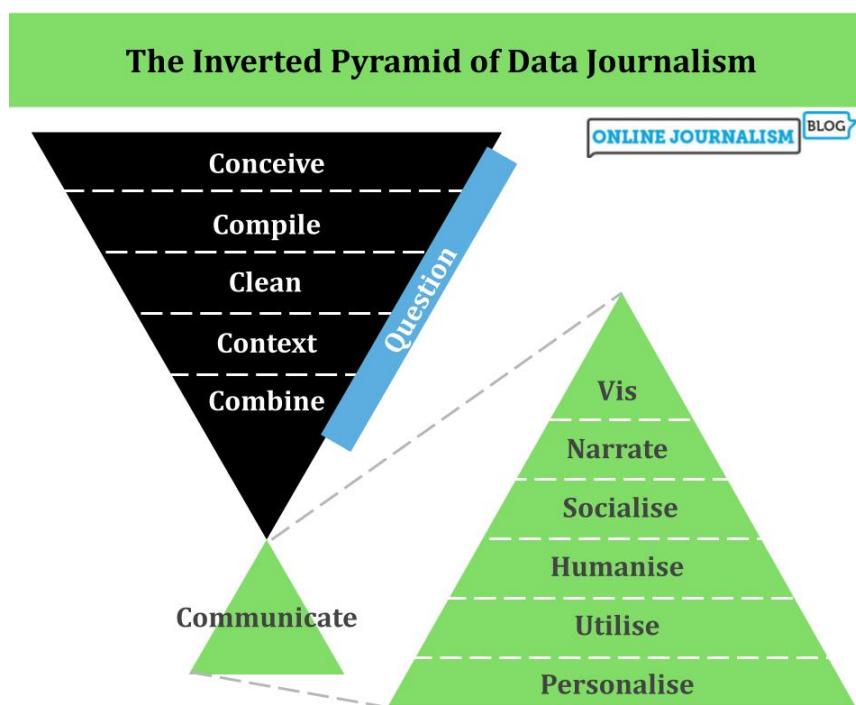
Propostas de fluxos de trabalho com dados vêm sendo estabelecidas desde o Jornalismo de Precisão. Meyer (2002), por exemplo, sugere as seguintes etapas para que o jornalismo se aproxime da ciência: *coleta* – busca de dados por meio de método científico; *armazenamento* – compreende a organização e as formas de armazenamento de dados; *recuperação* – pressupõe consultar dados disponibilizados por terceiros e oferecer formas de consulta (usabilidade e personalização); *análise* – congrega testes estatísticos e outras formas de investigar padrões em relação a um tema; *redução* – seleção do que é mantido na análise e do que não vai ser abordado; *comunicação* – apresentação dos dados, o que no JGD costuma ocorrer por meio do uso de gráficos, também chamados de visualizações<sup>63</sup>.

A Escola de Dados, uma das principais organizações que treina jornalistas no Brasil e no mundo, desenvolveu metodologia própria – o chamado *Data Pipeline* (BELISÁRIO *et al.*, 2020) – para organizar as etapas que formam um método: *pergunte, encontre, obtenha, verifique, limpe, analise e apresente*. A principal diferença em relação à proposta de Meyer diz respeito à inclusão de uma pergunta como ponto de partida para a obtenção de dados relacionados ao tema que se pretende abordar. Em consonância a essa lógica, modelo discutido previamente pelo professor Paul Bradshaw (2013) também enfatiza a existência de uma pergunta norteadora da investigação.

---

<sup>63</sup> A visualização de dados é uma área à parte no Jornalismo Guiado por Dados, envolvendo a etapa de comunicação e apresentação de números e narrativas por meio de gráficos. Não é o foco desta tese e, por isso, o tema não será aprofundado.

Figura 4 – Pirâmide invertida do Jornalismo Guiado por Dados



Fonte: *Online Journalism Blog*. Captura de tela em 4 jun. 2021.

Bradshaw utiliza um conceito bastante popular no jornalismo – a pirâmide invertida<sup>64</sup> – para explicar o procedimento de trabalho com dados (**Figura 4**). O autor distribui o passo a passo entre duas pirâmides. A primeira inclui a *concepção*, a *compilação*, a *limpeza*, a *contextualização* e a *combinação* (de bancos de dados) – todos esses processos são orientados por uma pergunta inicial. A segunda pirâmide (de pé), por sua vez, tem como destaque a *comunicação* das informações, que começa com a *visualização*, passando pela *narração*, *socialização*, *humanização*, *utilização* e *personalização* – todas essas são formas de envolver a audiência e promover interações a partir das narrativas apresentadas. No quesito *socialização*, por exemplo, leitores estão aptos a compartilhar uma história em suas redes sociais; a *utilização*, por sua vez, pressupõe uma participação ativa, como a possibilidade de fazer requisições via legislação de transparência em cada país.

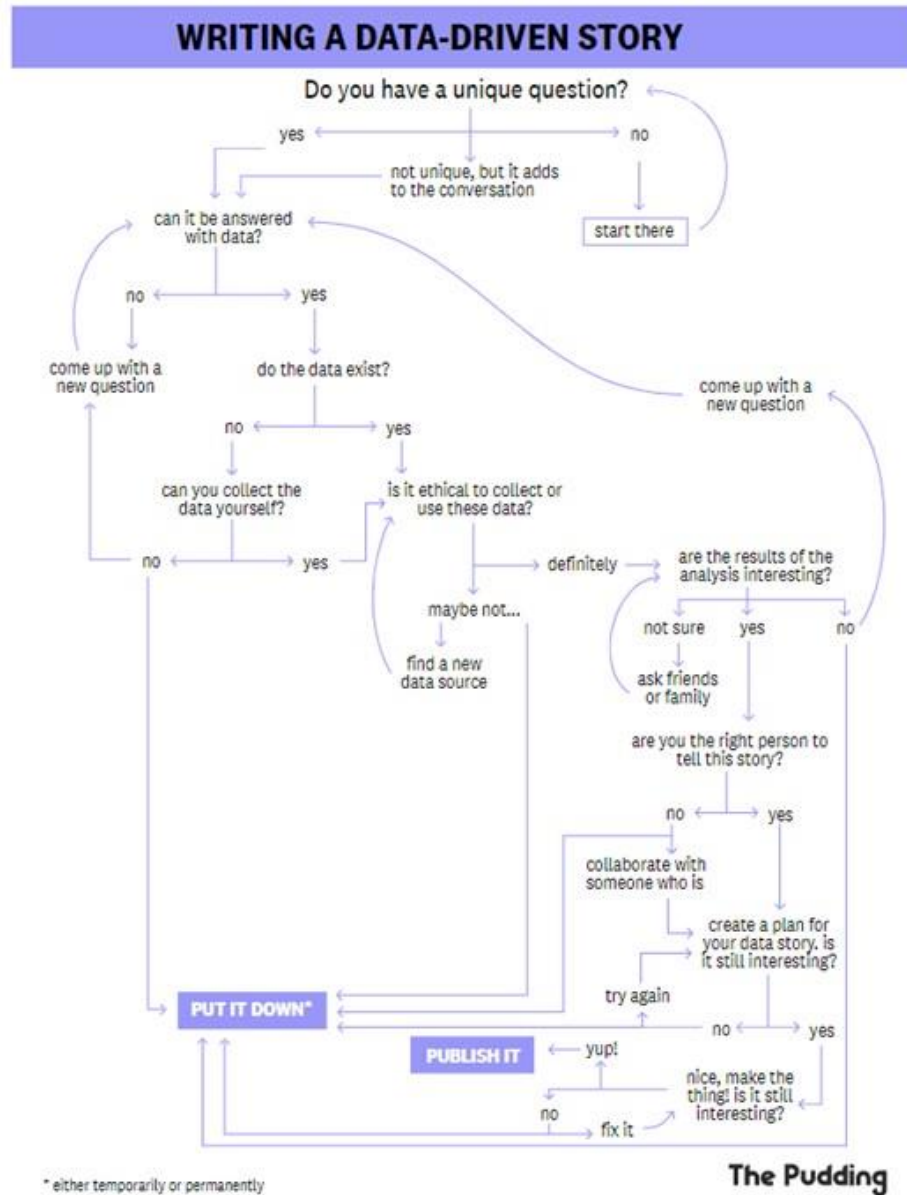
Iniciativas jornalísticas que surgiram no meio digital – os chamados nativos digitais – também apresentam sua própria configuração de método. O veículo estadunidense *The Pudding*, por exemplo, constrói suas narrativas a partir de um fluxograma<sup>65</sup> – que

<sup>64</sup> A pirâmide invertida é um formato textual para notícia que foi incorporado por redações em um formato ainda associado ao jornalismo industrial, em que o fato mais relevante ocupa o título e o primeiro parágrafo do texto.

<sup>65</sup> Disponível em: <https://pudding.cool/process/pivot-continue-down/>

lembra muito um algoritmo, visto que propõe algumas respostas binárias – sobre como seus profissionais devem conduzir uma história guiada por dados.

**Figura 5 – Fluxograma de trabalho do *The Pudding***



Fonte: *The Pudding*. Captura de tela em 15 jun. 2021.

Como é possível observar na **Figura 5**, a existência de uma pergunta inédita, ou ao menos fora dos assuntos normalmente discutidos na imprensa, é o primeiro passo para a construção de uma reportagem. Aspectos como a existência dos dados que respondem essa pergunta, ou sua disponibilidade e facilidade de coleta, também parecem

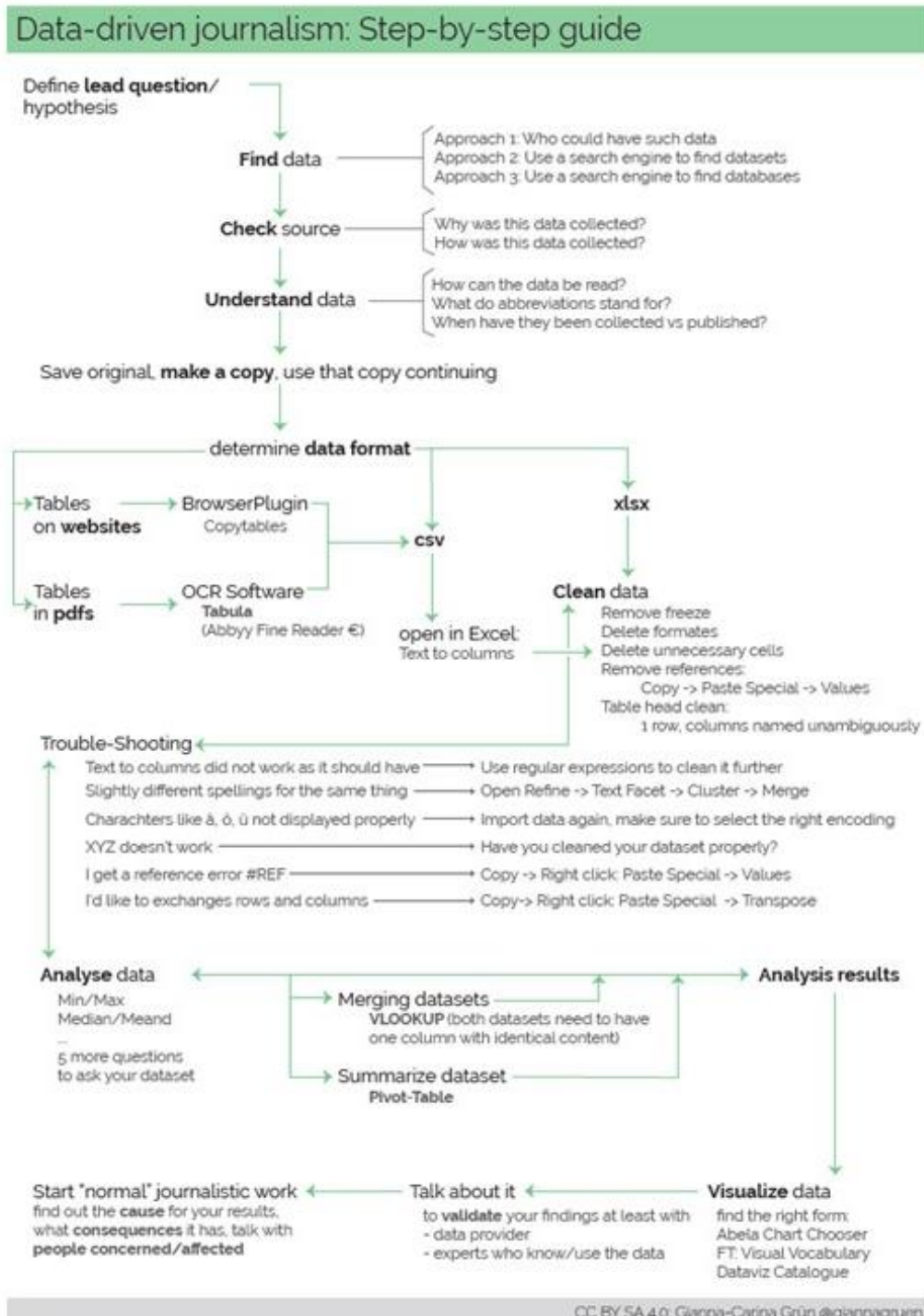
fundamentais para dar continuidade ao processo. A depender do cenário, uma investigação não necessariamente é descartada, mas pode sofrer mudanças para a execução.

Similarmente, a equipe de Jornalismo Guiado por Dados do veículo alemão *Deutsche Welle*<sup>66</sup> apresenta seu próprio método para encontrar histórias em bancos de dados. Existem etapas principais e adjacentes, que funcionam como desdobramentos da questão principal. Em linhas gerais, os procedimentos envolvem: *definir uma questão principal ou uma hipótese, encontrar os dados, verificar a fonte, compreender os dados* – e, depois disso, executar uma série de procedimentos técnicos que envolvem a determinação do formato de arquivo com o qual se pretende trabalhar –, *limpar os dados, analisar os dados, analisar os resultados obtidos, visualizar os dados, validar os resultados* e, por fim, *complementar este trabalho com técnicas de reportagem*.

---

<sup>66</sup> Disponível em: <https://github.com/gcgruen/Data-Journalism-Resources>

Figura 6 – Fluxograma de trabalho da *Deutsche Welle*



Fonte: *GitHub de Gianna Grün*. Captura de tela em 15 jun. 2021.

Em consonância com a abordagem do jornalismo em rede, do excesso de informação online e da necessidade de fazer sentido das informações por meio de uma abordagem científica, como propunha Meyer, Wihbey (2019) sugere a criação de princípios e padrões do fluxo de trabalho no jornalismo na era das redes. Em resumo, seguindo a lógica das

demais propostas aqui apresentadas, o profissional deve formular uma *hipótese* a partir de informações pré-existentes, buscar o *engajamento* das pessoas à medida que realiza entrevistas e busca documentos complementares. Deve, ainda, *mapear* os discursos existentes, *considerar o contexto* e *sintetizar* as informações encontradas. Por fim, tem de *analisar os dados*, *avaliar as evidências* e *praticar a transparência* – aspecto que não chega a ser detalhado pelo autor.

Dentre os modelos aqui apresentados, existe semelhança processual na tentativa de estabelecer um método próprio de trabalho no Jornalismo Guiado por Dados. Sob meu ponto de vista, o modelo da *Deutsche Welle* (**Figura 6**) é o que mais se aproxima da série de procedimentos que se repete na ciência, visto que prevê, em suas etapas finais, a validação dos resultados. Tal validação, em que os resultados são discutidos na equipe e entre especialistas que trabalham com o tema abordado, é um procedimento mais próximo da revisão por pares na ciência, o que normalmente ocorre na submissão de artigos científicos e até mesmo em congressos de uma área, em que há retorno dos pares por meio da discussão de trabalhos.

De maneira geral, as propostas apresentadas aqui ainda são insuficientes quanto à transparência e à abertura de dados, essencial ao processo científico no qual o JGD se baseia. As propostas também são limitadas quanto à checagem das informações, cujos números costumam ser interpretados literalmente pelos jornalistas quando apresentados por fontes oficiais, mas nem sempre exaustivamente verificados (BULLA, 2015; LAWSON, 2021). Ao observar esta lacuna, discuto uma proposta de método no **Capítulo 6**. De toda forma, é preciso reconhecer os esforços de praticantes e pesquisadores na criação de fluxos de trabalho a serem seguidos e demonstrados pela unidade do método (LIPPMANN, 1920). Como veremos no capítulo sobre transparência, este é um valor fortemente associado ao *ethos* do jornalista de dados, e não necessariamente às redações.

#### 4 TRANSPARÊNCIA COMO ABERTURA DO MÉTODO E DE DECISÕES EDITORIAIS

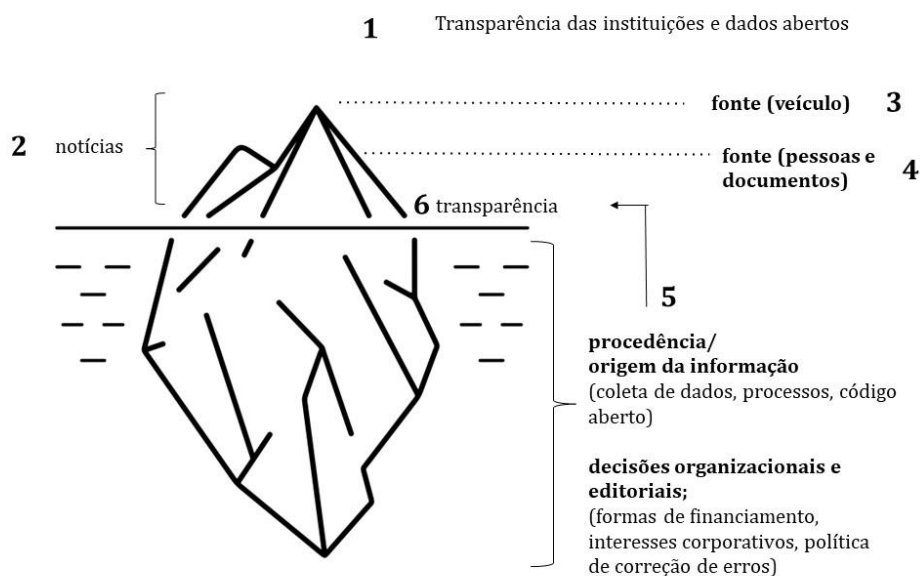
O Jornalismo Guiado por Dados é, sobretudo, jornalismo. E para que o jornalismo cumpra seus princípios fundamentais, como informar de modo qualificado e, a partir daí, gerar conhecimento, deve se comprometer com três princípios norteadores: a **verdade**, a **objetividade** e a **transparência**. Neste capítulo, cada um desses elementos será discutido individualmente, embora todos estejam conectados.

A verdade funcional, provisória e verificável, é a base de tudo; é o que garante ao jornalismo o amparo nos fatos, na realidade. A objetividade, por sua vez, diz respeito ao método que o jornalismo utiliza para acessar e relatar ao público as informações do mundo real. A transparência, então, é a abertura e a comunicação de condutas e procedimentos. Em nível organizacional, pressupõe a abertura de fontes de financiamento e política de correção de erros; em nível operacional, explica, por exemplo, como uma notícia foi feita, por que o tema importa, quais fontes foram consultadas, quais os resultados da análise e a existência de eventuais limitações.

A transparência, ponto central desta tese, pode ser comparada à parte visível de um *iceberg*: é o que o jornalista ou o veículo de comunicação decidem tornar visível; é o que emerge de decisões normalmente corriqueiras, tomadas dentro das organizações jornalísticas, e não abertas ao público. Em um cenário de cobrança por transparência e responsabilidade das instituições, é esperado que o jornalismo não apenas aponte os problemas sociais, mas saiba comunicar suas próprias escolhas e processos.



**Figura 7 – Modelo iceberg de transparência**



Fonte: A autora (2021), com ícone do *Noun Project*.

Em um contexto de transparência das instituições e dados abertos (1), o jornalismo, sendo também uma dessas instituições, trabalha com o compromisso de revelar seus processos. O conjunto de procedimentos pode ser comparado a um *iceberg*. A parte visível é notada a partir da publicação de notícias (2), em que as fontes se apresentam em dois níveis: (3) organizacional, ligado à ideia do veículo como fonte; e (4) operacional, relacionado às fontes ouvidas e consultadas pelos jornalistas para a produção do conteúdo. Em relação à apuração jornalística, um conjunto de processos (5), incluindo a procedência da informação (que, por sua vez, engloba a coleta de dados, o estabelecimento de métodos e processos, entre outros), em geral é desconhecido pelo público, situando-se na parte submersa do *iceberg*. As decisões organizacionais e editoriais estão presentes e são a base dessa seção, pois norteiam a prática jornalística. Quando existe clareza nas escolhas e nos métodos adotados – neste ponto se sobressai a objetividade – tais informações, que têm a verdade como base, retornam à superfície como transparência (6).

Percebe-se, na **Figura 7**, que as fontes desempenham um papel central, visto que formam um importante parâmetro para a qualidade da informação produzida e esperada pela audiência (LACY; ROSENSTIEL, 2015). Matéria-prima para a extração de

informações, as fontes são pessoas ou documentos que apresentam conhecimento sobre algum fato ou tema abordado (GEHRKE, 2018a). Em minha dissertação, identifiquei o uso recorrente de fontes documentais, especialmente com o uso de estatística, em notícias que apresentavam técnicas de Jornalismo Guiado por Dados. Justamente por apresentar um peso elevado na origem e na qualidade da informação, as fontes devem ser examinadas com cautela (STEENSEN, 2019).

No Jornalismo Guiado por Dados, as fontes estão associadas à busca da verdade, à objetividade metodológica e à transparência, quando ocorre a comunicação de uma análise e a possibilidade de reproduzir os procedimentos que levaram a um resultado. Desta forma, a qualidade do dado científico e dos números oficiais são pontos essenciais na cobertura de acontecimentos como a pandemia de Covid-19. Além de se certificar de que as informações são acuradas e confiáveis, os jornalistas precisam aprender a interpretar e comunicar seu significado.

#### **4.1 Verdade**

A prática do jornalismo está assentada sobre o paradigma<sup>67</sup> da verdade (BROERSMA, 2013). Trata-se de uma verdade funcional (KOVACH; ROSENSTIEL, 2014), provisória e não filosófica. Está amparada em uma correspondência com a realidade, visto que o jornalista reporta – ou deve reportar – fatos que verdadeiramente ocorreram, aproximando o relato jornalístico do fato em si (LISBOA; BENETTI, 2015). Para observar e dizer a verdade, o jornalista apresenta uma posição de exterioridade aos fatos, de quem os observa de fora e, assim, é capaz de fornecer um relato verificado e credível (CORNU, 1994). Em sua rotina de descrição e argumentação relacionada aos fatos, o jornalismo se empenha para mostrar que a verdade não é apenas uma condição subjetiva, e por isso trabalha de forma objetiva e metódica para afastar-se de enganos (GOMES, 2009).

A verdade é a primeira obrigação dos jornalistas (KOVACH; ROSENSTIEL, 2014) e condição primária para uma informação qualificada (CORNU, 1994); sua comunicação está ligada à ideia de autoridade do campo – se algo foi publicado nas notícias, então é

---

<sup>67</sup> O paradigma jornalístico corresponde a um sistema de valores compartilhados que envolve coleta, interpretação e validação de informações (BROERSMA, 2013).

real, e o leitor poderá encontrar informações verdadeiras e verificadas nesse espaço. Enquanto mediador social, o jornalismo tem o fato real como objeto (GUERRA, 2008).

Trata-se de um princípio normativo basilar do jornalismo, que opera a partir de elementos que especificam e delimitam a sua prática. Para Franciscato (2005), o jornalismo pressupõe a existência de uma ideia de verdade que pode ser apreendida em seus principais aspectos, de modo a se tornar notícia por meio de técnicas jornalísticas. A fidelidade ao relatar as ocorrências também é vista como aspecto fundamental pelo autor, que defende a publicação de conteúdo que auxilie o leitor a compreender o tempo presente. Lisboa e Benetti (2015) afirmam que a verdade é a primeira condição para o jornalismo tornar-se uma crença verdadeira justificada, o que, no final, viabiliza a confiança do leitor na instituição jornalística. Se a confiança nas instituições está associada à necessidade de correr riscos em alguma medida, a transparência, como veremos neste capítulo, é uma forma de minimizar esse risco – ou ao menos fazer parecer que o risco é menor (KARLSSON, 2020).

A verdade como ponto de partida atravessa os procedimentos jornalísticos na busca pela informação; trata-se de um princípio que deve orientar a disciplina da verificação. “Uma notícia que não seja orientada para a verdade – nem que essa verdade seja parcial e provisória, e consciente de o ser – não será uma informação” (CORNU, 1994, p. 394). Em linhas gerais, a verdade da qual o jornalismo se ocupa pode ser compreendida como “[...] uma declaração do que é mais provável em relação às evidências disponíveis no momento”<sup>68</sup> (KOVACH; ROSENSTIEL, 2011, p. 32). Dessa forma, a evolução de um acontecimento traz novas evidências com o passar do tempo. Admitir que a verdade tem caráter provisório implica considerar a existência de fragmentos sobre o conhecimento acerca de um objeto. “A verdade supõe uma procura, uma revelação dos fatos ocultos, uma relação entre esses fatos, que permitam destacar uma imagem (porque se tratará sempre de uma imagem!) da realidade” (CORNU, 1994, p. 75).

A verdade factual é a base para a objetividade jornalística, como veremos a seguir, associando-se a uma dimensão ética, segundo Cornu (1994). Em resumo, o que é noticiado é tomado como verdade pelo público à medida que há confiança nesta ética e na dimensão profissional do repórter. “O pacto da mediação cobra do jornalismo obrigações proporcionais à sua importância. Verdade, honestidade, correção, lealdade, respeito, equilíbrio,

---

<sup>68</sup> No original “[...] a statement of what is most probable in proportion to the evidence available at the time”.

justeza, imparcialidade são todos os valores e princípios que devem orientar uma ética do jornalismo [...]” (GOMES, 2009, p. 87). Tal honestidade, diz Cornu (1994), não pode ser definida ampla e vagamente; é preciso que o leitor confie no que está sendo comunicado porque existe um método de trabalho para apurar informações; e tal método é apresentado quando existe transparência no jornalismo.

Dentre as 12 finalidades jornalísticas propostas por Reginato (2019), entendo que duas apresentam maior relação com a busca pela verdade. São elas: informar de modo qualificado, que corresponde a uma informação verificada, relevante, contextualizada, plural e envolvente; e verificar a veracidade das informações, isto é, apurar dados e checar informações para entregá-las ao público de forma precisa e em consonância com a verdade dos fatos. Como diz Cornu (1994), uma informação reflete aspectos da realidade, mas não é a verdade, em si. Desta forma, existe uma realidade que pode ser observada, apreendida e comunicada.

A defesa da publicação de fatos verdadeiros também está associada à profissionalização do jornalismo para evitar o aumento da crítica pública. No início do século XX, segundo Picardo (2020), havia a preocupação com a perda de receita iminente se houvesse falta de confiança no jornalismo. Para evitar danos à credibilidade, valores como equilíbrio e objetividade foram adotados. Assim, cabia aos profissionais oferecer informações baseadas em fatos. Antes disso, na Europa, a estruturação das informações tendo a verdade como ponto de partida ajudou os primeiros jornais a se posicionarem no mercado da informação diferenciando-se de outras formas de comunicação (BROERSMA, 2017).

Em 1939, época da Segunda Guerra Mundial, o *The New York Times* explicou aos seus leitores por que eles poderiam e deveriam acreditar nas notícias que o jornal publicava sobre a guerra, o que incluía fatos que ultrapassavam o território dos Estados Unidos. Segundo Kovach e Rosenstiel (2011), o veículo fez circular inúmeros anúncios em suas páginas de cobertura internacional explicando como as informações de fora do país eram recebidas e editadas. Assim, a audiência estaria familiarizada com as práticas jornalísticas e até mesmo com a ideia de verdade dos fatos que permeava aquela cobertura. Dessa forma, o jornal oferecia, sobretudo, informações acerca da natureza de uma evidência e como ela foi avaliada, demonstrando sua preocupação com a verdade e em comunicar ao leitor seus modos de fazer.

Cabe lembrar que a verdade jornalística é sempre uma aproximação, e todo relato jornalístico será mediado por uma interpretação (LISBOA; BENETTI, 2015); a

subjetividade, assim, é condição para a objetividade. Para as autoras, existem algumas estratégias para que o leitor seja capaz de identificar elementos que auxiliam a medir o grau de fidelidade entre o fato propriamente dito e o relato que ocorre a partir dele. Entre os exemplos citados estão elementos como a fotografia, o acesso às fontes especializadas e o detalhamento de eventos. Ao citar o percurso interpretativo realizado no jornalismo durante o conhecimento e a apuração de um fato, pensando nas técnicas cognitivas que demanda, Guerra (2008) considera que a verdade e a relevância são parâmetros de desempenho para os jornalistas e de qualidade para a construção da notícia. A interpretação e a apresentação dos fatos geram impacto sobre a efetividade do discurso jornalístico, que, segundo Broersma (2013), está na capacidade de descrever e produzir fenômenos sociais simultaneamente.

Dito isso, é preciso reconhecer que o jornalista age como intérprete da realidade a partir de fragmentos dos fatos. Como diz Cornu (1994, p. 332-333):

[O jornalista] Não tem um acesso directo ao cerne dessa realidade presente. Não pode apreendê-la na sua verdade profunda, que lhe escapa. Deve tentar decifrá-la através dos fragmentos vividos ou relatados, actas, discursos. A complexidade, as contradições, as incoerências, as lacunas da actualidade – esta apresenta-se como um tecido esburacado – são evidentes.

Este ponto trazido por Cornu se relaciona com o que dizem Kovach e Rosenstiel (2014) sobre a verdade jornalística, funcional e praticável no dia a dia. Para que seja possível executá-la, dois conceitos são fundamentais: acurácia, referente à exatidão das informações, e verificação, ligada ao processo de checagem dos fatos. A acurácia, segundo os autores, é o ponto definidor para outros elementos fundamentais às notícias: contexto, interpretação, debate e comunicação pública. Para eles, a disciplina da verificação dos fatos, viabilizada por meio da objetividade no método, é o que diferencia o jornalismo de outras práticas. Assim, está no cerne de sua produção de conhecimento.

O jornalismo que tem a verificação em seu núcleo apresenta como propósito responder perguntas sobre os fatos e colocá-los em um contexto em que possam ser compreendidos. Em entrevista a Gehrke e Mielniczuk (2017, p. 11), Philip Meyer afirmou que os leitores não merecem apenas os fatos, mas a verdade dos fatos: para garantir a confiabilidade das informações, os jornalistas precisam usar “cinto e suspensório”, fazendo uma analogia à necessidade de dupla checagem.

Por meio de entrevistas em profundidade com jornalistas canadenses, estudo de Shapiro *et al.* (2013) mostra que, para os profissionais, a verificação é um processo circular, que ocorre durante todo o processo de elaboração das notícias; não seria, então, uma etapa de trabalho ou um processo marcadamente separado, mas inerente à busca pela verdade dos fatos. Em linhas gerais, os autores perceberam que documentos oficiais costumam ser alvo de maior confiança dos profissionais. Por esse e outros motivos, o processo de triangulação de informações, que compreende a comparação entre fontes independentes, não funciona da mesma forma no jornalismo e nas ciências sociais. Segundo Shapiro *et al.* (2013), os jornalistas usualmente verificam detalhes para não comprometer a acurácia do texto (nomes, referências geográficas etc.), mas não checam afirmações ou fatos mais amplos, a não ser que se trate de um caso de difamação.

Autores como Broersma (2013) apontam que o paradigma jornalístico pode estar fraturado com a chegada de novos agentes à cadeia da produção e difusão da informação. Neste contexto, o jornalismo tem sua autoridade enfraquecida, e o público precisa concordar que as normas e os valores desta prática social ainda fazem sentido na realidade que o jornalismo compreende e comunica. De todo modo, a audiência busca as notícias em fontes confiáveis porque parte do pressuposto de que a verdade é um valor caro ao jornalismo, e que existe uma obrigação de veracidade. Sob este aspecto, Gomes (2009, p. 11) ressalta que “[...] só é notícia um ato verbal que comporte uma pretensão de ser verdadeiro” e, ao fazê-lo, o jornalismo assume o compromisso de buscar a verdade e encontrar formas de evitar erros.

Compreender o que se entende por verdade jornalística parece fundamental neste cenário de desinformação e desconfiança em relação aos fatos noticiados pela imprensa. A verdade é o princípio fundamental do jornalismo, e está necessariamente ligada à sua matéria-prima, pois corresponde à base dos produtos jornalísticos informativos. A apuração, a acurácia e a verificação em relação aos fatos é o que permite a interpretação jornalística e a aproximação entre a realidade dos fatos e o relato jornalístico. Isso é possível por meio da objetividade jornalística, ou seja, pela via do método.

## 4.2 Objetividade

A objetividade jornalística corresponde aos procedimentos metodológicos adotados na prática da reportagem. Compreende a busca e a coleta de informações sobre um tema, bem como os processos de verificação relacionados a ele. A verificação e acurácia jornalísticas estão associadas à confiança depositada pela audiência no jornalismo. Como diz Cornu (1994, p. 391): “Um laço que o jornalista deve defender para que nunca se rompa”. A objetividade permite ao jornalista acessar e apreender a verdade verificável, ou seja, realizar a aproximação com o real a partir do relato; trata-se, segundo Cornu (1994, p. 394), de um “[...] método de orientação para a verdade”. Assim, a objetividade é o meio encontrado pelos jornalistas para gerar conhecimento em sociedade (CORNU, 1994).

Equivocadamente, o conceito de objetividade foi rejeitado ao longo dos anos por pensadores que falavam sobre a impossibilidade de praticá-la por conta da subjetividade humana. No entanto, a subjetividade é condição para que a objetividade jornalística seja viabilizada, visto que toda interpretação, característica fundamental da mediação, contém aspectos subjetivos. Associo meu argumento ao de Cornu (1994), Kovach e Rosenstiel (2014), Sponholz (2009) e tantos outros autores para quem a objetividade está no método, e não no jornalista.

Segundo Sponholz (2009), o método pode ser descrito como uma série de ações planejadas e adotadas para atingir determinado objetivo. No caso do jornalismo, esse objetivo é conhecer a realidade e mediar as informações obtidas junto à apuração dos fatos, que posteriormente tornam-se produto. Em outras palavras, a objetividade garante que o discurso produzido no jornalismo seja fiel aos fatos (GUERRA, 2008).

Para Cornu (1994, p. 391), não basta que o jornalista defina, de forma ampla e vaga, a honestidade que planeja manter junto ao leitor ao comunicar os fatos. A confiança de que a verdade é parte fundamental da apuração está no rigor do método jornalístico – ou seja, na objetividade da qual estamos falando. “O rigor é ir ao fundo das investigações materialmente realizáveis, no tempo dado para essas pesquisas. É a recolha de todos os fatos confirmados disponíveis. É tudo o que se opõe à falsificação, à deformação, à mentira.”

Entre os procedimentos jornalísticos relacionados à objetividade metodológica estão uma série de tomada de decisões, como a escolha do tema, o enquadramento e a seleção das fontes. Desta forma, a observação e o uso de documentos e entrevistas podem ser

encarados como objetos de evidência da qual o jornalismo se utiliza (ANDERSON, 2018). A objetividade jornalística está assentada sobre uma dimensão ética que tem como base a verdade verificável. Assim, dadas as dificuldades de apreender características de um acontecimento, o jornalista deve prezar pela descrição correta dos fatos, publicando apenas as informações cuja origem conhece (CORNU, 1994).

Sponholz (2009) lembra que a produção de notícias deriva também de processos oriundos das rotinas produtivas dos jornalistas, o que inclui aspectos da organização para a qual trabalham. Em clássico texto em que aponta a percepção da objetividade pelos jornalistas profissionais como ritual estratégico, Tuchman (1972) admite que exercem influência sobre a objetividade jornalística aspectos como a forma – o modo de apresentação de elementos das notícias –, as relações organizacionais e o conteúdo produzido, ou seja, aquilo que é resultado da apreensão da realidade social. Os rituais estratégicos dos quais a autora fala no texto são procedimentos de rotina que os jornalistas apontam como provedores de objetividade. Em estudo recente, McNair (2013) diz compreender a objetividade como uma extensão de rituais estratégicos sustentados pela percepção de verdade e confiança dos leitores no jornalismo.

A unidade do método já era defendida por Lippmann (1920) nas primeiras décadas do século XX, em um cenário de profissionalização da prática jornalística. Para o autor, o bom trabalho de reportagem dependia de um exercício de virtudes científicas. O conhecimento e o acesso à verdade estavam no escopo da liberdade que julgava ser essencial. Assim:

Há apenas uma unidade possível em um mundo diverso como o nosso. É a unidade do método [...] é a unidade do experimento disciplinado. Há apenas um vínculo pacífico que é permanente e enriquecedor: o crescente conhecimento do mundo em que o experimento ocorre. Com um método intelectual comum e uma área comum de fatos válidos, as diferenças podem se tornar uma forma de cooperação e deixar de ser um antagonismo irreconciliável (LIPPMANN, 1920, p. 67)<sup>69</sup>

---

<sup>69</sup> No original “There is but one kind of unity possible in a world as diverse as ours. It is unity of method [...] the unity of the disciplined experiment. There is but one bond of peace that is both permanent and enriching: the increasing knowledge of the world in which experiment occurs. With a common intellectual method and a common area of valid fact, differences may become a form of cooperation and cease to be an irreconcilable antagonism”.



Para Lippmann, não há liberdade para uma comunidade com pouca informação. Similarmente, Maras (2013) destaca que a objetividade jornalística é fundamental para a manutenção da democracia, da liberdade de imprensa e da performance dos veículos jornalísticos na cobertura de temas de interesse público.

É nesta abordagem científica do método, discutida por Lippmann, que retomo a discussão sobre ciência e jornalismo. Segundo Sponholz (2009, p. 153), o que diferencia a ciência e o jornalismo do senso comum é a utilização de regras profissionais para apreensão e comunicação da realidade. Ambos, à sua maneira, processam e estruturam o que é vivenciado a partir de técnicas que confirmam confiabilidade ao processo. “[...] as fontes e o método utilizado devem ser transparentes para que outros jornalistas ou experts possam repetir o mesmo procedimento dentro das mesmas condições. Eles também devem chegar ao mesmo resultado.”

Sob esta perspectiva, jornalistas e cientistas compartilham um ideal de objetividade viabilizada pelo método. Os propósitos de uso, no entanto, são diferentes. Para os cientistas, conforme Galison (2015), importa desempenhar um modo de codificação de conhecimento compartilhado para acessar objetos de seu campo de pesquisa, enquanto os jornalistas buscam um meio de atender maior parte de sua audiência e de seus anunciantes, formalizando o método em um formato de texto muito específico – o jornalístico –, em que o elemento do ineditismo é priorizado e disposto em texto não apelativo.

Anderson (2018) observa em sua pesquisa que, desde o começo do século XX, a resposta do jornalismo aos momentos de crise epistemológica tem sido o comprometimento com a objetividade, abraçando a certeza profissional em momentos incertos. No entanto, nem sempre a objetividade foi bem-vista. Enquanto os jornalistas a utilizam como resposta às críticas, o argumento daqueles que não acreditam em sua eficácia é que supostamente a objetividade é um ideal inatingível. Maras (2013) pontua que o conceito de objetividade jornalística responde a uma série de condições comerciais, culturais e organizacionais em uma sociedade e, por essa razão, sofre alteração ao longo do tempo.

A partir dos anos de 1830<sup>70</sup>, o jornalismo passou a enfatizar a coleta dos fatos, a separação entre opinião e informação. Antes desse período, segundo Schudson (2010),

---

<sup>70</sup> O cenário descrito por Schudson (2010) tem como base a imprensa estadunidense. Ao longo da história, esta mesma imprensa serviu de base para outros países, a exemplo do Brasil.

esperava-se que os jornais apresentassem um ponto de vista político ou partidário<sup>71</sup>. A ideia de uma notícia apartidária surgiu juntamente com a primeira agência de notícias norte-americana, a Associated Press, em 1846, logo após a invenção do telégrafo e a possibilidade de comunicação com outras partes do mundo.

A objetividade passa a se estabelecer como um valor nas últimas décadas do século XIX, época em que havia o surgimento da universidade norte-americana, das associações profissionais e da admiração pela ciência. De acordo com Schudson (2010), os repórteres enxergavam a si próprios, naquela época, como cientistas desvendando fatos políticos e econômicos da vida industrial. À época, a objetividade era compreendida como declarações consensualmente validadas sobre o mundo, com a separação radical entre fatos e valores.

Após a I Guerra Mundial, e motivados principalmente pela propaganda de guerra e pelo surgimento das Relações Públicas, os jornalistas perderam a confiança nos fatos, submetendo-os a uma série de normas e procedimentos. A objetividade enquanto método consistente para testar as informações aparece nos anos de 1920, nos Estados Unidos (KOVACH; ROSENSTIEL, 2014). Tal abordagem, afirmam os autores, permitiria que houvesse transparência em relação às evidências, evitando que possíveis vieses pessoais e culturais comprometessem a acurácia do trabalho dos repórteres.

Nos anos de 1930, ocorre o desenvolvimento da reportagem interpretativa, a partir da demanda dos leitores pelo aprofundamento e a interpretação diante da complexidade do mundo (SCHUDSON, 2010). A Sociedade Americana dos Editores de Jornais decidiu, em 1933, que deveria haver maior espaço e dedicação ao jornalismo interpretativo e explicativo com a apresentação de informações de bastidores para que os leitores pudessem entender a importância dos eventos da época (BARNHURST, 2014). Naquela década, a prática jornalística tornava-se mais visual e contextual (ANDERSON, 2018). Apesar do aumento nas notícias interpretativas de 1930 a 1950, nos anos de 1960 a objetividade tornou-se um termo amplamente utilizado e algo supostamente inatingível. Isso ocorre em um cenário de crise institucional generalizada no jornalismo e na época de desenvolvimento do *New Journalism*, com a ideia de um jornalismo próximo da literatura. De acordo com Pickard (2020), a maior parte das discussões sobre a objetividade jornalística

---

<sup>71</sup> É possível observar que a ideia equivocada de neutralidade está presente junto ao surgimento do princípio da objetividade. Trata-se, no entanto, de um valor diferente: enquanto a neutralidade é entendida como a abstenção de expor a própria opinião, a objetividade serve de pressuposto para uma cobertura objetiva.

ocorre em um período de democratização da sociedade dos EUA e mudança na postura dos jornalistas em relação à sua posição social. Contudo, alguns revisionistas históricos atribuem origens meramente econômicas para a objetividade jornalística: satisfazer os anunciantes.

Visto que os fatos e os acontecimentos são os objetos do jornalismo (GUERRA, 2008; SPONHOLZ, 2009), a objetividade tem duas implicações éticas principais, conforme Guerra: torna-se um parâmetro a ser buscado e referência para os profissionais, assim como cria elementos de aferição que respaldam a confiança do público na instituição jornalística.

Torna-se um dado inquestionável, portanto, que a imprensa tem como objeto um fato, que dá início à notícia e que norteia o trabalho dos profissionais tanto no que diz respeito à demanda e à credibilidade que vem do público – afinal atesta a validade do imperativo – quanto no que diz respeito ao objeto – a ocorrência ou a “trama” – em torno do qual a instituição se organiza. **Negar isso seria negar a existência da atividade** (GUERRA, 2008, p. 102, grifo meu).

Por isso, a verdade verificável, anteriormente abordada neste capítulo, é fundamental para a compreensão da objetividade. A ausência da objetividade, diz Cornu (1994, p. 393), abre caminho para o “vale tudo” no jornalismo, gerando um jornalismo “de aproximação” e “de preguiça”. Refutar a existência da objetividade corresponde a um argumento falacioso, frágil e improdutivo à medida que a realidade só pode ser apreendida por meio da linguagem (CORNU, 1998) e toda apreensão dos fatos demandará algum tipo de interpretação, que deve estar justificada no próprio discurso jornalístico (LISBOA; BENNETTI, 2015).

Sponholz (2009) propõe que a objetividade seja entendida como uma representação da realidade, originando-se a partir da relação entre as realidades social (primária) e midiática (secundária), enfatizando que sua concepção epistemológica está estruturada na verdade como correspondência dos fatos. Neste sentido, adiciona-se o elemento da verificação dos fatos, que serve de base para o jornalismo. Somente a partir da checagem das informações é que os fatos se transformam em produto jornalístico.

De forma geral, o termo “objetividade” ainda parece gerar dúvida entre os leitores. Relatório escrito por Toff *et al.* (2021) sugere que valores como objetividade, imparcialidade e equilíbrio normalmente são acionados pelos leitores ao falar sobre confiança nas

notícias. No entanto, esses termos geram confusão entre si. Para alguns leitores, ser objetivo significa manter opiniões fora das notícias; para outros, tem a ver com ampliar a transparência nos processos jornalísticos ao lidar com opiniões. A objetividade deve, sobretudo, estar associada ao método e à verificação das informações.

Além disso, o uso de números frequentemente está associado a uma suposta objetividade e abertura para confiança. Autores como Bulla (2015) discutem que há uma predisposição em aceitar o número como realidade objetiva tanto em termos de produção quanto de recepção de notícias. Ao aplicar Análise de Discurso em publicações do jornal *O Estado de S. Paulo* e entrevistar repórteres, Bulla identificou que existe uma apropriação dos números para a produção de argumentos e construção de sentidos conforme convém à organização jornalística. Assim, a autora chegou à conclusão de que o uso de números potencializa o valor da notícia por meio da quantificação da informação. A quantificação nas notícias, conforme Lawson (2021), gera simultaneamente um senso de magnitude, proporcionalidade e especificidade – efeito que só os números são capazes de gerar, visto que muitas vezes não se consegue obter o mesmo resultado apenas por meio das palavras.

Via de regra, os leitores não conseguem experienciar, em primeira mão, os acontecimentos que dão origem à notícia, recorrendo, então, ao jornalismo para acessar esses dados da realidade (McNAIR, 2013). Existe confiança de que o jornalismo apresentará informações apuradas e verificadas ao público. Assim, a disciplina da verificação, sobre a qual falam Kovach e Rosenstiel (2014), é fundamental para que a objetividade como método seja cumprida. A verificação, segundo os autores, reside no jornalismo como papel central – sua demonstração se dá pela via da transparência.

Em consonância com o que defendem Kovach e Rosenstiel, Cornu (1994) entende que o jornalismo deva prezar pela manutenção da verificação como cuidado ético central à prática jornalística, como algo do qual não se deve abrir mão. Para o autor, a efemeridade e a rapidez do desenrolar dos fatos devem ser alvo de atenção para que a objetividade não seja rompida pela velocidade e pela pressa; a objetividade, acompanhada do rigor metodológico, é o que de mais importante apresenta o jornalista ao leitor.

A verificação em tempos de jornalismo em rede e da velocidade que lhe é inerente é objeto de preocupação de autores como Cornu (1994) e Kovach e Rosenstiel (2014). Em consonância com a discussão sobre o acontecimento, que ocorre e é ressignificado à medida que o jornalismo fornece novos elementos sobre suas evoluções, pode-se refletir sobre a condição de pseudoambiente da qual fala Lippmann (1920). Uma vez que (na maior

parte dos casos) os leitores não têm acesso às informações primárias, a percepção da realidade concreta ocorre a partir de uma realidade apreendida e construída pela imprensa.

Trazendo o valor da objetividade como método para a atualidade, Anderson (2018, p. 160) indica o caminho do Jornalismo Guiado por Dados neste percurso. Após acompanhar a história do Jornalismo de Precisão e do desenvolvimento de práticas quantitativas similares ao longo da história, afirma que, atualmente, pode-se pensar sobre um conceito que chama de “objetividade de segunda ordem”, viabilizado pelo jornalismo estruturado. Compreendendo que o jornalismo estruturado é uma atualização e uma extensão do pensamento computacional na prática jornalística, avalia que as provas de certeza – e, portanto, de objetividade – não são apenas demonstradas, mas estão presentes na base de dados semântica que dá origem a este novo termo. O principal exemplar do jornalismo estruturado – uma versão ainda mais moderna do que o Jornalismo Guiado por Dados – é o projeto *Structured Stories*<sup>72</sup>, que opera com base de dados semânticas para oferecer ao leitor um nível diferenciado de aprofundamento nas notícias. São informações que recebem hiperlinks e admitem referências cruzadas. “Mostrar como esses objetos são codificados em um formato digital pode compensar a incerteza final da ‘realidade’ [...] a incerteza é eliminada à medida que a linha entre ‘fato’ e ‘verdade’ é redirecionada pelo banco de dados”<sup>73</sup> (ANDERSON, 2018, p. 164).

O emprego de números no jornalismo, porém, não se restringe ao Jornalismo Guiado por Dados ou suas formas mais modernas. Em pesquisa que discute o uso de bancos de dados pelos jornalistas britânicos na cobertura da crise humanitária na Europa, Lawson (2021) observa que existe confiança excessiva depositada pelos profissionais em estatísticas concedidas por fontes oficiais que seriam “apolíticas”, como a Organização das Nações Unidas (ONU) e organizações não governamentais de menor espectro. Assim, conclui que os jornalistas se escondem atrás de suas fontes, raramente verificando ou questionando os números apresentados. O que foi detectado por Tuchman há 50 anos, portanto, perdura no jornalismo apesar da digitalização e da facilidade de acesso às fontes: o uso acrítico de “evidência da evidência”, diz Lawson, permite que os jornalistas se

---

<sup>72</sup> O projeto *Structured Stories* pode ser considerado uma base de dados sobre fatos e acontecimentos, fugindo da forma tradicional de divulgação de notícias. Trata-se, de acordo com a descrição do projeto, de um repositório que permite coletar, utilizar e incrementar permanentemente novas histórias para cobertura. Disponível em: <http://www.structuredstories.com/#/intro>.

<sup>73</sup> No original “Showing how these objects are encoded in a digital format can compensate for the ultimate uncertainty of ‘reality’ [...] uncertainty is eliminated as the line between ‘fact’ and ‘truth’ is rerouted through the database.”

protejam das críticas caso utilizem estatísticas incorretas. Como não são eles os responsáveis por gerar tais estatísticas, existe a impressão de que basta culpar a entidade geradora dos números.

No início deste capítulo teórico, argumentei que, para tornar-se conhecimento, o jornalismo precisa cumprir esta tríade de valores essenciais: verdade, objetividade e transparência. Compreendo que o Jornalismo Guiado por Dados tem potencial para gerar um tipo de conhecimento específico pela via da transparência metodológica, cumprida a partir da objetividade. Na versão mais recente de seu livro sobre os elementos do jornalismo, Kovach e Rosenstiel (2014) pontuam justamente esta questão: que o surgimento de práticas como o JGD está mais próximo da promessa de fornecer um caráter científico à prática jornalística. Alertam, no entanto, que a autoridade não está em simplesmente mostrar os números e conquistar a audiência a partir de sua exatidão, mas da possibilidade de os números contribuírem para a criação de uma disciplina mais rígida de verificação. Os dados, assim, exigem maior transparência metodológica.

### 4.3 Transparência

A transparência no jornalismo corresponde, em linhas gerais, ao processo de **abertura** de informações da organização jornalística e dos processos por ela desenvolvidos. Sob este aspecto, pode-se considerar que envolve procedimentos anteriores à reportagem, mas ainda inerentes ao *modus operandi* de uma organização jornalística, como suas fontes de financiamento, políticas de diversidade e correção de erros. Em termos operacionais, abrange principalmente a clareza do método adotado e das fontes utilizadas, permitindo que o leitor acompanhe o que foi publicado e seja capaz de verificar e eventualmente reproduzir o passo a passo da reportagem.

Apresento, nesta seção, o argumento de que a transparência está dividida em duas categorias principais: como **operadora de uma prática ética** e **recurso discursivo de legitimação**. Busca-se, por meio dela, promover uma relação de respeito junto à audiência, reforçando laços de confiança e ampliando a percepção de credibilidade (JOHNSON; ST JOHN, 2021; KARLSSON, 2020; ZELIZER, 2019). Se a confiança parte de um processo em que uma das partes apresenta credibilidade suficiente para estimular essa sensação em quem a percebe, a transparência no jornalismo é também uma forma de minimizar riscos, já que toda relação de confiança implica lidar com o incerto em alguma medida (KARLSSON, 2020).

No jornalismo digital, a transparência vem sendo estudada há cerca de 20 anos, impulsionada pela publicação de conteúdo na Web. Karlsson (2020; 2022) identifica três ondas de transparência no jornalismo. A primeira, batizada pelo autor de “fundação”, corresponde ao início dos anos 2000 e é uma adaptação de contextos institucionais mais amplos deste tópico, como a demanda por abertura de processos de instituições públicas e como isso poderia se refletir no jornalismo. A segunda onda, chamada de “consolidação” e desenvolvida de 2010 a 2016, insere o conceito de transparência junto às redes sociais digitais, atualizando sua desenvoltura neste cenário. Também passa por um processo de institucionalização à medida que surge nos códigos de ética. De 2016 em diante, Karlsson (2022) enxerga uma terceira onda – atualmente em curso – que associa aos termos “esperança” e “escrutínio”, em que a transparência é vista como uma forma de permitir que o público avalie instituições socialmente estabelecidas. Tal cenário é um reflexo da desconfiança motivada em parte pelo excesso de desinformação disponível no cenário digital

e mobilizada por políticos e outras figuras públicas, ampliando a relação de confiança de parte da população em instituições tradicionalmente consolidadas.

A transparência jornalística pode ser vista sob diferentes ângulos. Explicada de forma ampla, trata-se de tornar públicas as informações usualmente privadas que influenciam no processo de elaboração de notícias (ALLEN, 2008). Também tem sido encarada como uma nova norma ou uma nova roupagem à objetividade (WEINBERGER, 2009) e vista com o sentido de abertura (*openness*) dos processos de reportagem (KARLSSON, 2010), assegurando o cumprimento de procedimentos objetivos que levam à verdade dos fatos e permitem que a audiência seja capaz de compreender tais evidências como verdadeiras.

Ainda que o sentido de abertura tenha prosperado nos estudos de transparência principalmente a partir de Karlsson (2010), o autor recentemente mudou de abordagem, afirmando que é mais útil apresentá-la como uma “visibilidade gerenciada estrategicamente” (KARLSSON, 2022, p. 78), visto que serve aos propósitos de um ator social (a organização jornalística, neste caso). Dessa forma, o autor atualiza a percepção de que a transparência serviria para corrigir uma assimetria, visto que os jornalistas têm acesso às informações que utilizam para produzir conteúdo, e que não são reveladas em sua totalidade ao público.

Compreende uma dimensão ética do jornalismo (KOLISKA; CHADHA, 2018; PLAISANCE, 2007), com iniciativas de *accountability* associadas à legitimidade (ALLEN, 2008) e credibilidade (ZELIZER, 2019). Sendo também uma forma de diferenciar o discurso jornalístico de outros discursos, a transparência pode ser vista como uma inovação normativa (GADE *et al.*, 2018), uma prática que visa delimitar a atuação do jornalismo em relação a outras áreas (ALLEN, 2008) e um efeito a ser mostrado para o público (LISBOA; FRANZONI, 2018). A transparência se apresenta como princípio normativo (SEIBT; FONSECA, 2019) para a prática de gêneros jornalísticos como a checagem de fatos, o *fact-checking*, em que o jornalista confronta discursos públicos com fatos, atribuindo graus de veracidade sobre um conteúdo utilizando-se de método objetivo (GRAVES, 2016). A transparência é defendida por Kovach e Rosenstiel (2014) como o aspecto-chave da disciplina da verificação.

Como tem sido verificado em produtos jornalísticos digitais, a transparência pode ser considerada uma estratégia para obter engajamento público, ampliar o conhecimento compartilhado e demonstrar valores jornalísticos, repaginando práticas e condutas que



normalmente são associadas às redações convencionais e analógicas, como a figura do *ombudsman*, cujo trabalho é fazer uma ponte entre o veículo e a audiência (MEIER, 2009).

Ao orientar a sociedade sobre o que deve esperar acerca do consumo de informações, Kovach e Rosenstiel (2011) enumeram seis pontos a serem observados. A audiência deve se questionar, portanto, sobre os seguintes tópicos: 1) que tipo de conteúdo está encontrando (jornalismo ou publicidade); 2) a completude das informações; 3) as fontes jornalísticas e sua credibilidade; 4) a apresentação de evidências e a possibilidade de testá-las; 5) a existência de explicações alternativas ao fato noticiado; 6) o aprendizado de informações sobre o fato noticiado.

Uma vez que – idealmente, com base em sua deontologia profissional – os jornalistas buscam a verdade dos fatos e a comunicam para o seu público respeitando as regras da objetividade, a busca pela transparência garantiria o acesso da audiência aos mesmos subsídios (ou grande parte deles) que os jornalistas acionaram na elaboração do conteúdo. Na prática, significa revelar informações sobre métodos e fontes adotadas. O que Kovach e Rosenstiel (2014) chamam de “Espírito da Transparência” tem como inspiração o método científico: explicar o que se sabe a respeito de um fato e como se sabe.

Sobre a relação entre as audiências e o jornalismo, Karlsson (2022, p. 8) enfatiza que é necessário encontrar um ponto em comum entre ambos. Dessa forma, pode ocorrer um alinhamento de expectativas entre o que o jornalismo é capaz de oferecer e o que a audiência espera receber. “Se a transparência não é traduzida em rituais performativos que são aceitos pelo público, é impossível decidir se um conteúdo jornalístico é ou não transparente e se há ou não implicações para a credibilidade percebida da perspectiva do público.”<sup>74</sup>. Entende, assim, que deve haver uma espécie de guia que oriente as práticas jornalísticas e viabilize a transparência.

Ao refletir sobre a relevância das informações às quais o público deve ter acesso, Kovach e Rosenstiel (2014) defendem que os jornalistas incluam, em seus textos, respostas para três questões: o que a audiência precisa saber para avaliar esta informação? Há algum aspecto do tratamento da informação utilizada que exija explicação? Quais são as respostas que não podem ser fornecidas no momento? Embora pareça contraintuitivo, é preciso reconhecer as incertezas, quando existem. Assentado sobre a premissa da geração

---

<sup>74</sup> No original “If transparency is not translated into performable rituals that are acceptable to the public, it would be impossible to decide whether a piece of journalism was transparent or not and whether this had any implications for perceived credibility from the perspective of the public.”

de forma (PARK, 1940) ou formas (NIELSEN, 2017) de conhecimento, o jornalismo deve capturar e comunicar inclusive a existência daquilo que não se sabe (WIHBEY, 2019).

Contribuições como as de Karlsson (2010) e Meier (2009) são fundamentais como ponto de partida para a criação de conceitos operacionais sobre transparência. Karlsson (2010), por exemplo, classifica a transparência em duas vertentes principais: *disclosure transparency*, que abrange explicações sobre como as notícias são selecionadas e produzidas, incluindo deixar claro ao público o quanto uma produção constante, sete dias por semana, pode afetar a qualidade do conteúdo; e *participatory transparency*, que envolve a participação da audiência em diferentes estágios do processo de produção das notícias e inclui elementos de interatividade. Anos mais tarde, o autor acrescentou o amplo conceito de *ambient transparency*, que busca aproximar a audiência da produção de conteúdo por meio de ferramentas e técnicas interativas que adicionem contexto ao conteúdo oferecido (KARLSSON, 2022).

Meier (2009), por seu turno, afirma que a transparência no jornalismo pode ocorrer de duas formas: externa (*outside transparency*) e interna (*self-transparency*). Enquanto a vertente externa diz respeito às tradicionais práticas jornalísticas – à medida que reporta fatos da economia, da política e da ciência, o jornalismo leva transparência à sociedade –, a interna contempla uma espécie de verificação por pares e autorregulação, em que a comunidade jornalística se encarrega de avaliar o que é produzido pelos veículos de comunicação.

A participação externa ligada à produção e verificação de conteúdo produzido pelo jornalismo tem sido particularmente valorizada pela audiência, segundo estudo conduzido por Uth, Badura e Blöbaum (2021). Ao indagar 1.029 leitores alemães sobre o que os faz confiar em um veículo, descobriram que a audiência considera importante participar ativamente da criação de conteúdo, seja sugerindo tópicos de cobertura, seja contribuindo para o processo de pesquisa ligado à reportagem. Em nível de conteúdo, práticas como o *fact-checking* também são valorizadas pelo público. “Uma vez que a transparência permite um melhor julgamento de produtos e processos jornalísticos, os destinatários estão em uma posição melhor tanto para julgar a confiança no jornalismo quanto os riscos

associados ao uso das informações veiculadas como base de suas ações”<sup>75</sup> (UTH; BADURA; BLÖBAUM, 2021, p. 77).

Na esteira da oferta de clareza ao leitor, a discussão sobre a transparência como operacionalização da objetividade resgata a possibilidade de contextualização na Web, cuja arquitetura permite o uso de hiperlinks. Anteriormente explicado nesta tese, o hipertexto, forma de escrita para a Web, possibilita camadas de informação, permite o aprofundamento e a contextualização em diferentes níveis, além de apresentar as referências originais que remetem à procedência da informação e permitem que o próprio leitor verifique de onde foram extraídas. “Aquilo em que costumávamos acreditar porque pensávamos que o autor era objetivo, acreditamos agora porque podemos ver, entre os textos do autor, as fontes e valores que o levaram àquela posição”<sup>76</sup> (WEINBERGER, 2009, s.p.). Em texto anterior (GEHRKE, 2018b), defendi a criação de três parâmetros essenciais para o cumprimento da transparência como valor no jornalismo: clareza nas fontes; abertura de pesquisas, testes e análises; e correção de erros e atualização.

Para Allen (2008), a transparência é uma forma de os jornalistas abrirem seus processos sem sair do método objetivo de reportagem. Embora os dois conceitos operem em consonância, o autor acredita que a transparência seja capaz de modificar a objetividade em diferentes formas. Enquanto a objetividade está ligada à verdade que deve ser descoberta, a transparência reconhece que o jornalismo trabalha com uma verdade verificável, provisória e complexa. Por meio da transparência, o jornalismo mostra suas provas de verdade, ou seja, apresenta evidências como documentos e outros registros que permitem ao público perceber que a informação é válida e passou por um processo de verificação.

Existem, é claro, críticas e limitações quanto à abordagem da transparência no jornalismo. Há quem perceba a abertura excessiva como um problema, visto que, historicamente, acredita-se no jornalismo como uma instituição legítima e confiável. O’Neill (2002, p. 6), por exemplo, entende que a confiança no jornalismo deveria ser suficiente. “Em algum momento, apenas temos que confiar. Não há uma resposta completa para a velha questão: ‘quem vai guardar os guardiões?’. Pelo contrário, a confiança é necessária

---

<sup>75</sup> No original “Since transparency enables better judgment of journalistic products and processes, recipients are in a better position to judge both the trustworthiness of journalism as well as the risks associated with using journalistic information conveyed by journalism as the basis of their actions”.

<sup>76</sup> No original “What we used to believe because we thought the author was objective we now believe because we can see through the author’s writings to the sources and values that brought her to that position”.

precisamente porque todas as garantias são incompletas.”<sup>77</sup> Karlsson (2022) adiciona que o estabelecimento de confiança depende da incerteza e da vulnerabilidade, admitindo que dificilmente haverá conhecimento sobre a íntegra de um processo.

Nesta mesma linha, autores como Allen (2008) e Meier (2009) citam o temor de que a autonomia e a autoridade jornalística sejam contestadas. Meier (2009) ainda levanta como possíveis problemas às redações a questão do tempo excessivo dedicado pelos jornalistas no detalhamento de suas próprias práticas e a competição por atenção que os materiais complementares ligados à transparência possam detonar, desviando os leitores dos temas prioritários, que servem de fio condutor das reportagens. Allen (2008, p. 328), por sua vez, caracteriza a transparência como um procedimento estratégico, ou uma espécie de “metodologia conveniente” – cientes de que muitas vezes realizam a cobertura de eventos pré-programados, os jornalistas estariam tentando se livrar do mundo “construído” que são designados a reportar.

Pensadores como Han (2017, p. 35) criticam de maneira mais ampla a sociedade da transparência. Geralmente abordado do ponto de vista discursivo para tratar da liberdade da informação, o tema abrange processos sociais complexos e sistêmicos, segundo o autor, em que transparência, controle e vigilância estão muito próximos. “Tudo deve tornar-se visível; o imperativo da transparência coloca em suspeita tudo o que não se submete à visibilidade.” Percepção semelhante havia sido descrita anteriormente por Marcondes Filho (2000, p. 112), para quem a transparência e o excesso de informação estão associados a uma ideia de “visibilidade total”, abertura e algo como uma “radiografia” de dados detalhados e individuais, o que é viabilizado pelas tecnologias de comunicação e informação. Neste caso, configura-se o estado de simetria ilusória do qual fala Han: a sociedade da transparência busca eliminar todas as relações assimétricas, assim como toda e qualquer lacuna de conhecimento, ou seja, o “não saber”. Em uma perspectiva crítica, Karlsson (2022) dá a entender que a postura de demonstração de transparência é cômoda à medida que, sinalizando o que foi feito, a instituição – neste caso, a jornalística – busca se eximir da responsabilidade de envolver o público em uma efetiva participação de tomada de decisões em processos.

Dentre as críticas aqui trazidas, talvez a mais significativa para refletir sobre as transformações do jornalismo seja o fato de que, por se dizer transparente e abrir seus

---

<sup>77</sup> No original “At some point we just have to trust. There is no complete answer to the old question: ‘who will guard the guardians?’. On the contrary, trust is needed precisely because all guarantees are incomplete.”

processos e métodos, o jornalismo tenta se colocar em uma posição horizontal juntamente com o leitor e outras instituições – e aí está sua força e sua fraqueza. O risco, assim, parece ser o jornalismo se equiparar a outros agentes que atuam na configuração do conteúdo noticioso – mas que não admitem seu trabalho como editores ou produtores de notícias – desde a produção até a distribuição, como os sites de rede social e outras plataformas (BELL; OWEN, 2017).

Assim como outras práticas, o desenvolvimento do jornalismo está atrelado à tecnologia. A tecnologia que gera questionamentos sobre a privacidade dos dados pessoais e o seu uso por parte do governo, por exemplo, é a mesma que permite às instituições abrirem informações de interesse público e, assim, possibilitar que o jornalismo monitore esses dados e transforme-os em conhecimento. Se uma sociedade da transparência apresenta riscos, uma sociedade opaca traz limitações sobre o conhecimento da realidade em que se vive. Por isso, cada vez mais cobra-se dos governos que dados de interesse público não estejam apenas disponíveis, mas acessíveis – o que pressupõe a disponibilização das informações em formato aberto (legível por máquina) e frequentemente atualizados.

Embora autores como Karlsson (2010) e Meier (2009) tenham discutido e sugerido diferentes matrizes para a transparência, citando principalmente âmbitos interno e externo às redações, sistematizo a transparência em dois macros conceitos, conforme dito no início desta seção: como **operadora de uma prática ética** e como **recurso discursivo de legitimação**, que serão a seguir explicados. Enquanto o primeiro tem em sua base a confiança entre o jornalista e o público e se baliza a partir das práticas operacionais da reportagem, o segundo tem relação com a credibilidade percebida pelo leitor em nível organizacional, identificando um veículo de comunicação como mais ou menos credível. Mesmo apresentando características diferentes, as categorias estão relacionadas – uma boa prática jornalística é fundamental para que se deposite confiança em uma empresa, cuja marca pode estimular e representar a existência de credibilidade. Ambas têm em seu horizonte o desenvolvimento de uma prática qualificada<sup>78</sup> e capaz de gerar conhecimento. Em outras palavras, a audiência espera receber informações verdadeiras e acuradas sobre

---

<sup>78</sup> Esta tese não se propõe a discutir o conceito de qualidade, embora avalie que o jornalismo, enquanto prática social, deva ter este conceito em seu horizonte para produzir conhecimento. A título de marco teórico, entendo que a qualidade de um produto está associada aos requisitos necessários para o seu funcionamento e ao que é esperado pela audiência. No jornalismo, os parâmetros ligados à qualidade partem do cumprimento de pilares essenciais desta prática profissional: verdade, relevância e pluralidade (GUERRA, 2008; 2010; 2014). Em se tratando de relevância e interesse público, um conteúdo de qualidade é elaborado para atingir um significativo estrato da sociedade, ainda que regionalmente, e não serve para atender aos interesses individuais (GUERRA, 2010; 2014; LACY; ROSENSTIEL, 2015).

fatos de interesse público, abordados sob múltiplas perspectivas e que façam sentido à vida em sociedade.

A transparência, por fim, não pode ser apenas uma estratégia adotada pelos jornais para aproximar-se da audiência. Uma vez que iniciativas de transparência são criadas (a exemplo de blogs que tratam dos bastidores da notícia, quadros informativos e contextuais nas reportagens, código aberto etc.), é preciso manter a funcionalidade desses elementos. Não basta desenvolver, é necessário manter, atualizar (UTH; BADURA; BLÖBAUM, 2021). Para isso, os veículos jornalísticos precisam, sobretudo, compreender a sua audiência e fornecer soluções correspondentes para que haja uma relação de honestidade e uma atmosfera de menor desconfiança por parte da população. Para tanto, as duas categorias discutidas nesta tese, que serão mostradas a seguir, precisam funcionar de maneira concomitante e interligada.

#### **4.3.1 Operadora de uma prática ética**

A transparência emerge no jornalismo contemporâneo como um princípio ético (PLAISANCE, 2007) a ser cumprido pelos jornalistas, ou seja, como parte de sua deontologia profissional. Em termos operacionais, integra a lista de fundamentos básicos com o qual o jornalista deve se comprometer, somado à honestidade, independência, verificação e responsividade para corrigir erros (RUSBRIDGER, 2020).

Segundo Bucci (2000, p. 25), a ética jornalística representa, acima de tudo, “[...] um pacto de confiança entre a instituição do jornalismo e o público, num ambiente em que as instituições democráticas sejam sólidas”. Por isso, ao mesmo tempo em que a ética jornalística é uma exigência ligada à prática profissional, é também um resultado de uma sociedade democrática balizada pela liberdade de imprensa. Na percepção de O’Neill (2002), a transparência funciona somente quando há uma predisposição do público em confiar em determinada instituição; neste caso, o jornalismo.

Como dissemos anteriormente, o jornalismo é norteado pela verdade provisória e verificável dos fatos, podendo ser acessada e aferida por meio de métodos de trabalho. Por isso, a empresa jornalística não é apenas mais uma companhia que produz e comercializa um tipo de produto – neste caso, a informação. A imprensa, de acordo com Bucci (2000), existe porque há uma relação de confiança entre as partes, vínculo fundamental para a percepção de credibilidade por parte do público. O autor entende, então, que a

transparência é condição fundamental para que as empresas jornalísticas lidem com as suas limitações. Nesta mesma linha, Wihbey (2019) afirma que a ligação entre a ética jornalística e a transparência implica os profissionais admitirem que seu trabalho é um produto inacabado, um processo em curso. Assim, os jornalistas estariam aptos a comunicar à audiência aspectos de incerteza e possíveis falhas.

Embora não seja integralmente possível dissociar a prática individual do jornalista da conduta coletiva esperada das redações, há algumas diferenças no tratamento da informação. Cornu (1998) distingue as responsabilidades dos jornalistas e dos veículos de comunicação. Segundo o autor, o profissional é pessoalmente responsável para com suas fontes e o público, para com as normas profissionais e para com as próprias convicções. Junto aos veículos de comunicação onde trabalha, porém, existe a responsabilidade compartilhada sobre a informação difundida e comunicada aos grupos de interesse, ao Estado e frente à sociedade como um todo.

Quanto à postura do jornalista, pode-se refletir sobre suas relações com as fontes pessoais e questões de privacidade (CHRISTOFOLETTI, 2019). Segundo Bucci (2000, p. 12), a ética no jornalismo pressupõe premissas institucionais, mas lida com o campo das decisões individuais dos profissionais. Ao mesmo tempo, reforça o autor, compreende valores que só fazem sentido quando seguidos por empregados e empregadores, e de modo que possa ter o público como fiscalizador. Para tanto, a ética jornalística tem como premissa fundamental a liberdade de imprensa, princípio inegociável “para beneficiar a sociedade democrática em sua dimensão civil e pública”.

A transparência está no centro na versão de 2014 do Código de Ética da Sociedade de Jornalistas Profissionais dos Estados Unidos. Uma seção do código é dedicada exclusivamente a essa questão. Segundo o texto, um jornalismo ético pressupõe assumir responsabilidade pelos atos e explicar ao público as decisões tomadas. Assim, do ponto de vista normativo, o código indica que os profissionais devem: explicar escolhas éticas à audiência, encorajando o debate civil; responder rapidamente às questões sobre acurácia e clareza; identificar e corrigir erros imediatamente, deixando claras as correções feitas; e expor eventuais condutas que falem com a ética no jornalismo (SOCIETY OF PROFESSIONAL JOURNALISTS, 2014). No Código de Ética dos Jornalistas Brasileiros<sup>79</sup>,

---

<sup>79</sup> Disponível em: <https://fenaj.org.br/codigo-de-etica-dos-jornalistas-brasileiros/>

disponibilizado pela Federação Nacional dos Jornalistas (Fenaj) e atualizado em 4 de agosto de 2007, não há menção à transparência.

Alguns manuais de redação e conduta no jornalismo estão desatualizados e não incorporam o ambiente da Web como um espaço de publicações. Em tese, este espaço, por ser aberto, permite facilmente a estruturação de informações em camadas – por meio do hipertexto, como apontamos anteriormente – e a rápida correção de erros. Após analisar as políticas de erros empregadas por cinco jornais digitais, Vieira (2014) sugeriu a necessidade de um código de ética específico para esse tipo de mídia no Brasil, bem como a atualização de manuais de redação, visto que os textos disponíveis não consideravam as especificidades do jornalismo praticado e publicado na Web. De maneira geral, os resultados de sua investigação indicaram falta de transparência, retificação não visível, revisão deficiente, pressa, perpetuação do erro, notícia original sem menção ao erro e falta de interação com o leitor. E descobriu que, quando a correção de erros ocorre de forma transparente, valores como confiança e credibilidade são reforçados entre leitor e veículo jornalístico (VIEIRA, 2014).

Do ponto de vista dos jornalistas, não há dúvidas de que a transparência surge como valor fundamental (CHRISTOFOLETTI, 2021; GEHRKE, 2020). Estudo de Toff *et al.* (2021) que reúne expectativas dos jornalistas e dos leitores sobre como o jornalismo deve ser conduzido conclui que a transparência é um valor bastante acionado pelos profissionais. No entanto, esta palavra-chave não chega a ser citada pela audiência – o público prioriza a familiaridade com a marca e a reputação do veículo de comunicação, valorizando, inclusive, a forma como uma notícia é apresentada. Neste estudo, grupos focais foram realizados com pessoas em diferentes países – inclusive no Brasil –, mas há de se destacar que o conceito de “transparência” pode simplesmente não fazer parte do vocabulário dos leitores, que acionam em seu discurso valores como isenção e equilíbrio.

Mesmo que pareça um contrassenso, admitir o erro é um dos pontos mais valorizados pelo público e levados em conta ao refletir sobre os veículos de comunicação cujas notícias irão consumir. Karlsson e Clerwall (2018) descobriram que, embora os leitores critiquem o equívoco jornalístico – que em princípio não deveria ocorrer, sendo evitado pelo processo de apuração –, valorizam a honestidade dos jornalistas e dos veículos ao admitirem que erraram. Assim, é possível pensar que há a construção de uma relação de confiança entre a audiência e os veículos de comunicação; como qualquer relação interpessoal, a confiança se apresenta como um elemento fundamental. “Mesmo que a



transparência não seja sempre uma condição suficiente para um comportamento mais ético, sua ausência é pré-requisito para o engano [...]”<sup>80</sup> (PLAISANCE, 2007, p. 193). Se a transparência não é um antídoto completo para falhas, sua ausência torna-se um risco.

Visto que o jornalismo tem como propósito informar a sociedade, é a este público que deve se mostrar honesto, vinculado a princípios éticos, e prestar contas. “Se um fato foi relatado de modo parcial ou inexato, se um artigo ou uma emissão divulgou alguma informação contendo erros materiais, os fatos devem ser precisados e os erros, corrigidos” (CORNU, 1998, p. 68-69). Uma conduta ética, ainda que pareça centralizada na prática da reportagem, deve se estender à tomada de decisões editoriais que norteiam toda a conduta de uma empresa jornalística. Isso envolve, por exemplo, deixar claro quem financia o jornal, visto que questões financeiras costumam gerar questionamentos quanto a possíveis conflitos de interesse.

#### 4.3.2 Recurso discursivo de legitimação

Em nível organizacional, a transparência pode servir de estratégia de legitimação para manutenção da credibilidade percebida<sup>81</sup> (LISBOA, 2012) pelos leitores. Em outras palavras, um veículo de comunicação abre seus métodos, processos e questões editoriais não só porque é eticamente recomendável, como discutido no tópico anterior, mas também para *parecer* transparente. A transparência editorial está associada à confiabilidade das informações produzidas e divulgadas pelos veículos de comunicação (GUERRA, 2014) e depende da relação de confiança entre o público e os jornalistas.

Numa época em que as informações se proliferam no ambiente da internet, em que todos se arvoram ao status de produtores de informação, a transparência em relação a quem se credencia a esse serviço e da forma como o faz pode ser uma atitude decisiva para se distinguir aqueles dignos de credibilidade dos demais (GUERRA, 2014, p. 207-208).

---

<sup>80</sup> No original “Even if transparency is not always a sufficient condition for more ethical behavior, its absence is a prerequisite for deception [...]”

<sup>81</sup> Em sua investigação de mestrado, Lisboa (2012) identificou cinco valores principais que usualmente são associados pelo leitor a um jornalismo confiável: independência, imparcialidade, honestidade, objetividade e coerência.

Do ponto de vista discursivo, sabe-se que a transparência nunca estará completa. O jornalista escolhe o que deseja comunicar e decide o quanto de seus métodos será conhecido pela audiência. Além disso, não existe literalidade no discurso, e os sentidos sempre serão produzidos a partir de um sujeito que interpreta. O jornalismo, enquanto um gênero discursivo particular, produz narrativas em condições bastante específicas. A partir de seus métodos de apuração, do cruzamento das fontes, da pluralidade das versões e da defesa da objetividade, adota estratégias fundamentais para construir sua própria narração diante do público (BENETTI, 2008). De qualquer forma, ao aparecer no discurso de jornalistas, a transparência também afeta suas práticas (ALLEN, 2008).

Quando escolhe não abrir ao público suas decisões, seus métodos de apuração e eventuais limites que possam ocorrer nas produções, o jornalismo busca manter seu papel de autoridade e assegurar a legitimidade necessária para seguir divulgando apenas aquilo que lhe convém (BENETTI, 2008). A narrativa construída visa, assim, promover um efeito de verdade e assegurar a credibilidade de quem enuncia (neste caso, o próprio jornalismo). Desta forma, a transparência também pode ser pensada como estratégia discursiva de legitimação associada a um *efeito* – mesmo que o público não reproduza os métodos detalhadamente apresentados nas notícias, o fato de ter esta opção disponível confere um efeito de transparência (LISBOA, FRANZONI, 2018). E, ainda que seja difícil medir ganhos reais de credibilidade, a demonstração da transparência pode servir para que o jornalismo simplesmente mantenha o *status quo* de mediador qualificado da realidade sobre temas de interesse público, evitando mudanças drásticas no modo como o público o percebe (KARLSSON, 2022).

Quando discutem o jornalismo como uma crença verdadeira justificada, Lisboa e Benetti (2015) entendem que o efeito de transparência se enquadra no âmbito da justificação. Assim, o relato jornalístico apresenta pistas e provas de que diz a verdade, sustentando a credibilidade do discurso. Neste caso, como diz Weinberger (2009, s.p.), as evidências podem ser ainda mais facilmente acessadas pelo público, tornando-se algo palpável: “[...] a transparência espelha a natureza da Web. O conteúdo é conectado por meio de hiperlinks, publicado, discutido e sempre sujeito à disputa e à revisão”.<sup>82</sup>

Para Guerra (2014), a credibilidade da organização jornalística provém de uma série de informações que atestam sua idoneidade (ausência de conflitos de interesse, por

---

<sup>82</sup> No original “[...] transparency mirrors the nature of the Web: Content is linked, public, discussed and always subject to dispute and revision.”

exemplo) e dados que garantem o cumprimento dos princípios editoriais no produto jornalístico. Assim, a transparência como recurso discursivo de legitimação inclui a demonstração de que as notícias observaram métodos rigorosos de apuração, cumprindo as premissas de verdade, relevância e pluralidade, que acompanham o jornalismo. Além disso, a promoção da transparência faz com que as organizações jornalísticas se credenciem como mediadoras relevantes no ambiente democrático. Como diz Guerra (2014, p. 208): “Quanto menor a transparência, maior o risco de perda da credibilidade”.

O significado de *accountability* está ligado à transparência como recurso de legitimação de suas práticas à medida que esse conceito remete à prestação de contas e responsabilidade, valores frequentemente relacionados às práticas de órgãos públicos. Para Fengler (2019), *accountability* é um conceito mais amplo do que transparência, pois envolve tornar pública a informação sobre processos editoriais e atores envolvidos. É nesta perspectiva, principalmente a partir dos anos de 1970, que empresas jornalísticas operam com o já mencionado cargo de *ombudsman* – sujeito que faz o papel de advogado do leitor. No jornalismo digital, fala-se, ainda, em *algorithm accountability*<sup>83</sup> (DIAKOPOULOS, 2019) para discutir não só os processos de análise de dados propriamente ditos, mas o modo como os dados são produzidos e estruturados.

Nos estudos de jornalismo, Porlezza e Splendore (2016, p. 199) citam códigos de ética, produção colaborativa de notícias e ferramentas que permitem a participação de leitores como instrumentos de *accountability*. As plataformas digitais, como as redes sociais, também são vistas pelos jornalistas como uma forma de obter aproximação com a audiência (FENGLER, 2019). Para que a *accountability* faça sentido e tenha instrumentos aplicáveis, as instituições precisam ter autonomia e estar em pleno funcionamento, a começar pela democracia e pela decorrente cultura de participação das pessoas.

Para citar alguns exemplos, empresas jornalísticas criam produtos em que buscam prestar contas à sociedade, especialmente quando o leitor financia o jornalismo. Em alguns países europeus, uma estratégia adotada é divulgar relatórios editoriais anuais. Trata-se de uma espécie de balanço geral sobre o que deu certo e o que não funcionou para determinada redação durante o período de um ano, incluindo erros de escrita, decisões éticas e, em casos positivos, prêmios recebidos pela equipe de jornalismo. “Relatórios

---

<sup>83</sup> Diakopoulos (2019, p. 11) chama de “*algorithm accountability reporting*” o desenvolvimento de métodos para auditar e explicar a formação de sistemas e bancos de dados diante dos processos de erros, discriminação e censura que podem ser inerentes ao processo da reportagem.

editoriais servem para legitimar o jornalismo sob novas circunstâncias que demandam tal justificação”<sup>84</sup> (EIDE, 2014, p. 683). Ao tornar pública essa espécie de autoavaliação, o jornalismo pede um voto de confiança e defende-se como credível.

Iniciativa responsável por investigações em fontes abertas, o *Bellingcat*<sup>85</sup> divulgou um relatório com o balanço anual<sup>86</sup> de suas práticas jornalísticas e suas receitas referente ao ano de 2020. Visto que o público é uma importante fonte de financiamento para este veículo, existe um compromisso de dar algum tipo de retorno. Uma vez que os veículos deixam claras as suas premissas de atuação, o leitor está apto a compreender se existe algum tipo de conflito de interesses ou origem duvidosa de financiamento. Quanto mais informações tiver, mais o leitor terá condições de analisar, por conta própria, o que o veículo tem feito com os recursos recebidos. Mesmo que práticas de *accountability* estejam associadas um *efeito de transparência*, tal efeito pode gerar implicações sobre o modo como a audiência deposita confiança em um veículo jornalístico.

Pesquisas como a de Johnson e St. John (2021) e Moran (2020) têm demonstrado que não só os jornais são os protagonistas da credibilidade percebida; os jornalistas responsáveis pelas notícias também são importantes nesse processo. Importa saber, por exemplo, quem está por detrás das histórias, sua expertise e até mesmo possíveis vieses.

Na pesquisa que busca associar graus de transparência à credibilidade percebida pelos leitores em relação ao veículo, Johnson e St. John (2021) examinam as percepções de 885 leitores dos Estados Unidos. O público foi distribuído em oito grupos diferentes para analisar que tipo de informação é útil para ampliar a percepção de credibilidade em relação ao veículo. Em um dos grupos, por exemplo, a audiência analisou apenas a notícia; em outro, a notícia, uma foto e inúmeras informações do repórter; em outro, ainda, um quadro explicativo sobre o tema da reportagem foi adicionado, e assim por diante.

Os autores descobriram que, além da notícia em si e de um quadro explicativo sobre por que a notícia virou pauta, veicular uma foto do jornalista que escreveu a notícia, bem como alguns detalhes sobre ele (educação, *hobbies* e experiência profissional), pode ampliar a percepção de credibilidade sobre o jornalista, a notícia e a organização para qual o jornalista trabalha. Ainda que, individualmente, cada um desses aspectos seja capaz

---

<sup>84</sup> No original “Editorial reports serve to legitimate journalism under new circumstances where the demand for such justification is required”.

<sup>85</sup> Baseado na Holanda, o *Bellingcat* se autodenomina um coletivo independente formado por pesquisadores, investigadores e jornalistas cidadãos que utilizam fontes abertas e redes sociais para investigar diversos temas.

<sup>86</sup> Disponível em: <https://www.bellingcat.com/app/uploads/2021/05/Bellingcat-Annual-Report-2020.pdf>

de agregar valor para o todo, é na combinação deles que se consegue o melhor resultado. Esse resultado dá pistas de como a transparência pode ser benéfica: “A chave parece ser encontrar um equilíbrio entre a abertura de informações sobre o jornalista e combiná-la com a informação sobre por que uma notícia entrou na cobertura jornalística” (JOHNSON; ST JOHN, 2021, P. 967).

Moran (2020), por sua vez, examinou o uso da plataforma *Slack*<sup>87</sup> – que funciona como uma espécie de fórum e bate-papo – como espaço de fornecimento de informações de bastidores do processo de produção da notícia (*newsmaking*). A pesquisadora analisou o *Rantt Media*, autointitulado veículo independente que busca combater a desinformação e o extremismo, localizado em Washington D.C, nos Estados Unidos. Por meio de entrevistas com jornalistas e leitores e acompanhando as publicações no Slack por cinco meses, verificou que os leitores utilizam a plataforma por dois motivos principais: a) como um agregador de notícias confiáveis; b) como espaço para aproximar-se da redação e compreender quem são os jornalistas que nela trabalham e quais são suas perspectivas.

Entre os leitores entrevistados, porém, não foi detectada uma alta demanda por transparência. Um dos motivos para isso pode ser a existência prévia de uma relação próxima e de confiança dos leitores com o *Rantt Media*. De todo modo, a autora identifica a abertura de espaço no Slack como um ato significativo de transparência – ou efeito de transparência, como eu definiria este caso –, visto que tal atitude demonstra aos leitores e potenciais leitores que o conteúdo está aberto ao escrutínio público.

Em outras palavras, significa dizer que a transparência como recurso discursivo de legitimação é um convite para que o leitor conheça os processos práticos e editoriais à medida que sugere a publicação de um conteúdo qualificado, satisfatório e credível. Associado ao conhecimento científico pela via da objetividade no método, o Jornalismo Guiado por Dados tem a transparência como tópico importante e diferencial de sua estrutura.

---

<sup>87</sup> *Slack* é uma plataforma de comunicação empresarial desenvolvida pela empresa *Slack Technologies*. Oferece a criação de canais (semelhantes a salas de bate-papo) por tópicos de interesse, grupos privados e mensagens diretas.

### 4.3.3 A transparência no Jornalismo Guiado por Dados

Para executar o método inspirado na ciência e, conseqüentemente, gerar um tipo de conhecimento especializado (TRÄSEL; GEHRKE, 2021), o Jornalismo Guiado por Dados deve submeter o seu trabalho ao escrutínio público, oferecendo à audiência informações sobre suas práticas (operadora de uma prática ética) e princípios editoriais (recurso discursivo de legitimação). A transparência, assim, funciona como reflexo da verdade e da objetividade.

Por operar principalmente na Web, o JGD herda uma tradição desenvolvida desde a época do seu surgimento, nos anos 2000: a transparência por meio do uso de hiperlinks. À época, autores de blogs utilizavam este recurso de abertura para assegurar a credibilidade de suas publicações (MEIER, 2009), assim como encaravam o próprio trabalho como uma construção, passível de atualizações frequentes e, por isso, correspondentes a uma verdade provisória (PHILLIPS, 2010). Oferecer a possibilidade de a audiência refazer o passo a passo adotado significa admitir que o público tem direito a acessar as mesmas informações que os jornalistas. “Enquanto os repórteres têm de confiar em conhecimento de segunda e terceira mão, o público está ainda menos equipado para avaliar a veracidade das notícias”<sup>88</sup> (KOVACH; ROSENSTIEL, 2011, p. 33).

Criador do Jornalismo de Precisão, principal precursor do JGD, Meyer (2002) entende que jornalistas e cientistas apresentam premissas parecidas na realização de seu trabalho. Ambas as categorias de profissionais, conforme o autor, adotam em seu trabalho: o ceticismo, ao questionar as informações recebidas; a abertura, ao explicar os procedimentos de investigação a partir de documentos; o instinto para operacionalização, visto que buscam testar uma hipótese no mundo real; o senso de obtenção da verdade, levando em conta que a verdade é provisória; e parcimônia, uma vez que permanecem atentos para escolher teorias que de fato tenham potencial para explicar um fenômeno.

Por trabalhar com planilhas, modelos matemáticos, análise estatística, formulações de *script* e outros – tudo isso sob a ética *hacker* de compartilhamento –, é esperado que os jornalistas troquem informações entre si e sejam capazes de tirar dúvidas e avançar no conhecimento enquanto comunidade. A rotina diária das redações, porém, parece

---

<sup>88</sup> No original “While reporters have to rely on second and third-hand knowledge, the public itself is even more poorly equipped to assess the truthfulness of news”.

gerar um impacto negativo sobre o potencial do Jornalismo Guiado por Dados em relação às políticas de transparência adotadas nas notícias.

Um dos poucos estudos que examina a transparência no JGD é o de Rodrigo Zamith (2019). Ao analisar cerca de 150 artigos dos jornais *The New York Times* e *The Washington Post* na primeira metade de 2017 a partir dos conceitos de transparência, interatividade, diversidade e proveniência da informação, Zamith (2019) conclui que o JGD praticado no dia a dia não é aberto e se mostra altamente dependente de fontes oficiais. A investigação mostrou, ainda, que são raros os casos em que os próprios jornalistas de dados efetuam a coleta das informações. Portanto, o autor entende que o JGD ainda tem de ser aprimorado, caracterizando-o como uma prática em evolução, mas não revolucionária.

Os resultados da pesquisa de Zamith convergem para a crítica feita por Träsel (2014): ao omitir os modelos estatísticos e bancos de dados originais, os jornalistas de dados e seus respectivos veículos deixam de incentivar o caráter público de auditoria que a transparência procura permitir. O autor exemplifica que, em se tratando da ciência – provavelmente a principal referência do JGD –, os artigos são revisados por pares e submetidos à avaliação metodológica de outrem, validando, ou não, as descobertas realizadas. “No caso do JGD, o acesso às bases de dados originais permitiria à audiência e a outros repórteres realizar suas próprias análises e julgar a validade das conclusões deles derivadas por uma determinada peça jornalística” (TRÄSEL, 2014, p. 128).

É claro que, diante de questões políticas e de saúde pública, como a própria pandemia de Covid-19, aumenta a pressão sobre a abertura de dados públicos e métodos utilizados nas análises. Em estudo sobre a cobertura do JGD da pandemia, Brenol e Träsel (2020) analisaram reportagens com técnicas de JGD e detectaram, entre os resultados, a falta de disponibilidade de informações consolidadas sobre a Covid-19, apontando a ainda presente dificuldade de obter acesso às informações públicas no Brasil. Dessa forma, segundo os autores, os jornalistas não só expuseram tais limitações, mas também adotaram um tom crítico quanto ao acesso aos dados. Há um temor de que esse cenário não seja exclusivo da pandemia de Covid-19, mas indique maior opacidade no Estado brasileiro, apesar dos avanços na transparência pública digital nas últimas décadas.

Observa-se que, nos últimos anos, surgiram iniciativas jornalísticas ancoradas no princípio da transparência. Os próprios jornalistas, como veremos a seguir, apresentam esta característica marcadamente em seu *ethos* profissional. Na esteira da colaboração e

da ética *hacker* diagnosticada por Träsel (2014) no Brasil, a transparência se soma como um valor acionado pela comunidade de jornalistas de dados.

No Brasil, a iniciativa *Vortex Media*, inaugurada em outubro de 2019 por Diego Escosteguy, surgiu para oferecer o que dizia ser um “jornalismo em movimento” e sob o discurso da “transparência radical de métodos e intenções” (GEHRKE; BENETTI, 2020), mas nunca relevou suas formas de financiamento e encerrou parcialmente as atividades dois meses após o início. Dessa forma, a equipe de jornalistas de dados, um dos grandes diferenciais do novo empreendimento, teve apenas uma curta trajetória.

Em 6 de janeiro de 2020, após o encerramento das atividades da equipe de dados do *Vortex*, o ex-editor desse pequeno grupo, jornalista Sérgio Spagnuolo, criou juntamente com Alexandre Orrico o *Núcleo Jornalismo*. A missão do veículo, segundo sua própria página<sup>89</sup>, é colocar a transparência no centro de tudo para estimular o uso “[...] principalmente de dados públicos, a partir de investigações, análises e aplicações que informarão e jogarão luz sobre políticas públicas, instituições e acontecimentos políticos e sociais” (NÚCLEO JORNALISMO, 2021, s.p.).

Neste cenário, o *Núcleo* busca executar, internamente, a transparência que cobra do poder público. “[...] a transparência que queremos das instituições também é a que devemos aplicar internamente. Isso significa transparência corporativa (quem nos financia e quem somos) e de processos (como fazemos nosso jornalismo)” (NÚCLEO JORNALISMO, 2021, s.p.). Para fazer isso, a empresa jornalística declara apostar no referenciamento sistemático de fontes, incluindo descrição e links; transparência de políticas públicas; e explicação sobre a importância de um tema.

Nos Estados Unidos, o nativo digital *The Markup*<sup>90</sup>, autointitulado uma instituição sem fins lucrativos, criou sua própria metodologia para cobrir a área de tecnologia e como as decisões dos grandes players (*Facebook* e *Google*, por exemplo) afetam a vida em sociedade. O *Markup Method*, segundo a editora Júlia Angwin (2021), consiste na publicação detalhada da metodologia empregada nas reportagens. São três passos principais em direção ao método científico: a elaboração de dados inéditos, a revisão por parte de especialistas e a comunicação dos dados.

A despeito dessas iniciativas, as redações jornalísticas convencionais ainda não estão completamente envoltas sob a aura da transparência, embora haja um movimento de

---

<sup>89</sup> Disponível em: <https://nucleo.jor.br/sobre>

<sup>90</sup> Disponível em: <https://themarkup.org/about>



mudança gradual nessa direção. Devido aos interesses comerciais, por exemplo, alguns veículos de comunicação optam por não abrir seus procedimentos. Essa premissa vai ao encontro do que mostra pesquisa realizada com jornalistas de dados brasileiros em 2019 (GEHRKE, 2020). O levantamento sugere que a transparência está mais conectada ao *ethos* do jornalista do que com as regras adotadas nas redações. Quase três quartos (72,2%) dos respondentes concordam com a afirmação de que “Alguns veículos de comunicação evitam ser transparentes (oferecer o código-fonte, por exemplo) devido às questões de competição”.

Ao entrevistar 17 jornalistas de 15 veículos de comunicação na Alemanha, Koliska e Chadha (2018) descobriram que, para a maior parte dos entrevistados, a transparência está relacionada à divulgação de informações de bastidores ao público, assim como também está ligada a valores como honestidade (inclusive sobre possíveis vieses dos jornalistas na cobertura das notícias). Ainda que os jornalistas considerem que uma maior transparência possa alavancar a credibilidade no jornalismo, afirmam que o tema é raramente discutido nas redações. Um tópico pouco debatido é, por consequência, pouco implementado.

Entre os jornalistas de dados, estudos de Gehrke (2020), com profissionais brasileiros, e de Beiler, Irmer e Breda (2020), com alemães, mostram que os praticantes entendem que a abertura do método é uma boa alternativa para gerar confiança e qualidade. Enquanto os brasileiros enxergam que a transparência no Jornalismo Guiado por Dados auxilia a estabelecer uma relação de respeito com a audiência e é uma forma de combater a desinformação por meio de sua clareza metodológica, os alemães entendem que o JGD produz práticas mais transparentes do que o jornalismo convencional.

A pesquisa realizada com 36 jornalistas de dados<sup>91</sup> no Brasil (GEHRKE, 2020) mostra que, entre os profissionais, há alto grau de confiança na capacidade de o JGD ser uma prática aberta, capaz de demonstrar métodos e procedimentos. Mais de três quartos (80,6%) dos participantes concordaram com a afirmação de que “Com raízes científicas, o jornalismo de dados é uma das formas mais eficientes de apresentar conteúdo transparente, uma vez que seus métodos permitem a reprodução”, enquanto 19,4% se

---

<sup>91</sup> A pesquisa envolveu a aplicação de um questionário a jornalistas de dados brasileiros. O convite para participação foi enviado a aproximadamente 200 profissionais, mas apenas 36 retornaram. Desses, metade (50%) afirma trabalhar em redações, enquanto um quarto (25%) diz ser empreendedor e 8,3% freelancer. Os outros 16,7% correspondem a trainees, servidores públicos, coordenadores de projetos e atividades similares. Em relação ao período de trabalho, 38,9% são jornalistas há mais de dez anos, 33,3% de seis a dez anos e 27,8% possuem até cinco anos de profissão (GEHRKE, 2020).

declararam neutros em relação a essa afirmação. Todavia, ninguém discordou. Além disso, muitos participantes (91,7%) indicaram que a transparência deveria ser demonstrada no dia a dia, em todos os textos, enquanto 8,3% acreditam que isso deva acontecer somente em reportagens especiais.

Para os profissionais brasileiros consultados no levantamento, a transparência é considerada um valor importante no jornalismo porque<sup>92</sup>: contribui para elevar a credibilidade (91,7%); mostra as evidências de um fato (86,1%); auxilia no combate à desinformação (80,6%); aumenta a qualidade do conteúdo (80,6%); ajuda a criar uma relação de respeito com o leitor (75%); mostra ao leitor qual a fonte acionada (52,8%); e envolve a abertura do código-fonte utilizado (47,2%).

Em entrevista a Gehrke e Mielniczuk (2017, p. 10), Meyer identifica a clareza no uso de fontes jornalísticas como uma das principais virtudes para a replicabilidade metodológica: “O melhor jornalismo descreve seus métodos com tal clareza e precisão que o leitor é convencido de que poderia seguir os mesmos procedimentos e alcançar as mesmas conclusões”. Phillips (2010) vai além do básico: para a autora, o ideal seria os veículos de comunicação divulgarem os metadados – os dados dos dados – para organizar as informações, mostrar as referências originais e estruturá-las de modo que os leitores possam compreendê-las em sua totalidade. Quando cumpre os requisitos de transparência, em especial quanto à abertura de métodos e fontes, o jornalismo aproxima-se da ciência. Neste ponto, a palavra-chave é **replicabilidade**, cuja centralidade será demonstrada na proposição final desta tese.

---

<sup>92</sup> Nesta questão, uma lista foi apresentada aos participantes, que poderiam assinalar mais de um item sobre por que consideram a transparência importante.

## 5 ELEMENTOS DE TRANSPARÊNCIA: PERCURSO METODOLÓGICO E RESULTADOS

Monitorada pela imprensa no mundo todo, a pandemia de Covid-19 exigiu dos jornalistas uma cobertura explicativa e especializada. Equipes de Jornalismo Guiado por Dados foram acionadas, em um primeiro momento, para contabilizar e contextualizar o número de casos confirmados e mortes por Covid-19. A falta de testes e dados confiáveis mostraram-se um problema em diferentes países, desafiando jornalistas a trabalhar com as informações disponíveis e admitir as limitações das estatísticas oficiais. Principalmente no segundo ano da pandemia, os números sobre a vacinação também passaram a ser consultados, interpondo desafios inerentes à qualidade dos dados públicos<sup>93</sup> e à necessidade de conhecimento científico para compreender e explicar a eficácia das vacinas.

No Brasil, o governo federal tornou-se o maior empecilho para uma cobertura jornalística qualificada, especialmente no início da pandemia. À época, o Ministério da Saúde apresentou um painel<sup>94</sup> limitado, utilizando-se apenas de dados agregados e desconsiderando o avanço da doença em municípios. Também ignorou a necessidade de registrar a série histórica dos casos e mortes. Esse tipo de registro longitudinal é primordial para compreender o avanço da Covid-19 em diferentes partes do Brasil e adotar estratégias de combate à propagação do vírus.

No início de junho de 2020, o presidente Jair Bolsonaro foi o responsável por promover uma espécie de “apagão” de dados, ordenando que o Ministério da Saúde atrasasse o horário de divulgação das informações – o que prejudicou a imprensa e a compilação de números para programas noticiosos – e modificasse o painel de Covid-19, limitando o acesso do público aos dados diários, sem informar o total acumulado de casos e mortes. Em resposta, veículos jornalísticos criaram uma força-tarefa<sup>95</sup> para atualizar casos e mortes diariamente, a partir de números oriundos dos municípios e de estados. O consórcio é formado por *O Estado de S. Paulo*, *Extra*, *Folha de S.Paulo*, *O Globo*, *G1* e *UOL* e, embora tenha sido criado para dar transparência aos números da pandemia, sua organização é apenas interna, uma vez que os dados não são abertos ao público. No início de novembro

---

<sup>93</sup> No Brasil, o sistema de saúde é tripartite e, por isso, as atualizações sobre o número de casos e mortes e sobre a vacinação exige uma mobilização do Poder Executivo em suas diferentes esferas. As secretarias municipais de saúde prestam contas às secretarias estaduais que, por sua vez, comunicam informações ao Ministério da Saúde.

<sup>94</sup> Disponível em: <https://covid.saude.gov.br/>

<sup>95</sup> Disponível em: <https://www1.folha.uol.com.br/equilibrioesaude/2020/06/veiculos-de-comunicacao-formam-parceria-para-dar-transparencia-a-dados-de-covid-19.shtml>

de 2020, um novo problema técnico foi atribuído ao sistema do Ministério da Saúde, fazendo com que notificações advindas dos estados brasileiros sobre novos casos e mortes sofresse atraso e, mais uma vez, prejudicasse a contagem relacionada ao avanço da doença. Os problemas diminuíram, mas não cessaram no segundo ano da pandemia: em diversas ocasiões, informações sobre a pandemia saíram do ar no site do ministério.

O cenário complexo que se apresenta no Brasil acerca da disponibilidade e do acesso aos dados públicos de Covid-19, e o quanto isso interfere na cobertura jornalística, motivou a delimitação do corpus da tese. Por essa razão, a proposta é estudar textos jornalísticos sobre Covid-19, que tenham empregado técnicas de Jornalismo Guiado por Dados e que tenham sido publicados em jornais de referência e nativos digitais. Integram o corpus apenas materiais de jornalismo informativo, ou seja, notícias e reportagens, além de projetos jornalísticos feitos com o uso de dados e elaborados exclusivamente para a cobertura de Covid-19.

Esta investigação de caráter qualitativo apresenta uma série de procedimentos metodológicos para reunir e analisar os elementos de transparência empregados por jornais brasileiros durante a cobertura da Covid-19 em 2020 e 2021. Visto que a transparência no Jornalismo Guiado por Dados é o objeto teórico de investigação, parte-se da pesquisa bibliográfica para estabelecer o estado da arte neste campo e identificar suas lacunas. Por isso, esta é uma pesquisa de caráter exploratório, cuja principal finalidade é proporcionar um panorama em relação a um tópico ainda pouco estudado. Trata-se de uma pesquisa que sistematiza conhecimento e poderá servir de base para novos estudos, com novas hipóteses. Diz Gil (2019b, p. 26) sobre a pesquisa exploratória: “O produto final deste processo passa a ser um problema mais esclarecido, passível de investigação mediante procedimentos mais sistematizados”.

Concomitante à revisão de literatura presente antes e durante o processo de escrita da tese, realizei observação sistemática de produtos jornalísticos informativos para a composição de uma amostragem por intencionalidade (GIL, 2019b). Nesta tese, utilizamos como metodologia principal o estudo de casos múltiplos (STAKE, 2006) à medida que a transparência se apresenta como objeto central de estudo. A seleção dos casos, assim, é formada por produtos jornalísticos informativos que apresentam elementos de transparência como característica comum.

Embora a discussão sobre transparência ocorra há cerca de duas décadas nos estudos de jornalismo, ainda é bastante recente no Jornalismo Guiado por Dados. Dessa

forma, esta pesquisa inspira-se em esforços metodológicos realizados anteriormente no campo do jornalismo digital, como a tese de Barbosa (2007). De caráter inovador, a autora optou por selecionar casos representativos – o estudo de caso como ilustração (MACHADO; PALACIOS, 2007) – para fundamentar as discussões conceituais do Paradigma Jornalismo Digital em Base de Dados (JDBD).

Nesta investigação, o estudo de casos múltiplos permite descrever características de transparência que são inerentes às notícias e reportagens desenvolvidas a partir de técnicas do Jornalismo Guiado por Dados. Para exemplificar cada elemento descrito, utilizamos o estudo de caso como ilustração – na próxima seção, iremos detalhar essas escolhas metodológicas.

Após a sistematização e descrição dos elementos de transparência, formulo as seguintes propostas: o **Protocolo de Transparência no Jornalismo Guiado por Dados** para analisar produtos jornalísticos já publicados, cuja estrutura é inspirada no Protocolo de Análise de Cobertura Jornalística (SILVA; MAIA, 2011), e o **Método para o Jornalismo Guiado por Dados com foco na transparência e na reprodutibilidade**. Ambas serão descritas no sexto capítulo da tese.

### 5.1 Estudo de casos múltiplos e como ilustração

Esta tese de abordagem qualitativa apresenta duas principais inspirações metodológicas: o estudo de casos múltiplos (STAKE, 2006) para a análise de fenômenos e a modalidade híbrida do estudo de caso como ilustração (MACHADO; PALACIOS, 2007) adotada por pesquisadores do Grupo de Pesquisa em Jornalismo Online (GJOL), da Universidade Federal da Bahia (UBFA). Ambos os procedimentos são particularmente úteis para a observação sistemática de objetos empíricos digitais ainda em desenvolvimento.

Na metodologia híbrida de estudo de caso como ilustração, ocorrem três etapas principais, segundo Machado e Palacios (2007, p. 343): revisão preliminar da bibliografia e análise de organizações jornalísticas; delimitação do objeto com formulação de hipóteses e estudo de caso somado à pesquisa junto às organizações jornalísticas; e, por fim, elaboração de categorias de análise para processamento do material coletado e definições conceituais acerca dos objetos de estudo. “Nossa metodologia permite que o pesquisador, por um lado, revise a bibliografia corrente sobre o objeto e, por outro, possibilita que esta

produção conceitual seja testada em estudos de casos específicos.” Nesta tese, tomo a metodologia híbrida como inspiração porque não há, propriamente, observação junto aos processos produtivos nas redações jornalísticas. Aqui, analiso somente o produto.

O estudo de casos múltiplos (STAKE, 2006), por sua vez, é utilizado para examinar um fenômeno passível de ser explorado a partir de diferentes ângulos ou partes. Trata-se de uma abordagem efetiva para a análise de fenômenos. Nesta tese, o fenômeno estudado é a transparência no Jornalismo Guiado por Dados, cuja amostragem por intencionalidade (GIL, 2019b), não probabilística, envolve a seleção de casos semelhantes entre si. Na presente pesquisa, isso se materializa por meio de produtos jornalísticos informativos, nomeadamente notícias, reportagens e projetos orientados por dados.

Embora nesta metodologia cada caso seja estudado de maneira distinta, busca-se entender como os casos individuais podem contribuir para o entendimento do fenômeno como um todo. No estudo de casos múltiplos, o objeto que se pretende estudar é chamado de quintão (*quintain*, em inglês), formado por casos singulares que compartilham alguma característica em comum. Neste estudo, todas as notícias ou reportagens que compõem o *corpus* possuem elementos de transparência como ponto em comum. Em geral, seleciona-se de quatro a dez casos para compor uma investigação guiada por esse método. Isso porque, em um mesmo produto, torna-se difícil encontrar todas as características que se pretende estudar.

Para Stake (2006), uma das vantagens de adotar o estudo de casos múltiplos é o fato de poder levar em conta o contexto ou o sistema de forma ampla, características que por vezes se perdem no estudo de caso convencional. Ambos, é claro, apresentam limitações pontuais – tanto o estudo de caso quanto o estudo de casos múltiplos correspondem a uma realidade particular e não apresentam validade externa, ainda que possam se repetir em outras circunstâncias. Este tipo de método é adotado, portanto, para compreender um fenômeno específico e pertencente a um recorte. Assim, trata-se de uma alternativa para obter respostas motivadas por um objeto central não necessariamente amplo.

Neste ponto, a metodologia híbrida proposta no estudo de caso como ilustração (MACHADO; PALACIOS, 2007) converge para o propósito da metodologia dos estudos de caso múltiplos. Não havendo descrição exaustiva dos objetos como ocorre no estudo de caso convencional, o estudo de caso como ilustração é uma etapa do processo metodológico usada para ilustrar argumentos e demonstrar a validade ou refutar hipóteses. Dessa forma, essa metodologia apresenta de certa maneira um formato aberto, que possibilita

formulações futuras para abarcar aspectos até então desconsiderados pelos pesquisadores.

Conforme indicado anteriormente, o corpus empírico da tese é constituído por materiais informativos publicados em 2020 e 2021 por jornais de referência e nativos digitais brasileiros, tendo como tópico principal a pandemia de Covid-19. O período de análise do corpus da tese começa em março de 2020, quando foram detectadas as primeiras reportagens sobre o tema, e termina em agosto de 2021, por conta dos prazos da tese, visto que a cobertura da pandemia de Covid-19 não tem data para acabar. Analiso somente o material jornalístico já publicado, sem levar em conta a etapa de produção noticiosa.

Para selecionar os textos que integram esta amostra de conveniência, realizei, ao longo da pandemia, navegação flutuante e observação sistemática que me permitiu identificar, em nível de produto, aspectos relacionados à transparência no jornalismo à luz da teoria. E, para fins de rigor metodológico, observo orientações de composição de corpus para a pesquisa qualitativa, em que os materiais apresentam apenas um foco temático e, portanto, relevância, homogeneidade – a análise está restrita aos textos, por exemplo – e sincronicidade, que diz respeito à presença do recorte na história (BAUER; AARTS, 2015). Em relação ao *corpus* selecionado para compor a análise de casos múltiplos, a segmentação é bastante clara: a sincronicidade é garantida pela seleção de textos referentes ao mesmo período histórico, ou seja, ao início e ao desenvolvimento da pandemia de Covid-19. Além disso, as notícias e reportagens foram publicadas em língua portuguesa e no contexto brasileiro.

Integram o corpus os textos jornalísticos publicados pelos seguintes veículos jornalísticos: *Agência Lupa*<sup>96</sup>, *Agência Pública*<sup>97</sup>, *El País*<sup>98</sup> (Brasil), *Fiquem Sabendo*<sup>99</sup>, *Folha de*

---

<sup>96</sup> A *Agência Lupa* é uma agência digital de checagens brasileira fundada em novembro de 2015. Conta com 16 profissionais e apresenta um braço educativo para promover capacitações em fact-checking e debunking. Disponível em: <https://piaui.folha.uol.com.br/lupa/>

<sup>97</sup> A *Agência Pública* é um nativo digital fundado em 2011, e autointitulada a primeira agência de jornalismo investigativo sem fins lucrativos do Brasil. Em seu site, há uma seção específica para a publicação de reportagens guiadas por dados. Disponível em: <https://apublica.org/>

<sup>98</sup> *El País Brasil* é a versão digital do site do jornal *El País*, diário espanhol fundado em 1976. Apresentava reportagens esporádicas em português com o uso de técnicas de JGD. A redação brasileira teve suas atividades encerradas ao final de 2021. Disponível em: <https://brasil.elpais.com/>

<sup>99</sup> *Fiquem Sabendo* é uma agência especializada em Lei de Acesso à Informação e, portanto, reúne pedidos via Serviço de Informação ao Cidadão (SIC) e oferece, publicamente, os dados obtidos a partir de suas próprias investigações. Esta iniciativa passou a operar extraoficialmente em 2015, e reconhecida como empreendimento a partir de 2018. Disponível em: <https://fiquemsabendo.com.br/>

*S.Paulo*<sup>100</sup>, *GaúchaZH*<sup>101</sup>, *Jota*<sup>102</sup>, *Nexo*<sup>103</sup>, *Núcleo*<sup>104</sup>, *O Estado de S. Paulo*<sup>105</sup> e *Piauí*<sup>106</sup>. Ainda que alguns desses veículos tenham profissionais ou equipes que trabalham exclusivamente com a prática de Jornalismo Guiado por Dados, somente o *Núcleo* apresenta ênfase na produção de materiais derivados de técnicas de JGD. A agência *Fiquem Sabendo*, ainda que especializada em Lei de Acesso à Informação, atua muito mais como fonte para outros veículos, eventualmente publicando reportagens em seu próprio site. Parte dos jornais escolhidos ainda integra o *Projeto Credibilidade*<sup>107</sup>, que busca oferecer ferramentas e técnicas para ampliar a transparência no jornalismo.

As notícias e reportagens que serviram de base para a sistematização dos elementos de transparência no Jornalismo Guiado por Dados e que serão a seguir apresentadas representam, obviamente, apenas parte das possibilidades de um estudo desta natureza. Em outras palavras, esta pesquisa se propõe a detectar elementos de transparência à luz da teoria – não significa que esses são os únicos elementos possíveis de serem encontrados no jornalismo digital ou Jornalismo Guiado por Dados.

A escolha por jornais brasileiros se deu porque o JGD é uma prática crescente no país – somente em 2021, dez projetos nacionais estiveram entre os finalistas do maior

---

<sup>100</sup> O jornal *Folha de S.Paulo* teve sua versão impressa iniciada em 1921, e seu site foi criado em 1995, com o nome de *Folha Online*. Publica cerca de 160 notícias diariamente, e apresenta uma equipe voltada especificamente para o trabalho com dados: o *Delta Folha*. Disponível em: <https://www.folha.uol.com.br/>

<sup>101</sup> O portal *GZH* (*GaúchaZH*) é um site jornalístico que existe desde 2017 e integra o *Grupo RBS*, ao qual também pertence o jornal anteriormente chamado de *Zero Hora*, fundado em maio de 1964. Em 2020, o jornal criou o projeto *Fluência em Notícias* para mostrar à audiência como ocorre sua produção noticiosa. Disponível em: <https://gauchazh.clicrbs.com.br/>

<sup>102</sup> O *Jota* é uma empresa jornalística digital fundada em 2014 para cobrir o Poder Judiciário brasileiro. Apresenta uma seção específica para publicação de reportagens guiadas por dados. Disponível em: <https://www.jota.info/>

<sup>103</sup> O *Nexo* é um jornal digital lançado em novembro de 2015 com o objetivo de levar contexto às notícias e ampliar o acesso do público aos dados e estatísticas. Disponível em: <https://www.nexojournal.com.br/>

<sup>104</sup> O *Núcleo* é um site jornalístico lançado em 6 de janeiro de 2020 para liderar investigações a partir de dados, buscando trazer mais transparência ao governo, ao debate político e às políticas públicas. Tem a transparência como ponto principal de sua missão. Disponível em: <https://nucleo.jor.br/>

<sup>105</sup> O jornal *O Estado de S. Paulo* circula em versão impressa desde 4 de janeiro de 1875, ainda na época do Brasil Império (à época, o nome era *A Província de S. Paulo*; o nome atual veio em janeiro de 1890). Posteriormente, passou a publicar notícias também na internet. Seu site tem como ano inicial 1995. Em sua equipe, as produções de JGD estão associadas à editoria de Infografia. Disponível em: <https://www.estadao.com.br/>

<sup>106</sup> Lançada em 2006 como revista impressa – ainda em circulação, cabe ressaltar –, a *Piauí* apresenta também uma seção específica para publicações em JGD, chamada de *Igualdades*. Disponível em: <https://piaui.folha.uol.com.br/>

<sup>107</sup> O Projeto Credibilidade, também conhecido por seu nome em inglês – *Trust Project* – desenvolve, desde 2016, padrões de transparência para ajudar o público a avaliar a qualidade e a credibilidade do jornalismo. Dentre os veículos brasileiros mencionados nesta tese, integram o projeto os seguintes jornais: *Agência Lupa*, *Folha de S.Paulo*, *GaúchaZH* e *Nexo*. Os indicadores do *Projeto Credibilidade* para estimular a abertura de processos e métodos são a identificação clara dos seguintes tópicos: 1) Melhores práticas; 2) Jornalista; 3) Tipo de matéria; 4) Citações e referências; 5) Método; 6) Apuração local; 7) Diversidade de vozes; e 8) Feedback acionável. Disponível em: <https://www.credibilidade.org/>



prêmio de jornalismo de dados no mundo, chamado *Sigma Awards* –, porque o JGD desenvolvido na América Latina ainda é pouco estudado (DE LIMA; MESQUITA, 2021) e porque conheço com propriedade o contexto de publicação jornalística no Brasil durante a pandemia de Covid-19. Envolver publicações de outros países, ainda que pudesse trazer contribuições originais e ampliar os elementos de transparência detectados, demandaria uma análise de conjuntura muito específica em relação a cada uma dessas nações, visto que a produção jornalística depende do contexto em que ocorre.

Assim como Barbosa (2007), que buscava exemplos de caráter inovador e diferencial para a composição do material empírico à luz da teoria, selecionei casos jornalísticos que me permitiram observar a articulação dos elementos de transparência na cobertura da Covid-19. Alguns veículos jornalísticos se repetem entre os exemplos trazidos, devido à maior frequência na utilização de certos elementos de transparência. Em um mapeamento inicial, identifiquei cerca de 150 materiais informativos observados principalmente entre esses 11 veículos de comunicação, fazendo a leitura dos textos e selecionando os casos mais relevantes para compor o corpus empírico desta tese. Quando possível, restringi minhas buscas às seções de Jornalismo Guiado por Dados oferecidas pelas empresas (a exemplo do *Delta Folha*, da *Folha de S.Paulo*, da seção *Igualdades*, da *Piauí*, e da área de *Dados da Agência Pública*).

Por fim, destacamos que a vantagem de apresentar um conjunto formulado a partir de amostra de conveniência é sua especificidade e relação com o objeto teórico abordado. Além disso, a inspiração no estudo de casos múltiplos (STAKE, 2006) e estudo de caso como ilustração (MACHADO; PALACIOS, 2007) permite que a conformação de categorias possibilite que novas pesquisas avancem para uma etapa quantitativa de análise para além desta primeira sistematização, gerando novos tipos de conhecimento.

## 5.2 Transparência e seus níveis: organizacional e operacional

Diferentemente de seu precursor chamado de Reportagem Assistida por Computador (RAC), em que os dados eram encarados como mais uma fonte, sem necessariamente apresentar protagonismo na reportagem, o Jornalismo Guiado por Dados inclui a transparência como aspecto-chave de seus processos e resultados (CODDINGTON, 2014), visto que, como dissemos anteriormente, está ancorado em uma filosofia de código aberto e colaboração (TRÄSEL, 2014), oriunda da comunidade de desenvolvedores e cientistas da computação.

Para que se sustente como **operadora de uma prática ética e recurso discursivo de legitimação**, a transparência, a rigor, tem de ser um elemento que repercute tanto em nível organizacional quanto em nível operacional – neste caso, deve se refletir também no produto. Isso significa que, idealmente, não basta que os jornalistas adotem e demonstrem clareza no seu percurso metodológico. É necessário que as empresas jornalísticas ofereçam clareza quanto a tópicos mais amplos, como a política de correção de erros, a política de diversidade adotada na empresa e, não menos importante, suas formas de financiamento. O conhecimento sobre a origem dos recursos tende a ser um bom aliado da audiência para avaliar o tipo de cobertura feita (ou ignorada) por um jornal.

Como veremos a seguir, dificilmente um mesmo veículo jornalístico irá apresentar, sozinho, todos os elementos de transparência possíveis e viáveis. Ainda que a cobertura de Covid-19 seja marcante em nível operacional, é preciso considerar que a publicação de reportagens integra um contexto mais amplo de decisões e processos editoriais presentes no jornal como um todo. Por isso, não há como ignorar esta etapa anterior, inerente às decisões editoriais que repercutem em toda a estrutura organizacional (o nível organizacional da transparência), e não somente no conteúdo informativo do Jornalismo Guiado por Dados (o nível operacional da transparência).

Além disso, o uso de um ou mais elementos de transparência não garante, necessariamente, que o veículo de comunicação esteja atento às premissas científicas ligadas ao Jornalismo Guiado por Dados, ou que siga todos os métodos sugeridos para esta prática. Os elementos apresentados a seguir possuem como objetivo a sistematização de conhecimento e a exemplificação, e não têm a finalidade de apontar que um veículo de comunicação é mais ou menos transparente do que o outro.

Apresento, a seguir, os tópicos correspondentes aos dois níveis verificados quanto à transparência no Jornalismo Guiado por Dados: **organizacional** (quanto ao veículo, envolvendo também aspectos sobre o jornalista que trabalha nessa organização) e **operacional** (quanto ao método).

Os exemplos listados seguem a inspiração metodológica do estudo de casos múltiplos (STAKE, 2006) e do estudo de caso como ilustração (MACHADO; PALACIOS, 2007). Por conta das escolhas metodológicas desta tese, os elementos de transparência são listados e analisados de forma individual e exclusiva, ou seja, não apresenta o detalhamento de explicações que normalmente ocorre no estudo de caso. Em alguns casos, cada elemento terá apenas o exemplo mais significativo para a sua seção – conforme preconizado pela metodologia de estudos de casos múltiplos, para o qual a compreensão de um fenômeno não demanda exploração quantitativa, mas qualitativa; há algumas exceções, porém, em que mais de um caso é apresentado por elemento para fins de contextualização.

Cabe esclarecer, finalmente, que não estou avaliando se um mesmo elemento de transparência é constantemente utilizado e atualizado dentro de uma empresa jornalística. Esta pesquisa, assim como outros estudos qualitativos, representam uma fotografia ampla do momento em que a análise foi empreendida.

### 5.2.1 Nível organizacional: o jornal

Em nível organizacional, a transparência aparece como **recurso discursivo de legitimação**, isto é, o quanto um jornal se mostra transparente a partir de suas estratégias de aproximação com a audiência. Nesta seção, apresento oito elementos relacionados à postura do profissional e do veículo de comunicação: **assinatura, perfil e expertise; contato; correção de erros; fontes de financiamento; manual da redação e princípios editoriais; política de diversidade; política de privacidade; e tipo de conteúdo.**

### 5.2.1.1 Assinatura, perfil e expertise

Indicar o nome do profissional responsável pela apuração de uma reportagem costuma ser um requisito básico presente nos manuais de redação. Como estratégia de transparência, porém, deve ir além, envolvendo também informações sobre o perfil do jornalista, como os trabalhos por ele ou ela desenvolvidos anteriormente; junto a essa descrição, devem aparecer aspectos sobre a sua própria formação, indicando suas credenciais. Se as habilidades e o conhecimento do autor de uma reportagem forem de conhecimento público, o leitor terá condições de julgar o nível de especialização do profissional e o quanto ele está preparado para cobrir determinado tema.

Por isso, nesta seção, apresentamos três características que estão entrelaçadas: a **assinatura**, que se refere à clareza de divulgação do nome da pessoa responsável pelo texto; o **perfil**, que diz respeito à possibilidade de acessar uma página de histórico de outras reportagens publicadas por essa mesma pessoa; e **expertise**, que mostra as competências desenvolvidas pelo profissional. Nem sempre os veículos jornalísticos irão apresentar todas essas características; o mais comum é oferecer apenas o nome do repórter, sem detalhamento.

No exemplo a seguir, observamos a descrição do perfil do repórter na publicação intitulada *TrateCov indicava cloroquina e outros remédios sem eficácia a quase todos pacientes*, veiculada pelo *Jota* (MARCELINO, 2021).

**Figura 8 – Conhecimento prévio do repórter é listado em notícia no *Jota***

The image shows a screenshot of a news article on the Jota website. The article title is "TrateCov indicava cloroquina e outros remédios sem eficácia a quase todos pacientes". Below the title, there is a sub-headline: "Após polêmica, saiu do ar aplicativo desenvolvido pelo Ministério da Saúde que orientava médicos em 'tratamento precoce'". The author's name, Daniel Marcelino, is displayed in a small box with a profile picture. Below the author's name, there is a video call interface showing Daniel Marcelino speaking. At the bottom of the article, there is a bio box for Daniel Marcelino, which reads: "DANIEL MARCELINO – Cientista de dados em Brasília, é especialista em métodos quantitativos, modelos de previsão e pesquisas de opinião. Antes do JOTA, foi pesquisador em universidades e órgãos de governos: Universidade de York e Universidade de Montreal (Canadá), IPEA (Instituto de Pesquisa Econômica Aplicada) e Prefeitura de Curitiba. Email: daniel.marcelino@jota.info".

Fonte: *Jota*. Captura de tela em 18 maio 2021.

A **Figura 8** apresenta em destaque o nome do autor do texto, logo abaixo do título da notícia, e mais detalhes sobre sua área de formação e expertise que funciona como uma nota de rodapé ao final do texto. Não há, porém, hiperlink que leve o público a conhecer outros trabalhos desenvolvidos por ele.

No caso da Covid-19 e em outros tantos exemplos que exigem conhecimento técnico prévio, trata-se de uma estratégia de transparência que dialoga não apenas com a credibilidade do veículo, mas também com a confiabilidade assegurada pelas credenciais do repórter. Nesta reportagem em questão, o profissional desvendou o código-fonte utilizado pelo Ministério da Saúde para recomendar o uso de medicamentos sem eficácia para supostamente prevenir a doença. Contudo, a transparência para por aí nesta reportagem do *Jota*: o código-fonte examinado na notícia – correspondente ao elemento de reprodutibilidade, como veremos adiante – sequer é aberto ao público.

Em outro exemplo, na reportagem intitulada *Cidades mais desenvolvidas do país vacinam mais rápido contra a Covid* (BARBON *et al.*, 2021), publicada pela *Folha de S.Paulo*, a lista de autores é divulgada ao final da notícia (ao centro na **Figura 9**).

**Figura 9 – Nome dos repórteres funciona como indexador na *Folha***

DELTA FOLHA · CORONAVÍRUS

## Cidades mais desenvolvidas do país vacinam mais rápido contra a Covid

Rede de saúde forte acelera imunização, já falta de uniformidade dos municípios é fator de entrave

RIO DE JANEIRO e SÃO PAULO Cidades mais ricas e desenvolvidas têm avançado mais rápido na vacinação contra a Covid-19. Em geral, municípios com melhor IDH (Índice de Desenvolvimento Humano) já conseguiram imunizar mais fatis da sua população do que aqueles com índice mais baixo.

Isso se deve principalmente, segundo especialistas e gestores, à existência de redes de saúde básica já fortes, calendários mais eficientes e conscientização constante. Esses locais também têm mais idosos e profissionais da saúde grupos vacinados antes e atraem gente de regiões vizinhas.

A **Folha** analisou microdados do Ministério da Saúde dos 326 municípios com mais de 100 mil habitantes do país até o último dia 21, para evitar distorções por atrasos na notificação. Naquela data, o Brasil havia imunizado com a primeira dose 30% da população em geral — hoje são 36%.

Entre as 130 cidades que estavam acima dessa marca, 102 têm IDH superior à média dos grandes municípios. Quase todas ficam em estados do Sul e do Sudeste — apenas 12 se localizam em outras regiões, sendo que nove delas são capitais.

"Isso mostra que o melhor desempenho na vacinação tem a ver com a organização da cidade, não só com a riqueza. O IDH se baseia em renda, educação e expectativa de vida, portanto é uma população mais idosa, com mais acesso à informação e menos problema para se deslocar", analisa Christovam Barcellos, geógrafo e sanitarista da Fiocruz.

Pesquisadores da Fiocruz calcularam que 15% das doses usadas até o dia 16 no país foram aplicadas em pessoas que se deslocaram para outros municípios. A distância média percorrida por esses brasileiros foi de 252 km.

"Isso é resultado da falta de padronização nas campanhas. Pelo menos as regiões metropolitanas deveriam ter um calendário único de vacinação para evitar esse desespero e essa correria entre as cidades", diz o sanitarista Barcellos.

O estado do Rio de Janeiro, um dos mais atrasados na campanha, é um exemplo da disparidade entre capital e outras cidades: 25 dos seus 30 municípios com mais de 100 mil habitantes estavam abaixo da média nacional vacinada até o último dia 21.

"Um dos maiores problemas foi a falta de uniformização dos grupos prioritários e do calendário. Cada município adotou um comportamento. Agora, eles se comprometeram a seguir um cronograma único junto ao estado e observamos uma melhora", diz Alessandra Nascimento, subcoordenadora de Saúde e Tutela Coletiva da Defensoria Pública do RJ.

Há que se ponderar, ainda, a defasagem nas notificações da imunização feitas pelas secretarias estaduais e municipais ao Ministério da Saúde. A **Folha** adotou as cidades com mais de 100 mil habitantes porque elas costumam ter um volume de dados mais constante.

Júlia Barbon, Flávia Faria, Diana Yukari e Daniel Mariani

tópicos  
LEIA TUDO SOBRE O TEMA E SIGA:  
coronavirus

**Júlia Barbon**  
Repórter no Rio de Janeiro  
É repórter na Folha desde 2014 e correspondente no Rio de Janeiro desde 2018. Escreve sobre temas como segurança pública, educação e saúde.

CORONAVÍRUS

## Casos de síndrome respiratória voltam a subir no país, calcula Fiocruz

Tendência de queda da SRAG se reverteu pela primeira vez em quase três meses  
11 ago. 2021 às 19h42

RIO DE JANEIRO

### Paes diz que Rio não recebeu vacinas suficientes e pede ajuda a Doria

Prefeitura suspendeu a imunização contra a Covid nesta quarta (11) e agora também nesta quinta (12)  
11 ago. 2021 às 19h47

PODER

### Bolsonaro repete discurso pró-voto impresso e cita 'legitimidade' do Congresso após motociata em SC

Em Florianópolis, presidente criticou STF e fez aceno a parlamentares, Câmara vai analisar PEC em plenário  
7 ago. 2021 às 19h53

Fonte: *Folha de S. Paulo*. Captura de tela em 13 ago. 2021.

Em destaque no centro e à direita da **Figura 9**, é possível observar que o nome de cada repórter funciona, por meio de hiperlink, como uma espécie de indexador de trabalhos anteriores. Além disso, é possível encontrar, junto ao nome, as principais informações profissionais do repórter (no topo da **Figura 9**, à direita).

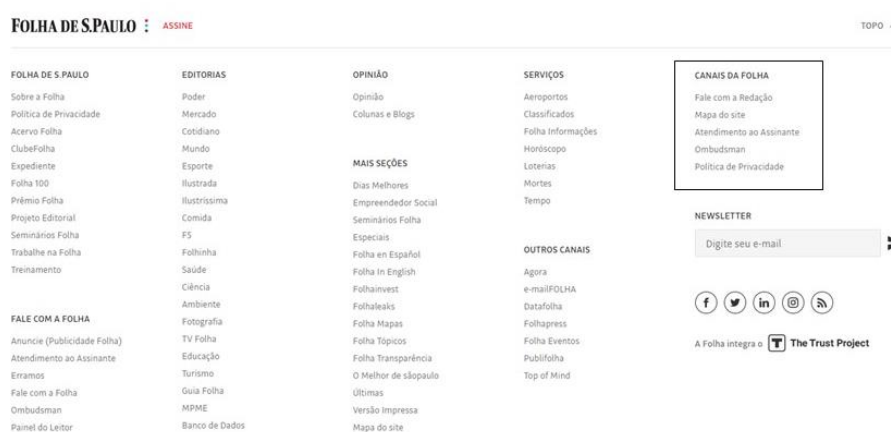
Além das informações sobre os profissionais responsáveis pela reportagem, idealmente o veículo jornalístico deveria incluir dados sobre o editor que coordena cada seção. Em linhas gerais, o leitor tem condições para julgar o trabalho do repórter se cumpridos os principais preceitos desta seção; menos frequente, porém, é o conhecimento do público em relação à estrutura de trabalho de uma empresa jornalística, em que a cadeia de produção noticiosa é muito mais complexa, passando, sempre, por um ou mais editores antes da publicação.

### 5.2.1.2 Contato

Listas e formulários de contato com as redações não são novidade na imprensa brasileira e representam uma ponte entre o veículo jornalístico e o leitor. Este tipo de abertura torna-se fundamental em tempos de crise sanitária, em que o jornalismo reforça seu papel de prestação de informações públicas relevantes. As formas de contatar a empresa jornalística para comunicar um erro ou mesmo comentar alguma notícia devem ser evidentes, portanto.

A *Folha de S.Paulo* oferece uma lista de contatos a partir da seção *Fale com a Folha*, disponível no rodapé do site<sup>108</sup>. O jornal apresenta diferentes canais de comunicação com o leitor, incluindo contato pelo *WhatsApp* e diferentes endereços de e-mail, que variam de acordo com a editoria a ser contatada. Além disso, a *Folha* é aqui mencionada como exemplo porque apresenta a figura do *ombudsman*, que aparece como uma possibilidade de contato no rodapé do site, e analisa criticamente, em sua coluna<sup>109</sup> (**Figura 10**), o conteúdo jornalístico publicado.

**Figura 10 – Contato com o ombudsman é um dos destaques na Folha**



Fonte: *Folha de S.Paulo*. Captura de tela em 21 jul. 2021.

<sup>108</sup> Disponível em: <https://www1.folha.uol.com.br/falecomafolha/>

<sup>109</sup> Disponível em: <https://www1.folha.uol.com.br/ombudsman/>

A função de *ombudsman* é ocupada desde 10 de maio de 2021 por José Henrique Mariante, com o qual o leitor pode ter contato direto via e-mail ou telefone (em destaque no canto inferior direito na **Figura 10**); antes dele, Flávia Lima era a responsável. A *Folha* foi o primeiro jornal brasileiro a adotar essa função, em 1989. Além de funcionar como um elo entre o leitor e o jornal, a figura do *ombudsman* é fundamental para a demonstração da correção de erros de um veículo, como veremos a seguir.

### 5.2.1.3 Correção de erros

A política de correção de erros diz respeito ao modo como a empresa lida com problemas de apuração e falta de acurácia detectados após a publicação de um conteúdo. Normalmente está presente no seu próprio site e/ou em espaços como o manual da redação, reunindo o passo a passo de como um erro deve ser esclarecido textualmente e comunicado ao leitor. Trata-se de um elemento básico para que haja uma relação de honestidade com o público.

No *Nexo*, a política de correção de erros está explícita dentro da seção de padrões editoriais<sup>110</sup> e, quando aplicada nas notícias que necessitam de algum tipo de retratação, uma errata é registrada ao final da notícia. De acordo com o material disponibilizado pela redação (na parte superior da **Figura 11**), busca-se corrigir os dados equivocados imediatamente e de forma transparente.

---

<sup>110</sup> Disponível em: <https://www.nexojornal.com.br/sobre/Padr%C3%B5es-editoriais-do-Nexo>



**Figura 11 – Nexo apresenta errata ao final das notícias**



Fonte: *Nexo Jornal*. Captura de tela em 18 ago. 2021, com edições da autora.

Na publicação do dia 29 de março de 2021, intitulada *Por que tanta gente ignora regras no pior momento da pandemia* (ROCHA, 2021) e demonstrada do centro para a parte inferior da **Figura 11**, o jornal traz uma nota clara apresentando o que estava errado e o que foi modificado em relação à notícia original, indicando, ainda, data e horário da correção.

Similarmente, a *Folha de Paulo* apresenta uma política aberta de correção de erros, cujas decisões são detalhadas no projeto editorial disponível publicamente no site<sup>111</sup>. O acesso a este material ocorre por meio do rodapé da página, em que primeiramente o leitor encontra a seção *Erramos*<sup>112</sup>, onde estão listados os textos que sofreram algum tipo de alteração.

A **Figura 12** mostra um exemplo de como isso ocorre na prática. Em 2 de julho de 2021, o jornal publicou a reportagem intitulada *Registros indicam que milhares no Brasil tomaram vacina vencida contra Covid; veja se você é um deles* (GAMBA; RIGHETTI, 2021). A partir de dados do Ministério da Saúde, o texto indica que 26 mil doses da vacina Astra-Zeneca fora da validade foram aplicadas em 1.532 municípios.

<sup>111</sup> Disponível em: <https://temas.folha.uol.com.br/folha-projeto-editorial/manual-de-redacao-conduta/erramos.shtml>

<sup>112</sup> Disponível em: <https://www1.folha.uol.com.br/erramos/>

**Figura 12 – Erramos é destacado pela *Folha* junto à data e ao final do texto**



Fonte: *Folha de S.Paulo*. Captura de tela em 18 jul. 2021.

No dia 6 de julho de 2021, o jornal adicionou a errata, cujo aviso passa a ficar exposto junto à data original de publicação da reportagem (em destaque à esquerda na **Figura 12**) e ao final do texto (em destaque na parte inferior da **Figura 12**). A correção informa o seguinte: “Versão anterior desta reportagem deixou de alertar que os dados sobre 26 mil doses aplicadas fora do prazo de validade poderiam decorrer de erros do sistema do Ministério da Saúde. O texto foi atualizado” (GAMBA; RIGHETTI, 2021, s.p.).

A breve correção da *Folha* gerou insatisfação dos leitores, que criticaram a publicação original e os problemas na verificação do banco de dados que gerou a reportagem original. O *ombudsman* José Henrique Mariante repercutiu o caso na coluna de 10 de julho de 2021, intitulada “Jornalismo vencido”<sup>113</sup>.

Por último, ao final dos textos informativos, a *Folha de S.Paulo* apresenta ícones para compartilhamento do conteúdo nas redes sociais digitais, espaço para comentários e um botão para a comunicação de erros (Na **Figura 13**, sinalizado pela seta).

<sup>113</sup> Disponível em: <https://www1.folha.uol.com.br/colunas/jose-henrique-mariante-ombudsman/2021/07/jornalismo-vencido.shtml>

**Figura 13 – Botão para comunicação de erros aparece na *Folha***

The image shows a screenshot of the Folha de S. Paulo website. On the left, there is a blue box with the text 'sua assinatura vale muito' and 'Mais de 180 reportagens e análises publicadas a cada dia. Um time com mais de 120 colunistas. Um jornalismo profissional que fiscaliza o poder público, veicula notícias proveitosas e inspiradoras, faz contraponto à intolerância das redes sociais e traça uma linha clara entre verdade e mentira. Quanto custa ajudar a produzir esse conteúdo?' Below this box is an orange button that says 'ASSINE A FOLHA'. To the right of the blue box are two buttons: 'ENVIE SUA NOTÍCIA' and 'ERRAMOS?'. An arrow points from the 'ERRAMOS?' button to the 'Comunique erros' form on the right. The form has a title 'Comunique erros' and a subtitle 'Para comunicar erros nos textos publicados pela Folha de S. Paulo e Agora, preencha os dados abaixo e clique em Enviar.' Below the subtitle is a warning icon and the text 'ATENÇÃO Se você quiser falar com Ombudsman, Assinaturas ou deseja enviar e-mail sobre outros assuntos, clique aqui.' The form is titled 'PREENCHA OS DADOS' and has fields for 'Nome:', 'Sobrenome', 'E-mail\*', and 'Comentário\*'. At the bottom of the form is a checkbox labeled 'Não sou um robô' and a CAPTCHA logo.

Fonte: *Folha de S. Paulo*. Captura de tela em 18 out 2021.

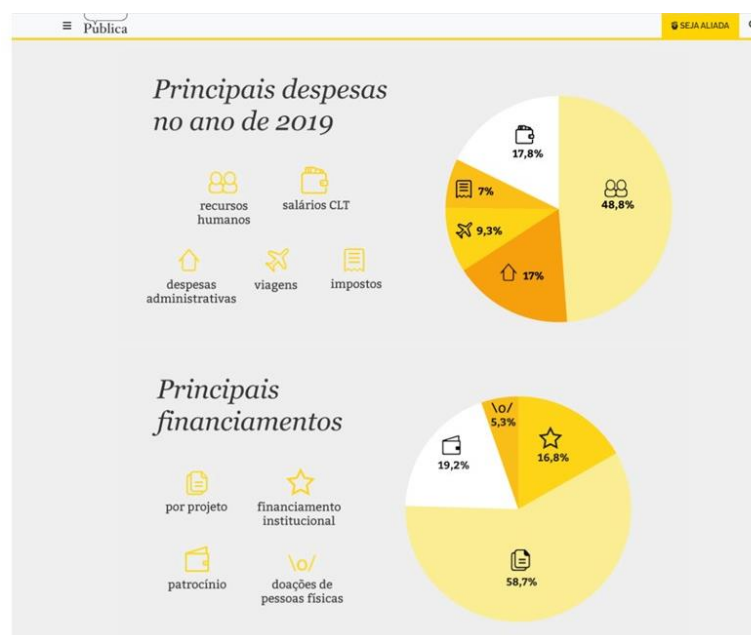
Ao acionar o botão *Erramos?*, o leitor encontra um formulário (à direita na **Figura 13**) em que tem a oportunidade de descrever o erro encontrado ou mesmo ser redirecionado para o contato direto com o *ombudsman* do jornal.

#### 5.2.1.4 Fontes de financiamento

Divulgar a origem do dinheiro que mantém um empreendimento jornalístico é um dos aspectos básicos da transparência, visto que a sustentabilidade de uma redação ou consórcio de profissionais é o primeiro passo para a realização de qualquer atividade que exige remuneração. Rastrear a origem do dinheiro permite, em suma, que os leitores avaliem a existência de conflitos de interesse, por exemplo. Por isso, deve-se deixar claro quais são as principais fontes de financiamento, com detalhes sobre receitas e despesas, em especial quando o leitor é parte primordial para angariar recursos (por meio de assinaturas ou financiamento coletivo, por exemplo).

A *Agência Pública* oferece visualizações com os principais dados (**Figura 14**).

**Figura 14 – Agência Pública apresenta gráficos sobre despesas e receitas**



Fonte: Agência Pública. Captura de tela em 18 jul. 2021.

Fundada em 2011, a *Agência Pública*, empreendimento digital de jornalismo investigativo sem fins lucrativos e que apresenta uma seção dedicada ao uso de dados no jornalismo, disponibiliza em sua página de transparência<sup>114</sup> o nome das principais entidades patrocinadoras, além das fontes de financiamento ao longo de um ano (**Figura 14**) e as regras<sup>115</sup> para quem deseja apoiar, garantindo que não haja interferência nas publicações jornalísticas ou que não haja leitura prévia, por parte das fontes, antes da publicação de um texto ou investigação.

Outra organização jornalística que prioriza a divulgação de receitas e despesas em uma seção de seu site, intitulada prestação de contas<sup>116</sup>, é a agência *Fiquem Sabendo*, especializada em Lei de Acesso à Informação.

<sup>114</sup> Disponível em: <https://apublica.org/transparencia/>

<sup>115</sup> Disponível em:

<https://apublica.org/wp-content/uploads/2021/06/politica-financiamento-publica2021.pdf>

<sup>116</sup> Disponível em: <https://fiquemsabendo.com.br/prestacao-de-contas/>

**Figura 15 – Fiquem Sabendo possui seção específica de prestação de contas**



A Fiquem sabendo é uma organização sem fins lucrativos, independente e apartidária. Conheça nossos documentos e confira nossos relatórios anuais.

Receitas

Nossas Receitas	2019	2020
Prestação de serviços	49,59%	21,01%
Bolsa para pesquisa	4,39%	22,32%
Doação Pessoas Físicas	12,99%	51,45%
Financiamento coletivo	20,16%	5,22%
Venda de itens promocionais	12,87%	0,00%

Fonte: *Fiquem Sabendo*. Captura de tela em 16 ago. 2021.

Além de mostrar ao leitor uma tabela com informações sobre percentuais e receitas obtidas anualmente, a agência *Fiquem Sabendo* reúne outras documentações na seção de prestação de contas, a exemplo de demonstrativo financeiro e balanço patrimonial. Outro dado relevante é a identificação nominal de seus principais financiadores individuais, além de um mural com uma lista de todos os apoiadores.

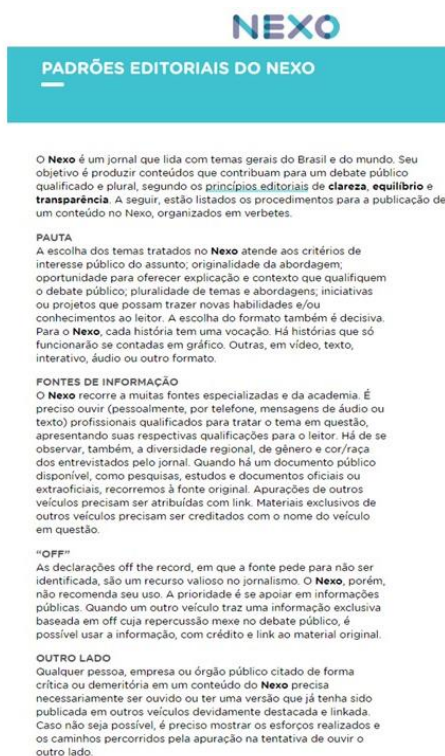
Mencionados esses exemplos, é preciso pontuar, ainda, que todo veículo jornalístico irá divulgar somente aquilo que considerar necessário e prudente, provocando o efeito de transparência sobre o qual falamos anteriormente. Esse tipo de conduta – abertura de finanças com certo nível de detalhamento – costuma ser mais frequente entre os nativos digitais, que normalmente dependem de doações de grandes investidores, organizações sem fins lucrativos ou mesmo assinaturas para sobreviver.

### 5.2.1.5 Manual da redação e princípios editoriais

Detalhes sobre a conduta do repórter em seus processos de apuração, bem como informações acerca da missão de uma empresa jornalística e sua hierarquia, além de normas técnicas e de escrita, costumam estar presentes no manual da redação e/ou em um espaço destinado à política editorial do jornal. Por meio desse material, o leitor torna-se capaz de compreender a prática jornalística que vai além de questões estilísticas, acentuada no âmbito das decisões editoriais.

Assim, o manual da redação serve como um guia ao qual os usuários e assinantes podem recorrer quando desconfiam que o jornal não está cumprindo premissas básicas do jornalismo. Este material normalmente inclui, por exemplo, o código de conduta dos jornalistas. O jornal digital *Nexo*, que possui uma relação estreita com seus leitores, inclusive por meio de grupos fechados nas redes sociais digitais, disponibiliza em sua página a *Política Editorial*<sup>117</sup>:

Figura 16 – *Nexo* apresenta aos leitores sua política editorial



Fonte: *Nexo Jornal*. Captura de tela em 18 jul. 2021.

<sup>117</sup> Disponível em: <https://www.nexojornal.com.br/sobre/Padr%C3%B5es-editoriais-do-Nexo>

Os padrões editoriais preconizados pelo Nexo (**Figura 16**) trazem informações sobre a escolha dos temas que se tornam pautas e como as fontes relacionadas a esse tópico são escolhidas. Revela, ainda, quais são os critérios levados em conta para o emprego de fontes anônimas e a consulta ao “outro lado”, quando a informação veiculada exige um pronunciamento oficial.

Por fim, a política editorial apresentada pelo *Nexo* também dialoga com questões éticas e a política de diversidade proposta pela redação. Fundado em novembro de 2015, o *Nexo* possui participação<sup>118</sup> no Projeto Credibilidade, assim como outros jornais mencionados nesta tese, que pressupõe padrões de qualidade a serem cumpridos em nome da acurácia das informações.

#### 5.2.1.6 Política de diversidade

Assegurar a existência de diferentes temas de interesse público é um dos compromissos do jornalismo e, para cumpri-lo, essa premissa deve estar prevista e exposta entre os propósitos de uma empresa jornalística. Para tanto, observa-se a necessidade de contratar profissionais com diferentes formações e vivências, de modo que se possa administrar uma variedade de temas para cobertura e debate.

Na *Folha de S.Paulo*, por exemplo, a política de diversidade está localizada entre as condutas previstas pelo *Manual da Redação*<sup>119</sup>.

---

<sup>118</sup> Disponível em: <https://www.nexojornal.com.br/sobre/Sobre-o-Nexo>

<sup>119</sup> Disponível em: <https://temas.folha.uol.com.br/folha-projeto-editorial/manual-de-redacao-conduta/politica-de-diversidade.shtml>

**Figura 17 – Política de diversidade é apresentada na *Folha***



Fonte: *Folha de S. Paulo*. Captura de tela em 22 jul. 2021.

A **Figura 17** mostra que, em seu comunicado, a *Folha* reforça a importância de implementar uma política de diversidade para assegurar a pluralidade de vivências de seus jornalistas, permitindo novos pontos de vista em relação aos temas de cobertura jornalística. Em termos práticos, o jornal abriu um cargo específico: em maio de 2021, Flávia Lima, ex-ombudsman, assumiu a função de editora de diversidade do jornal<sup>120</sup>.

A seção já havia entrado em operação dois anos antes para ampliar a variedade de profissionais que trabalham no jornal, bem como fontes ouvidas levando em conta questões raciais, de gênero e geográficas. Uma das primeiras medidas em consonância a essa política foi a criação de uma turma do programa de *trainees* da *Folha* destinado exclusivamente para pessoas negras. A primeira turma, cujas atividades se iniciaram em maio de 2021, é formada por 18 integrantes.

Uma das limitações inerentes à política de diversidade da *Folha* e de outros veículos que se propõem a fazê-lo diz respeito à dificuldade de monitoramento dessas ações, ou seja, além do que é abertamente comunicado pelo veículo, é difícil de monitorar e cobrar o cumprimento dessas políticas.

<sup>120</sup> Disponível em: <https://www1.folha.uol.com.br/poder/2021/05/jornalismo-nao-fala-mais-so-para-classe-media-branca-diz-nova-editora-de-diversidade-da-folha.shtml>

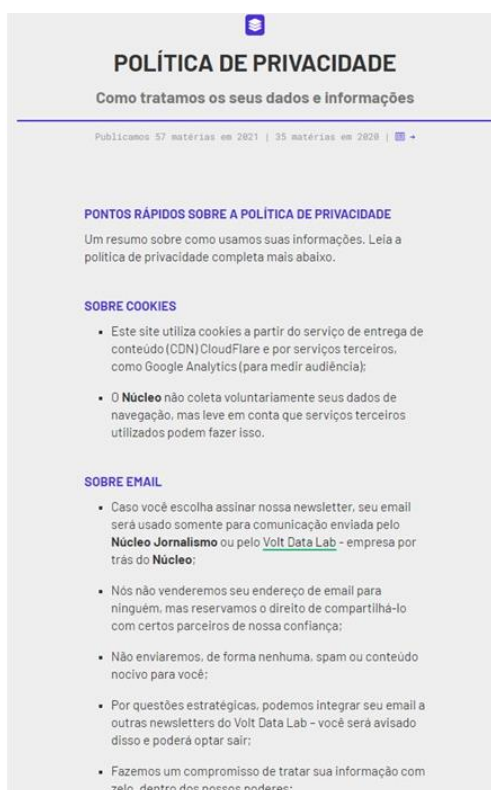


### 5.2.1.7 Política de privacidade

A Lei Geral de Proteção de Dados Pessoais (LGPD), aprovada em agosto de 2018 e em vigor no Brasil dois anos mais tarde, assegura, entre outros aspectos, que os cidadãos recebam informações claras e acessíveis sobre o tratamento atribuído aos seus dados pessoais. Desde meados de 2020, então, é comum que as empresas que exigem algum tipo de cadastro disponibilizem informações sobre adequações e usos dos dados pessoais dos brasileiros.

No jornalismo, isso também ocorre em se tratando de rastros digitais e cadastros em serviços de *newsletter*, por exemplo. O site jornalístico *Núcleo* detalha o tipo de informação que é coletada pelo site, disponibilizando uma lista em uma página específica<sup>121</sup> sobre privacidade dos dados coletados pelo site:

**Figura 18 – *Núcleo* abre política de privacidade aos leitores**



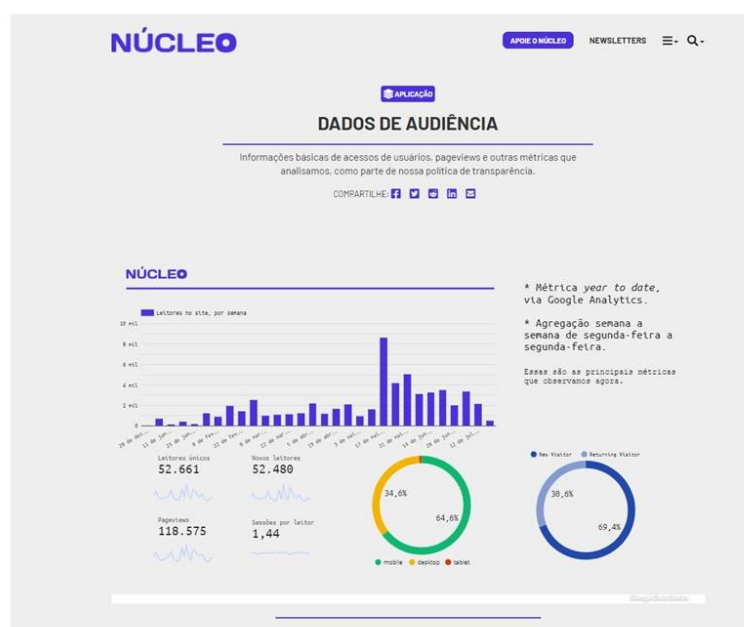
Fonte: *Núcleo Jornalismo de Dados*. Captura de tela em 18 jul. 2021.

<sup>121</sup> Disponível em: <https://nucleo.jor.br/privacidade>

Visto que as plataformas digitais se tornam uma forma de os veículos ganharem dinheiro ou comercializarem publicidade direcionada, torna-se essencial esclarecer o que é feito com as informações coletadas dos usuários.

Adicionalmente, o *Núcleo* fornece dados sobre a audiência<sup>122</sup> que acessa o site. As informações provêm do *Google Analytics*, ferramenta utilizada para contabilizar dados de acesso a uma página Web.

**Figura 19 – Núcleo traz dados de tráfego de seu site**



Fonte: *Núcleo Jornalismo de Dados*. Captura de tela em 18 jul. 2021.

Neste caso, é perceptível a busca do *Núcleo* em se mostrar transparente à audiência. Por questões de concorrência, a maior parte dos veículos de comunicação não divulga esse tipo de informação. No final de 2021, o jornal também passou a divulgar alguns dados demográficos, bem como os principais sites e redes sociais de entrada para acesso de seu conteúdo.

<sup>122</sup> Disponível em: <https://nucleo.jor.br/audiencia>

### 5.2.1.8 Tipo de conteúdo

A distinção entre conteúdo informativo (notícias e reportagens) e opinativo (colunas e editoriais, por exemplo) normalmente é explicitada por meio do uso de cartolas ou outros recursos gráficos bastante claros em páginas de jornais e páginas Web.

Durante a pandemia de Covid-19, alguns veículos de comunicação inclusive passaram a liberar gratuitamente o acesso ao conteúdo informativo relacionado à doença, visto que se trata de um problema universal de saúde pública. Justamente por isso, tornou-se ainda mais fundamental indicar a cobertura noticiosa e diferenciá-la do conteúdo opinativo. Neste período de pandemia, jornalistas precisaram avaliar com cuidado os especialistas ouvidos como fonte.

Alguns veículos de comunicação aproveitaram as circunstâncias da pandemia para criar estratégias relacionadas à educação midiática junto à audiência. Imprensa de referência no Sul do Brasil, o site *GaúchaZH* criou a série *Fluência em Notícias*. Uma de suas publicações, divulgada em 15 de outubro de 2020, recebe o título de *Veja como identificar o que é notícia e o que é opinião em GZH*<sup>123</sup>:

**Figura 20 – GZH criou série chamada *Fluência em Notícias***



Fonte: *Gaúchazh*. Captura de tela em 22 jul. 2021.

<sup>123</sup> Disponível em:

<https://gauchazh.clicrbs.com.br/geral/noticia/2020/10/veja-como-identificar-o-que-e-noticia-e-o-que-e-opinio-em-gzh-ckg9qjytv000d015xwqi0piff.html>

Como é possível visualizar na **Figura 21**, cada tópico do projeto apresenta uma breve explicação textual sobre sua proposta e um vídeo que apresenta, em detalhes, como o leitor ganha autonomia para identificar as diferenças entre informação e opinião. A série *Fluência em Notícias* funciona como uma iniciativa de alfabetização midiática e ainda apresenta temas como o combate à desinformação, sugestões de pauta e informações sobre como uma notícia é feita na redação. Reconhece, então, que a abertura de parte dos processos editoriais é fundamental para que a audiência compreenda como os produtos jornalísticos são criados.

### 5.2.2 Nível operacional: o método

Em nível operacional, a transparência aparece no produto jornalístico como **operadora de uma prática ética**, isto é, o quanto demonstra, em seus métodos e procedimentos, informações sobre fontes consultadas, documentos originais e permissão para verificação e reprodutibilidade da metodologia. Nesta seção, apresento sete elementos: **data de publicação; documentos; fontes de dados; graus de incerteza; marcas de apuração; metodologia; e reprodutibilidade e código aberto.**

#### 5.2.2.1 Data de publicação

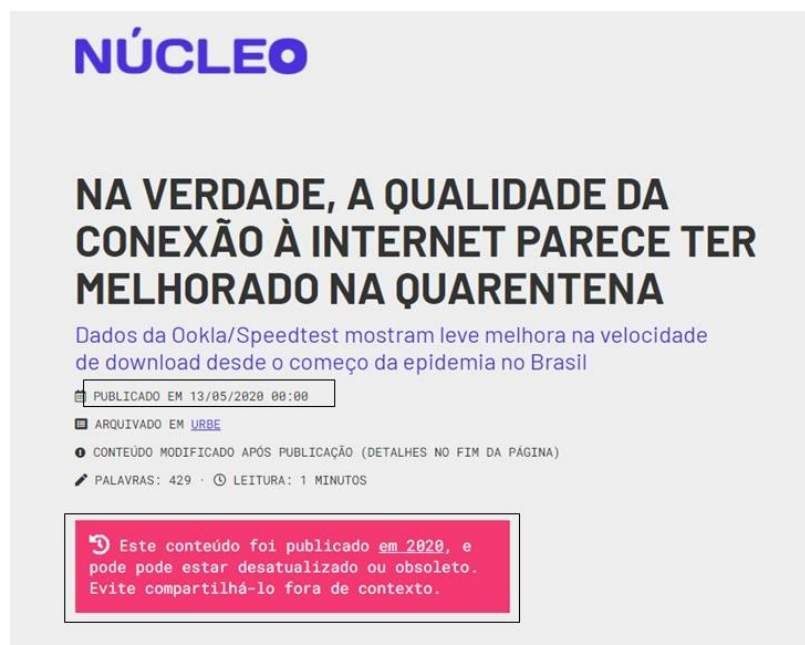
A data de publicação de um conteúdo é um elemento indispensável, pois abrange a ideia de atualidade e de uma verdade verificável, passível de atualizações constantes. Embora pareça óbvio, a data de publicação de uma notícia ou reportagem nem sempre está clara em um portal noticioso, ou então não aparece no topo da página, para a identificação rápida e imediata dos leitores.

Mais do que informar um dado elementar, a disponibilidade da data de publicação é particularmente útil para combater a desinformação, visto que uma das principais formas em que isso ocorre no ambiente digital é via recirculação de conteúdo antigo. Na pandemia de Covid-19, isso se torna ainda mais perigoso, visto que podem ser veiculados dados antigos e fora de contexto.

Na **Figura 21**, o site do *Núcleo*, especializado em Jornalismo Guiado por Dados, apresenta um aviso em destaque, ressaltando que a notícia em questão, intitulada *Na*

*verdade, a qualidade da conexão à internet parece ter melhorado na quarentena* (SPAGNUOLO, 2020), foi publicada um ano antes da data de acesso pelo leitor.

**Figura 21 – Banner que relembra data de publicação é realçado no Núcleo**



Fonte: *Núcleo Jornalismo de Dados*. Captura de tela em 18 maio 2021.

Os destaques da **Figura 21** mostram que a data original de uma publicação pode ser realçada. Alerta, ainda, para que o usuário evite repassá-la fora de contexto, evitando que ocorra a recirculação das informações de maneira distorcida. Conforme discutido no **Capítulo 2**, a inclusão de alertas para evitar o compartilhamento distorcido é uma medida efetiva de combate à desinformação.

### 5.2.2.2 Documentos

A divulgação de documentos consultados como fonte em uma notícia, via hiperlink ou mesmo no corpo do texto – ainda mais importante no caso de estudos científicos, em que a documentação nem sempre é aberta –, confere uma espécie de prova de verdade (LISBOA; BENETTI, 2015) referente ao conteúdo apurado. Implica abrir ao leitor o documento na íntegra, e não apenas determinado trecho, como geralmente ocorre.

Na reportagem a seguir, veiculada pela Agência Pública sob o título *Influenciadores digitais receberam R\$ 23 mil do governo Bolsonaro para propagandear "atendimento precoce" contra Covid-19* (FLECK; MARTINS, 2021), o acesso ao montante de dinheiro público pago aos influenciadores digitais foi obtido por meio da Lei de Acesso à Informação (LAI). De forma complementar à lista fornecida, a publicação traz o acesso, por meio de hiperlink, ao documento original que acompanha a resposta ao pedido via LAI.

**Figura 22 – Ofício acompanha reportagem da Agência Pública**

De acordo com os documentos obtidos pela Pública, a ex-EBB Flávia Viana recebeu R\$11,5 mil para ação de marketing 'Cuidados Precoce COVID-19'

No texto-guia, ou *briefing*, desta ação, os quatro influenciadores foram orientados a posar de maneiras diferentes, alguns com a máscara no rosto e álcool gel na mão, outros lavando a mão. O texto pelo qual os influenciadores deveriam se guiar para fazer seus posts dizia: "Hoje quero falar de um assunto importante, quero reforçar algumas formas de se prevenir do coronavírus. Vamos nos informar e buscar orientações em fontes confiáveis. Não vamos dar espaços para fake news. Com saúde não se brinca. Fiquem atentos! E se identificar algum sintoma como dor de cabeça, febre, tosse, cansaço, perda de olfato ou paladar, #NãoEspere, procure um médico e solicite um atendimento precoce". O texto trazia ainda a recomendação do uso da máscara e higienização das mãos com água e sabão ou álcool gel.

Em ofício que acompanha a resposta da LAI, a Secom esclareceu que, do valor total, R\$987,2 mil foram destinados à produção das peças – filmes para TV, spot para rádio, vídeos e banners para internet e peças para mídia exterior – enquanto o valor restante (R\$ 18,9 milhões) foi destinado à veiculação e divulgação do material produzido. Não há detalhamento dos gastos com ações de marketing de influência.

Uma das peças de TV veiculadas em outubro focava na hashtag "NãoEspere", presente no texto-guia entregue aos influenciadores. As ações foram pensadas no bojo da campanha anunciada pela pasta no fim de setembro de 2020 para estimular o cuidado precoce.

Fonte: Agência Pública. Captura de tela em 18 maio 2021.

Conforme é possível observar na **Figura 22**, o documento utilizado como fonte é disponibilizado para consulta na íntegra. À esquerda na imagem, a palavra ofício aparece sublinhada e, ao acessar o hiperlink, o leitor obtém a publicação (à direita na **Figura 22**). Ou seja, a Agência Pública não disponibiliza somente uma parte do documento na própria notícia, mas permite que o leitor o confira na íntegra.

### 5.2.2.3 Fontes de dados

Revelar as fontes de informação utilizadas na elaboração de uma reportagem faz parte da conduta ética do jornalista. Quando se trata de uma fonte documental cujo acesso do leitor pode ocorrer em primeira mão, a exemplo dos bancos de dados públicos, existe a possibilidade de compartilhamento, o que geralmente ocorre por meio da disponibilidade das fontes originais via hyperlink.

No texto publicado pelo *El País* Brasil e intitulado *Mortes entre caixas, frentistas e motoristas de ônibus aumentaram 60% no Brasil no auge da pandemia* (SOARES, 2021), é possível observar duas situações relacionadas ao uso e à disponibilidade de fontes de dados. A primeira, conforme pode ser observado na **Figura 23**, diz respeito à disponibilidade de hyperlink externo direcionado à fonte original.

**Figura 23 – Hiperlinks remetem às fontes de dados no jornal *El País***

The image shows a screenshot of a news article from *El País* with three hyperlinks highlighted in boxes. Arrows point from these boxes to three external data sources:

- Link 1:** [Novo Caged](#). This link points to the **PDET** (Programa de Determinação das Estatísticas do Trabalho) website, which displays a bar chart titled "Evolução do Saldo no Novo CAGED em 2020 - Com Ajustes". The chart shows monthly data for 2020, with values ranging from approximately -1,000,000 to 1,000,000.
- Link 2:** [os epidemiologistas costumam usar o conceito de "excesso de mortes"](#). This link points to the **CONASS** (Conselho Nacional de Secretários de Saúde) website, which features a section titled "Painel de análise do excesso de mortalidade por causas naturais no Brasil em 2020".
- Link 3:** [O Sivep-Gripe](#). This link points to the **DataSUS** (Ministério da Saúde) website, which displays the "SRAG 2021 - Banco de Dados de Síndrome Respiratória Aguda Grave - incluindo dados da COVID-19".

Fonte: A autora (2021). Captura de tela em 18 maio 2021.

A **Figura 23** mostra em destaque, à esquerda, os fragmentos de texto que apresentam hiperlink externo; à direita, por sua vez, estão destacados os portais que serviram de fonte documental para a reportagem. Embora a fonte original seja referenciada, uma limitação, neste caso, é a ausência dos bancos de dados primários utilizados nas análises. Ou seja, o leitor não sabe exatamente quais foram as operações feitas pelos jornalistas em sua análise de dados e posterior publicação – não necessariamente por escolha do profissional, mas muitas vezes por conta de normas internas dos jornais.

Na **Figura 24**, que apresenta visualizações de dados, as planilhas utilizadas para a produção dos gráficos estão disponíveis aos leitores, ainda que não haja clareza sobre os filtros e seleções adotados na etapa anterior aos dados disponibilizados.

**Figura 24 – Gráficos apresentam opção para download dos dados em *El País***



Fonte: A autora (2021). Captura de tela em 18 maio 2021.

Neste caso da **Figura 24**, não é possível fazer o download de toda a base de dados que serviu de fonte para a notícia; pode-se apenas fazer o download das informações utilizadas para gerar os gráficos, conforme os destaques presentes na imagem. Mesmo assim, o material permite a comparação entre os gráficos e as informações disponibilizadas na planilha.



Ainda que apresente certas limitações quanto à abertura dos dados e das fontes originais, o jornalismo digital possui considerável vantagem em relação aos meios analógicos para praticar a transparência metodológica – por exemplo, pela oferta de hiperlinks para fins de contextualização.

Como se tem visto, abrir espaço para links externos ainda não é uma prática completamente comum entre as empresas jornalísticas por questões de concorrência ou mesmo por conta de padrões de publicação. Em geral, os jornais acabam apelando para a autorreferência em notícias publicadas anteriormente, resgatando o histórico de notícias publicadas apenas dentro de seu próprio site.

#### **5.2.2.4 Graus de incerteza**

Especialmente no começo da pandemia, ainda nos primeiros meses de 2020, a disponibilidade e organização dos dados ligados à Covid-19 apresentavam lacunas e problemas, a exemplo da subnotificação de casos. A própria organização do setor da saúde é complexa no Brasil e envolve diferentes segmentos governamentais, sendo que cada instância é responsável por uma etapa da coleta e/ou disponibilização das informações.

Dessa forma, especialmente no início, jornalistas de dados precisaram buscar alternativas. Ao encontrar bancos de dados que apresentavam números sobre a pandemia, porém, uma espécie de explicação sobre as incertezas e limitações dos bancos de dados precisava acompanhar as reportagens.

Em texto intitulado *Base de dados de cartórios traz falhas que impedem calcular efeito real do coronavírus no Brasil*, Diegues, Faria e Takahashi (2020) revelaram que a base de dados dos cartórios, até então utilizada para contabilizar as mortes por Covid-19 na falta de outras opções, continha informações desatualizadas não só sobre 2020, mas também em relação a anos anteriores. Assim, tais lacunas prejudicavam comparações.

**Figura 25 – Limitações em base de dados foram levantadas pela *Folha***

ANÁLISE · DELTAFOLHA · CORONAVÍRUS

## Base de dados de cartórios traz falhas que impedem calcular efeito real do coronavírus no Brasil

Plataforma é atualizada de maneira desigual em diferentes cidades pelo país

Leonardo Diegues  
Flávia Faria  
Fábio Takahashi

**SÃO PAULO** Utilizada por pesquisadores e pela imprensa para tentar [medir o impacto da Covid-19](#), a base de dados que mostra o número de mortes no país tem defasagem de informações que praticamente impede a análise do avanço da doença, mesmo em capitais e cidades grandes.

Um dos principais problemas da plataforma é que há dados desatualizados não só das últimas semanas, mas até de anos anteriores, o que prejudica a comparação com 2020.

Mantido pela associação de cartórios, a Arpen, o sistema chamado Portal de Transparência do Registro Civil mostra o total de registros de óbitos no país. Em tese, é uma base com a qual se poderia estimar o “excedente de mortes” causado pelo novo coronavírus.

Nesse cálculo, considera-se o número de óbitos em anos anteriores e se compara com 2020. O que ultrapassar o padrão é, provavelmente, decorrente do vírus, o elemento novo na equação. Ou seja, poderia se estimar o impacto da doença, mesmo com a carência de testes no Brasil.

A **Folha** analisou o volume de mortes em cada uma das 518 cidades presentes na base, com dados coletados até a segunda (11). Foi verificada uma variação tão grande de mortes, nos meses de 2020 em relação aos mesmos períodos de 2019, que o cálculo para o impacto da Covid-19 ficou prejudicado.

Consideradas as 27 capitais, 6 tiveram variação no volume de óbitos acima dos 20% já em janeiro e fevereiro, quando a doença ainda não havia se espalhado pelo país.

Porto Velho chegou a ver aumento de 1.754% (de 11 para 204) em fevereiro de 2020 em relação a 2019.

O cálculo do impacto da doença seria mais seguro se o número de mortes em janeiro e fevereiro deste ano fosse próximo ao de 2019; então, o eventual crescimento em março e abril poderia ser explicado pela Covid-19.

Fonte: *Folha de S.Paulo*. Captura de tela em 18 maio 2021.

Torna-se interessante observar, na **Figura 25** e ao longo da reportagem, que os jornalistas não deixaram de utilizar o banco de dados dos cartórios para cobrir a pandemia. Contudo, há o reconhecimento das imperfeições de tal levantamento e, por isso, o alerta para que suas informações sejam utilizadas com cuidado e parcimônia. Em mais de um ponto do texto, essas limitações foram comunicadas ao leitor:

A *Folha* analisou o volume de mortes em cada uma das 518 cidades presentes na base, com dados coletados até a segunda (11). Foi verificada uma variação tão grande de mortes, nos meses de 2020 em relação aos mesmos períodos de 2019, que o cálculo para o impacto da Covid-19 ficou prejudicado (DIEGUES; FARIA; TAKAHASHI, 2020, s.p.)

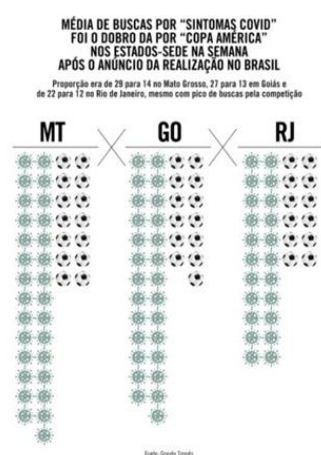
Ao longo da reportagem, os jornalistas explicaram quais cálculos puderam efetuar e quais não puderam ser realizados a partir da posse dos dados de cartório. Nota-se, ainda, que o emprego do banco de dados revela a desigualdade na disponibilidade de informações no Brasil – em alguns municípios, especialmente nos maiores, há detalhamento no número de mortes e suas causas. Em municípios menores, porém, as estatísticas costumavam estar incompletas e por isso prejudicam o monitoramento e qualquer

comparação. De qualquer forma, o não uso de bancos de dados públicos não parecia ser a solução diante da existência de um problema complexo de saúde pública; indicar suas limitações e trabalhar com o que se tem, com os devidos cuidados e ressalvas, foi a saída encontrada por alguns jornais.

### 5.2.2.5 Marcas de apuração

As marcas de apuração referem-se aos registros jornalísticos que normalmente mantêm-se nos bastidores e não são expostos nos textos, mas que ajudam o leitor a ter um panorama sobre as dificuldades do trabalho de apuração. Difere-se da metodologia empregada porque é mais pontual e sutil; indica, por exemplo, quantas vezes o jornalista ligou para uma mesma fonte ou entrou em contato com uma assessoria, quantos bancos de dados precisou percorrer para encontrar um tópico ou valor de análise, além de erros encontrados nos documentos originais, entre outros problemas. A clareza na data de consulta a uma fonte pessoal ou documento igualmente integra a lista de possibilidades das marcas de apuração, como mostra a imagem a seguir.

**Figura 26 – Piauí registra o dia e horário da análise**



Entre o dia 1 de junho, quando o presidente Jair Bolsonaro confirmou que o Brasil sediará a Copa América, e o dia 8, data marcada pelos jogadores da Seleção para se pronunciarem oficialmente sobre sua insatisfação com o torneio, o campeonato foi uma das principais pautas do noticiário nacional. No dia 4, às 23 horas, entre o fim de Brasil x Equador e a entrevista do volante Casemiro, a busca por "Copa América" explodiu e atingiu os 100 pontos, patamar mais alto da escala do Google. Mesmo assim, nos três estados-sede, a média de buscas por "sintomas covid" foi duas vezes maior que a por "Copa América" no período. A exceção é o Distrito Federal, onde a diferença foi pequena (17 a 13), mas também desfavorável à Copa.

Dados verificados no dia 8 de junho às 13h.



Na madrugada do dia 7 de junho, a busca por "sintomas covid" no Mato Grosso chegou a 90 pontos, 13 vezes maior que o interesse por Copa América registrado naquele momento, de 7 pontos. É a maior disparidade de procura entre os termos nos estados-sede naquela semana. Mesmo durante Brasil x Equador, quando a busca pelo torneio atingiu o ponto máximo da escala (100 pontos) nas quatro sedes, a diferença era bem menor: a busca pelo campeonato era cinco vezes maior que a por sintomas (100 para 18 pontos). Após a confirmação de que os atletas jogariam a Copa, ocorrida no fim da manhã daquela segunda-feira, as buscas por "Copa América" arrefeceram. Durante o jogo contra o Paraguai pelas Eliminatórias, no dia seguinte, a procura ate aumentou, mas não passou dos 40 pontos na escala.

Dados verificados no dia 9 de junho às 14 horas.

Fonte: Revista Piauí. Captura de tela em 21 jul. 2021.

Na notícia intitulada *O 7×1 da Covid*, Guimarães e Buono (2021) comparam as pesquisas realizadas no Brasil, no mecanismo de buscas do *Google* – o *Google Trends* –, sobre a Copa América e sobre os sintomas da Covid-19. Após explicar as expressões utilizadas para reunir os dados, as autoras comunicam o dia e o horário em que os dados foram consultados, conforme os destaques na **Figura 26**. Embora essa ação possa aproximar-se da reprodutibilidade do método, não o é, pois não apresenta detalhadamente o passo a passo metodológico adotado; trata-se de uma das marcas do processo de reportagem.

### 5.2.2.6 Metodologia

A metodologia empregada em uma reportagem de Jornalismo Guiado por Dados costuma aparecer ao final do texto, em formato de nota metodológica. Trata-se de um trecho que busca traduzir uma série de procedimentos feitos pelos repórteres para assegurar a clareza no recorte de seleção e coleta de dados, além de servir para justificar escolhas, como a adoção de certas fontes e procedimentos.

O método do JGD apresenta inspiração científica e, por isso, normalmente parte da detecção de um problema ou de um teste de hipóteses – ainda que esse detalhamento em geral não seja comunicado. A metodologia costuma ser apresentada na forma de narrativa e por vezes se utiliza de hiperlink, mas nem sempre envolve a abertura de código de programação.

Em reportagem intitulada *Eventos de Bolsonaro durante a pandemia* (CONCONI *et al.*, 2021), que calcula a quantidade de aglomerações provocadas pelo presidente Jair Bolsonaro na Covid-19, os jornalistas de *O Estado de S. Paulo* utilizaram uma nota metodológica para indicar as fontes e os procedimentos adotados.

**Figura 27 – Metodologia é esclarecida pelo *Estadão***



Fonte: *O Estado de S. Paulo*. Captura de tela em 18 jul. 2021.

Conforme é possível observar em destaque na **Figura 27**, a metodologia oferece a descrição dos procedimentos e a justificativa de escolhas, mas não informa se houve a utilização de código de programação, ou mesmo como foi o passo a passo que resultou na análise apresentada. Neste caso aqui demonstrado, a nota metodológica está situada ao final da reportagem.

Exemplo semelhante é detectado no *Núcleo*. Em geral, uma nota metodológica detalhada é oferecida ao final dos textos publicados nesse veículo. Na publicação intitulada *Ministério da Saúde camuflou dezenas de tweets sobre 'tratamento precoce'* (GELAPE; MARTINS, 2021), a seção *Como fizemos isso* traduz o passo a passo empregado para obter os resultados de uma análise.

**Figura 28 – Núcleo usa *Como fizemos isso* para detalhar método**

**COMO FIZEMOS ISSO**

Para contabilizar o número de tweets por gestão no Ministério, dividimos o período da pandemia pelo tempo de cada ministro à frente da Saúde, considerando a data de demissão de cada ministro como término de um período, e o dia subsequente como início da gestão seguinte.

Para identificar menções à COVID, utilizamos as seguintes palavras-chave: covid, coronavirus, pandemia, novocoronavirus, respiradores, sars-cov2, coronavac, sinovac, astrazeneca, pfizer, butantan, ifa, segunda dose, mascara, ventilador, entrada de estrangeiro, eficacia, oxford — além de uma combinação de classificação manual e por palavras-chave de tweets que mencionavam vacina.

As publicações sobre a vacinação contra o coronavirus foram todas as que mencionavam vacina e não atendiam aos seguintes critérios:

- Tweets até 31/07/2020, pois as vacinas contra covid eram tema pouco abordado, e o foco era em outras campanhas;
- Tweets com as seguintes palavras-chave: gripe, influenza, sarampo, informatizacaosus, conectesus, todas as idades, curso, poliometite, 15 anos, caes e gatos, crianca, diad, dia d, paralisia infantil, tetano, rubeola, caderneta de vacinacao, novembroazul, adolescente, adolescencia, cartao de vacina, golpe, celular.

Existe a possibilidade de falsos negativos (tweets marcados como não mencionando a covid, quando se referiam a ela). Contudo, é improvável que eles afetem a comparação entre gestões: é necessário que existam 326 publicações falso-negativas dentre as publicações da gestão Pazuello para que ele alcance o patamar da gestão Queiroga (64,9%) em menções a Covid (considerando que não existam falsos negativos entre esta última). Ou seja, isto implicaria em um erro de classificação de 23% dos tweets categorizados como não-covid nesta gestão.

Fonte: *Núcleo Jornalismo de Dados*. Captura de tela em 18 out. 2021.

A extensa nota metodológica apresentada pelo *Núcleo* é seguida pela disponibilização dos bancos de dados originais utilizados na análise, permitindo que os leitores acessem as informações completas.

### 5.2.2.7 Reprodutibilidade e código aberto

Inspirada pela vertente científica do Jornalismo de Precisão e com uma comunidade de profissionais mobilizados pelo *ethos* de colaboração presente na área da computação, o Jornalismo Guiado por Dados tende a abrir, em suas reportagens, o código-fonte utilizado para construir projetos e análises. Significa que as linhas de código que possibilitam a criação de algoritmos estão documentadas e são colocadas à disposição para consulta.

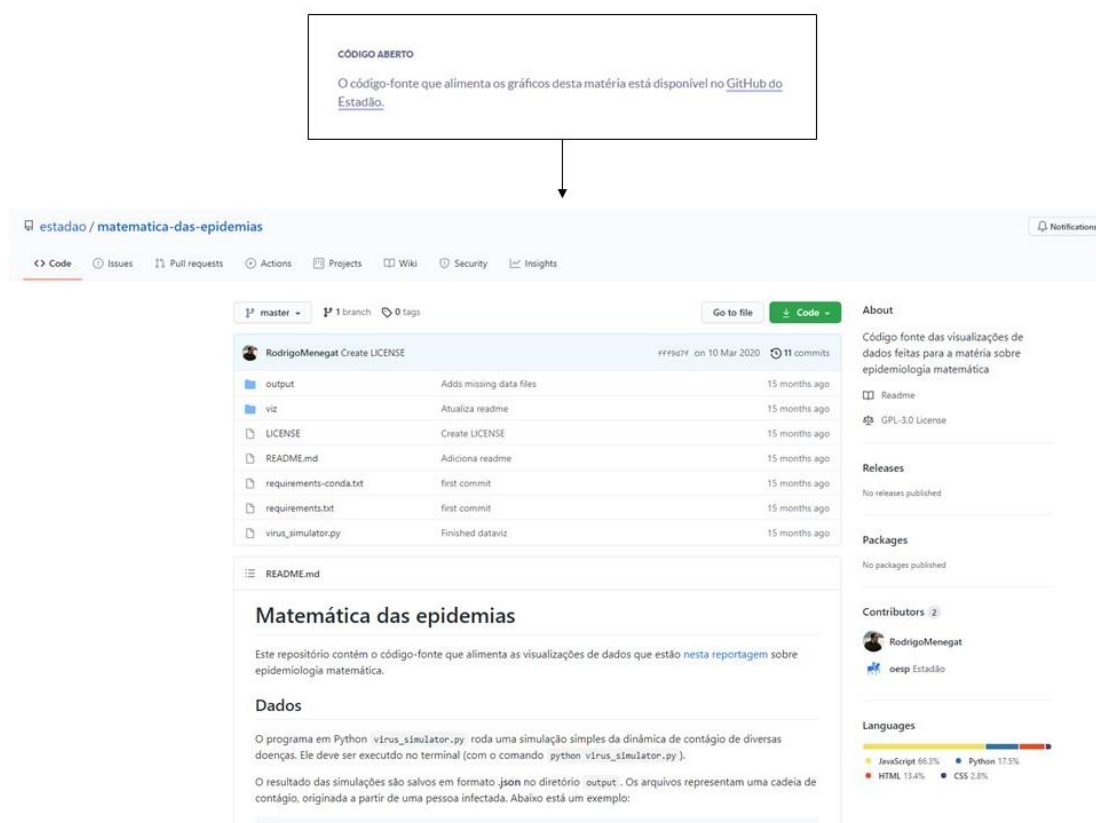
Além da possibilidade de descrever o método de trabalho apenas em formato narrativo, conforme demonstrado no tópico 5.2.2.6, um dos diferenciais do JGD é justamente organizar seus métodos em formato de código e, assim, permitir a reprodutibilidade e a verificação das informações. Dessa forma, outros jornalistas e redações estão aptos a verificar o conteúdo produzido, como em uma espécie de revisão entre pares, e eventualmente reaproveitar o código; também permite que leitores com conhecimento técnico avaliem a construção de uma narrativa de dados e a pertinência de suas informações. A disponibilidade de código-fonte ocorre tanto em notícias quanto em projetos jornalísticos de maior fôlego. Como veremos adiante, esse tópico dialoga com a ideia de ciência aberta.

No exemplo a seguir, a reportagem *Como a matemática pode ajudar a entender (e combater) epidemias*, publicada pelo jornal *O Estado de S.Paulo* (MENEGAT, 2020), oferece linhas de código no *GitHub*<sup>124</sup>, uma espécie de repositório público costumeiramente utilizado no Jornalismo Guiado por Dados e também em comunidades que valorizam o uso de linguagens de programação de código aberto, como *Python* e *R*.

---

<sup>124</sup> O *GitHub* é uma plataforma de hospedagem para código-fonte e arquivos. Costuma ser utilizada por programadores, pesquisadores, jornalistas e demais usuários cadastrados na plataforma, que contribuem em projetos privados e/ou *open source*.

**Figura 29 – Reportagem do *Estadão* divulga código-fonte no *GitHub***



Fonte: *O Estado de S. Paulo*. Captura de tela em 1 jun. 2021.

A parte superior da **Figura 29** mostra trecho extraído ao final da reportagem citada, indicando a possibilidade de acesso ao código-fonte. O hiperlink leva a uma página no *GitHub*. Esta página, por sua vez, apresenta um arquivo que documenta o passo a passo adotado, além de arquivos e pastas que compreendem o código-fonte na íntegra.

Em outro exemplo, no projeto de reportagem *No Epicentro*<sup>125</sup>, desenvolvido pela *Agência Lupa*, vinculada à *Piauí*, e um dos vencedores do prêmio *Sigma Awards*, o código-fonte também é disponibilizado no *GitHub*<sup>126</sup> em diferentes pastas, diferenciando as etapas que pressupõem sua elaboração e seu uso posterior por parte dos leitores.

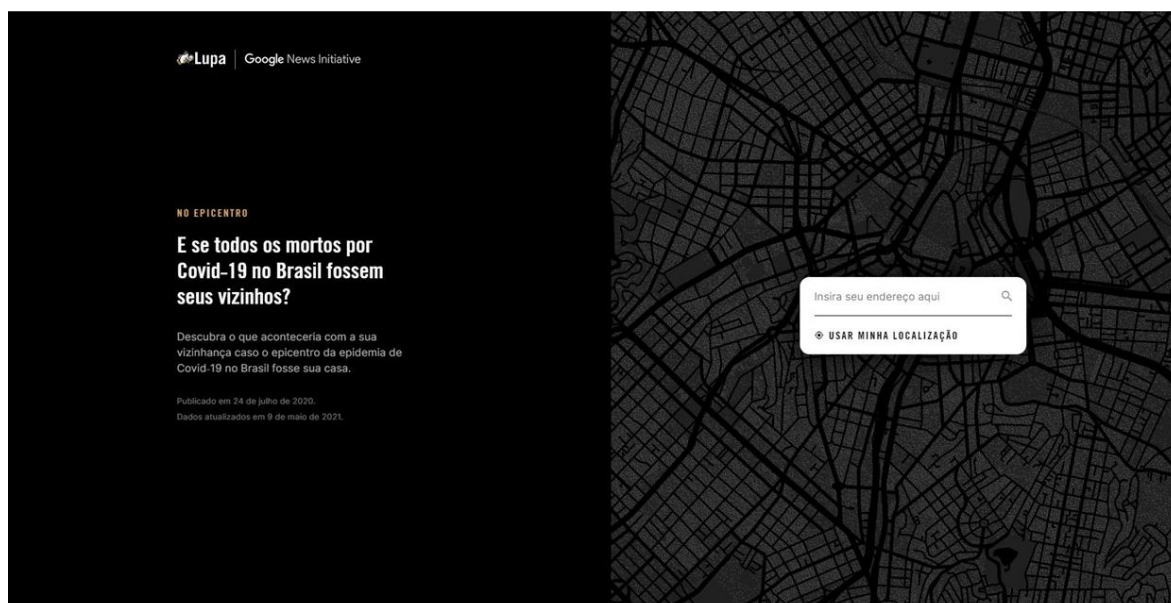
Em primeiro lugar, ao acessar a página inicial do projeto, o leitor é convidado a inserir seus dados geográficos para que compreenda o tamanho da tragédia da Covid-19 ao seu redor, conforme mostra a **Figura 30**.

<sup>125</sup> Disponível em: <https://piaui.folha.uol.com.br/lupa/epicentro/>

<sup>126</sup> Disponível em: <https://github.com/noepicentro/>



**Figura 30 – Projeto *No Epicentro* apresenta mapa para inserção de dados**



Fonte: A autora (2021). Captura de tela em 1 jun. 2021.

A seguir, na **Figura 31**, é possível observar diferentes pastas que reúnem a totalidade dos códigos deste projeto. Como dissemos, quando o leitor acessa o site em questão, insere o código postal de sua cidade para observar, no mapa, o raio de mortos em decorrência da Covid-19. Por se tratar de um processo complexo e distribuído em etapas, existem códigos de programação presentes em diferentes partes do desenvolvimento do site e, por isso, estão separadas em pastas no *GitHub* (*front* e *back*, por exemplo, que dialogam com etapas visíveis e não visíveis à audiência, respectivamente, e que podem ser vistos na imagem a seguir).

**Figura 31 – Projeto *No Epicentro*, da Agência Lupa, divulga código-fonte**

**Como funciona a simulação**

O raio de mortes exibido ao seu redor depende da sua localização e da quantidade de pessoas que vivem próximas a você.

Para definir o tamanho desse raio, o aplicativo usa a população que vive dentro de cada setor censitário do país. Estes setores são pequenos blocos territoriais definidos pelo IBGE.

Alguns setores representam prédios de uma cidade, Outros, quadras ou bairros. No último censo, em 2010, o IBGE contou quantos moradores havia em cada uma dessas áreas. Assim, cada ponto no mapa representa uma pessoa que vive naquela região.

O aplicativo soma a população dos setores censitários ao seu redor até atingir um número próximo ao total de mortes por Covid-19 no país.

Usamos número de óbitos divulgado pelo **Brasil.io**, grupo de voluntários que tem reunido dados dos boletins epidemiológicos estaduais desde o início da pandemia.

A **metodologia** completa envolve desenhar raios crescentes, calcular interseções e fazer estimativas estatísticas.

O código-fonte está disponível no [GitHub](#).

The screenshot shows the GitHub repository page for 'noepicentro'. It lists three repositories: 'front' (JavaScript, MIT license, 15 stars, updated 23 Apr), 'back' (Python, 14 stars, updated 29 Mar), and 'coldfoot.github.io' (JavaScript, 1 star, updated 25 Jul 2020). The page also shows a search bar, navigation links, and a footer with copyright information for 2021 GitHub, Inc.

Fonte: *No Epicentro*, da Agência Lupa. Captura de tela em 1 jun. 2021.

No exemplo em questão, a linguagem de programação *Python* foi a mais utilizada. Conforme Belisário *et al.* (2020), linguagens de código aberto, como *Python* e *R*, costumam ser empregadas não apenas entre cientistas da computação – mas também são acionadas por jornalistas e pesquisadores. Ambas são de uso livre e gratuito, bem como contam com milhares de bibliotecas ou pacotes para expandir suas funcionalidades, como estruturação de dados e visualização, criação de modelos estatísticos e inserção de rotinas de automação.

Na próxima seção, este projeto e outras reportagens consultadas ao longo do capítulo serão sistematizadas no **Quadro 3**. O resumo dos elementos de transparência, por sua vez, poderá ser conferido no **Quadro 4**.

### 5.2.3 Resumo dos resultados

A formação do corpus empírico da tese, em relação aos materiais informativos que foram utilizados para compor os níveis organizacional e operacional, está estruturada resumidamente a seguir, de acordo com a ordem em que o conteúdo aparece ao longo desta pesquisa. A listagem apresenta somente notícias, reportagens e projetos, visto que o endereço de consulta aos sites dos jornais foi indicado anteriormente.

**Quadro 3 – Lista dos textos informativos que compõem o corpus**

<b>Título</b>	<b>Veículo</b>	<b>Data</b>	<b>URL</b>
TrateCov indicava cloroquina e outros remédios sem eficácia a quase todos pacientes	<i>Jota</i>	22/01/21	<a href="https://www.jota.info/dados/tratecov-indicava-cloroquina-e-outras-remedios-sem-eficacia-a-quase-todos-pacientes-22012021">https://www.jota.info/dados/tratecov-indicava-cloroquina-e-outras-remedios-sem-eficacia-a-quase-todos-pacientes-22012021</a>
Cidades mais desenvolvidas do país vacinam mais rápido contra a Covid	<i>Folha de S.Paulo</i>	04/07/21	<a href="https://www1.folha.uol.com.br/equilibrioe-saude/2021/07/cidades-mais-desenvolvidas-do-pais-vacinam-mais-rapido-contra-a-covid.shtml">https://www1.folha.uol.com.br/equilibrioe-saude/2021/07/cidades-mais-desenvolvidas-do-pais-vacinam-mais-rapido-contra-a-covid.shtml</a>
Por que tanta gente ignora regras no pior momento da pandemia	<i>Nexo</i>	29/03/21	<a href="https://www.nexojornal.com.br/expresso/2021/03/29/Por-que-tanta-gente-ignora-regras-no-pior-momento-da-pandemia">https://www.nexojornal.com.br/expresso/2021/03/29/Por-que-tanta-gente-ignora-regras-no-pior-momento-da-pandemia</a>
Registros indicam que milhares no Brasil tomaram vacina vencida contra Covid; veja se você é um deles	<i>Folha de S.Paulo</i>	02/07/21	<a href="https://www1.folha.uol.com.br/equilibrioe-saude/2021/07/milhares-no-brasil-tomaram-vacina-vencida-contra-covid-veja-se-voce-e-um-deles.shtml">https://www1.folha.uol.com.br/equilibrioe-saude/2021/07/milhares-no-brasil-tomaram-vacina-vencida-contra-covid-veja-se-voce-e-um-deles.shtml</a>

Na verdade, a qualidade da conexão à internet parece ter melhorado na quarentena	<i>Núcleo</i>	13/05/20	<a href="https://www.nucleo.jor.br/reportagem/2020-05-13-qualidade-internet-conexao-quarentena/">https://www.nucleo.jor.br/reportagem/2020-05-13-qualidade-internet-conexao-quarentena/</a>
Influenciadores digitais receberam R\$ 23 mil do governo Bolsonaro para propagandear “atendimento precoce”	<i>Agência Pública</i>	31/03/21	<a href="https://apublica.org/2021/03/influenciadores-digitais-receberam-r-23-mil-do-governo-bolsonaro-para-propagandear-atendimento-precoce-contr-a-covid-19/">https://apublica.org/2021/03/influenciadores-digitais-receberam-r-23-mil-do-governo-bolsonaro-para-propagandear-atendimento-precoce-contr-a-covid-19/</a>
Mortes entre caixas, frentistas e motoristas de ônibus aumentaram 60% no Brasil no auge da pandemia	<i>El País (Brasil)</i>	05/04/21	<a href="https://brasil.elpais.com/brasil/2021-04-05/caixas-frentistas-e-motoristas-de-onibus-registram-60-a-mais-de-mortes-no-brasil-em-meio-ao-auge-da-pandemia.html">https://brasil.elpais.com/brasil/2021-04-05/caixas-frentistas-e-motoristas-de-onibus-registram-60-a-mais-de-mortes-no-brasil-em-meio-ao-auge-da-pandemia.html</a>
Base de dados de cartórios traz falhas que impedem calcular efeito real do coronavírus no Brasil	<i>Folha de S. Paulo</i>	13/05/20	<a href="https://www1.folha.uol.com.br/equilibrio-e-saude/2020/05/base-de-dados-de-cartorios-traz-falhas-que-impedem-calcular-efeito-real-do-coronavirus-no-brasil.shtml">https://www1.folha.uol.com.br/equilibrio-e-saude/2020/05/base-de-dados-de-cartorios-traz-falhas-que-impedem-calcular-efeito-real-do-coronavirus-no-brasil.shtml</a>
O 7x1 da Covid	<i>Piauí</i>	21/06/21	<a href="https://piaui.folha.uol.com.br/o-7x1-da-covid/">https://piaui.folha.uol.com.br/o-7x1-da-covid/</a>
Eventos de Bolsonaro durante a pandemia	<i>O Estado de S. Paulo</i>	06/06/21	<a href="https://arte.estadao.com.br/politica/2021/06/deslocamentos-jair-bolsonaro-pandemia/">https://arte.estadao.com.br/politica/2021/06/deslocamentos-jair-bolsonaro-pandemia/</a>
Ministério da Saúde camuflou dezenas de tweets sobre 'tratamento precoce'	<i>Núcleo</i>	02/08/21	<a href="https://www.nucleo.jor.br/reportagem/2021-08-02-ms-tweets-tratamento-covid/">https://www.nucleo.jor.br/reportagem/2021-08-02-ms-tweets-tratamento-covid/</a>
Como a matemática pode ajudar a entender (e combater) epidemias	<i>O Estado de S. Paulo</i>	13/03/20	<a href="https://www.estadao.com.br/infograficos/saude,como-a-matematica-pode-ajudar-a-entender-e-combater-epidemias,1082298">https://www.estadao.com.br/infograficos/saude,como-a-matematica-pode-ajudar-a-entender-e-combater-epidemias,1082298</a>
No Epicentro	<i>Agência Lupa/Piauí</i>	24/07/20	<a href="https://piaui.folha.uol.com.br/lupa/epicentro/">https://piaui.folha.uol.com.br/lupa/epicentro/</a>

Fonte: A autora (2021).

Como é possível visualizar no **Quadro 3**, optou-se pela variedade de veículos de comunicação e de data de publicação dos textos informativos, conforme preconizado pela metodologia adotada. Esse conteúdo representa a sistematização de uma parcela das publicações lidas a partir do início da pandemia de Covid-19 e a detecção do primeiro caso no Brasil, ou seja, a partir de março de 2020. Por questões de viabilidade de coleta, engloba o período até agosto de 2021.

Reunidos, os níveis organizacional e operacional apresentam **15 elementos de transparência** detectados em análise empírica realizada durante a pandemia de Covid-19 em observância a veículos jornalísticos brasileiros. São eles, por ordem alfabética:

#### Quadro 4 – Os elementos de transparência no Jornalismo Guiado por Dados

Elemento	Descrição
Assinatura, perfil e expertise	Nome do profissional que assina a reportagem, bem como breve descrição biográfica e lista de notícias escritas pelo jornalista anteriormente.
Contato	Múltiplas formas de contato com a empresa jornalística. Pressupõe interação entre leitor e redação.
Correção de erros	Indicação de alteração no conteúdo jornalístico que apresentou erro e posterior correção, além de apresentação de política de correção de erros.
Data de publicação	Demonstração da data em que a reportagem foi produzida, bem como estratégias de contexto e alerta para combate à desinformação.
Documentos	Inserção de arquivos empregados como fonte, revelando não apenas um trecho de seu conteúdo, mas o material na íntegra.
Fontes de dados	Apresentação dos bancos de dados utilizados como fonte, de preferência com dados primários e hiperlink para acesso externo.
Fontes de financiamento	Indicação detalhada de receitas e despesas, bem como apresentação nominal de apoiadores e patrocinadores.
Graus de incerteza	Alerta em relação às limitações e ao contexto dos bancos de dados utilizados como fonte.
Manual da redação e princípios editoriais	Presença de manual da redação e/ou lista de princípios editoriais e de conduta a serem seguidos e praticados pela empresa.
Marcas de apuração	Indicação de aspectos inerentes ao processo de apuração da reportagem.
Metodologia	Apresentação de nota metodológica que indica escolhas editoriais para uma reportagem, bem como recorte e seleção de dados empregados.

Política de diversidade	Existência de normas que demonstrem a preocupação do jornal com a diversidade de temas a serem discutidos e profissionais que abordarão esses temas.
Política de privacidade	Apresentação do que é feito com os rastros digitais dos leitores que consomem notícias em um determinado site.
Reprodutibilidade e código aberto	Disponibilidade de código aberto para verificação e reprodução das análises de dados.
Tipo de conteúdo	Clareza na separação entre conteúdo informativo e opinativo.

Fonte: A autora (2021).

Pontuo, mais uma vez, que esses foram os elementos encontrados na cobertura realizada pelo Jornalismo Guiado por Dados em 2020 e 2021 sobre a pandemia de Covid-19 a partir da metodologia proposta para a tese, que é uma combinação entre os métodos de estudo de casos múltiplos (STAKE, 2006) e estudo de caso como ilustração (MACHADO; PALACIOS, 2007), de caráter qualitativo e exploratório.

Como foi possível notar neste percurso metodológico, as empresas jornalísticas aqui mencionadas não atendem por completo a todos os elementos de transparência. Embora alguns jornais apareçam mais de uma vez entre os tópicos, não há julgamento de valor sobre o quão transparente é esta empresa em questão. Por isso, essa tese não se propõe a discutir qual veículo é mais ou menos transparente. Trata-se, sobretudo, de um trabalho de sistematização e embasamento para novas pesquisas. Entendemos que a cobertura de Covid-19 provocou mudanças no jornalismo como um todo, mas especialmente em segmentos como o Jornalismo Guiado por Dados, e por isso tais alterações nos processos de reportagem e publicação poderão inspirar mudanças a serem implementadas mesmo após a pandemia.

Para inspirar novos estudos, apresento, a seguir, o **Protocolo de Transparência no Jornalismo Guiado por Dados**, composto pelos elementos de transparência que aqui detectamos, e a proposta de **Método para o Jornalismo Guiado por Dados com foco na transparência e reprodutibilidade** a partir do que foi estudado nesta tese, e como forma de combater a desinformação. Entendemos que este conhecimento é basilar e forma as principais contribuições desta tese para avanços no campo do jornalismo.

## 6 CONTRIBUIÇÕES

Nesta investigação, o estudo de casos múltiplos e o estudo de caso como ilustração permitem descrever características de transparência presentes no conteúdo informativo desenvolvido a partir de técnicas do Jornalismo Guiado por Dados. Após a sistematização e descrição dos elementos, proponho as seguintes contribuições: o **Protocolo de Transparência no Jornalismo Guiado por Dados**, destinado à análise de produtos jornalísticos já publicados, e a proposta de **Método para o Jornalismo Guiado por Dados com foco na transparência e na reprodutibilidade** para implementação em iniciativas jornalísticas orientadas por dados.

### 6.1 Protocolo de Transparência no Jornalismo Guiado por Dados

O Protocolo de Transparência no Jornalismo Guiado por Dados, que proponho aqui, é inspirado em estudos anteriores (SILVA; MAIA, 2011; TRÄSEL; LISBOA; REIS, 2018) e tem como base a teoria associada à transparência no jornalismo, desenvolvida por diversos autores (KARLSSON, 2010; KOLISKA; CHADHA, 2018; PLAISANCE, 2007; KOVACH; ROSENSTIEL, 2014; LISBOA; FRANZONI, 2018; MEIER, 2009; WEINBERGER, 2009). Desdobra-se dos elementos de transparência encontrados em notícias, reportagens e projetos conduzidos com técnicas de Jornalismo Guiado por Dados durante a cobertura de Covid-19 e já examinados nesta tese.

Como amplamente discutido no capítulo anterior, entendemos que a transparência está ancorada em dois princípios norteadores, como **operadora de uma prática ética e recurso discursivo de legitimação**. Seu caráter pragmático, por sua vez, pressupõe a existência de dois níveis: **organizacional**, ligado ao veículo de comunicação onde o conteúdo foi publicado, e **operacional**, em que a atenção está voltada ao passo a passo da reportagem.

Juntos, os dois níveis apresentam 15 elementos. O primeiro, organizacional, possui os seguintes parâmetros, por ordem alfabética: **assinatura, perfil e expertise; contato; correção de erros; fontes de financiamento; manual da redação e princípios editoriais; política de diversidade; política de privacidade; e tipo de conteúdo**. Quanto ao segundo nível de análise, **operacional**, os elementos de transparência a serem

observados são os seguintes: **data de publicação; documentos; fontes de dados; graus de incerteza; marcas de apuração; metodologia; e reprodutibilidade e código aberto.** O **Quadro 4**, apresentado na seção anterior, permite observar cada um desses parâmetros em detalhes.

Este protocolo tem como objetivo auxiliar pesquisadores e pessoas interessadas em avaliar os níveis de transparência de uma empresa jornalística e das reportagens publicadas. Se o pesquisador considerar pertinente, pode implementar um sistema de pontuação conforme o cumprimento dos critérios detalhados a seguir, que serão apresentados em ordem alfabética. Para executar a primeira etapa, deve-se coletar as informações disponíveis na página inicial do site do jornal estudado. O segundo nível compreende a avaliação das reportagens de Jornalismo Guiado por Dados. Neste caso, o cumprimento dos indicadores – com ênfase nas fontes e nos métodos – é avaliado a partir da leitura do material que se pretende analisar. Embora este protocolo seja destinado à análise de reportagens guiadas por dados, pode ser adaptado para outros contextos e produções jornalísticas.

A seguir, apresentamos cada elemento e uma relação de indicadores, em forma de perguntas, que guiam a avaliação da transparência daquele tópico. Novos indicadores podem ser acrescentados a partir de características próprios do objeto.

## **1) Nível organizacional**

### **Assinatura, perfil e expertise**

- O texto está assinado?
- Permite acessar um perfil com informações biográficas do autor?
- Permite acessar outros trabalhos do autor?
- Há indicação do nível de conhecimento do autor sobre o tema?

### **Contato**

- O site do jornal exibe uma lista de contatos de fácil localização e acesso?
- Há modos variados de contato (e-mail, aplicativos de mensagem, telefone etc.)?
- Há ombudsman ou ouvidor?
- Há uma editoria de audiência?



- Há uma seção de comentários do leitor? Essa seção é mediada pelo veículo ou apenas remete a uma rede social digital?

### **Correção de erros**

- Há uma política clara sobre a correção de erros?
- A descrição dessa política editorial é de fácil ou difícil acesso aos leitores?
- Há uma seção específica para as retificações?
- Há indicação, no texto original, de que houve algum tipo de equívoco?
- A correção de erros ocorre no texto original de forma clara (indicando, por exemplo, data e horário em que foi feita)?
- Há um botão que indica a possibilidade de o leitor comunicar erros?
- Existe preocupação do veículo em fornecer espaço adequado à comunicação de erros, de modo que a retificação receba um espaço significativo para ganhar visibilidade e de fato funcionar como uma correção?

### **Fontes de financiamento**

- O site do jornal explicita as fontes de financiamento ou patrocinadores?
- Há uma seção de prestação de contas com a indicação de receitas e despesas?
- É possível acessar o balanço patrimonial detalhado?
- A receita gerada pelo pagamento da audiência é monitorada e disponibilizada?
- Quando integra uma corporação ou conglomerado, o jornal permite o amplo acesso do leitor, indicando as demais empresas que o compõem?

### **Manual da redação e princípios editoriais**

- O veículo apresenta claramente seus princípios editoriais?
- A transparência é um dos princípios citados pelo jornal?
- Há um manual da redação disponível aos leitores?
- Neste manual, há preocupação em explicitar a conduta ética do jornalista e suas relações com as fontes?

**Política de diversidade**

- O jornal apresenta uma política de diversidade de jornalistas, temas e fontes?
- Há uma editoria de diversidade, de modo a tratar dessa política?
- Há subsídios ou ferramentas para o leitor monitorar o cumprimento da política de diversidade?

**Política de privacidade**

- O leitor tem acesso a informações claras sobre o tratamento de seus dados pessoais ao acessar o site do jornal?
- O veículo detalha o tipo de informação coletada pelo site?
- O jornal apresenta os dados sobre o perfil da audiência?

**Tipo de conteúdo**

- Há marcadores que permitam identificar se um conteúdo é informativo ou opinativo, diferenciando-os de conteúdo pago?
  - No caso do conteúdo opinativo, existe clareza de que o texto é uma análise feita por um especialista?
  - Há marcadores de texto indicando conteúdo editorial?
  - O veículo investe em projetos ou iniciativas para o letramento do leitor, de modo que ele possa identificar diferenças entre informação e opinião?
  - O veículo investe em projetos ou iniciativas para o letramento do leitor, de modo que compreenda o processo de produção de notícias e reportagens?

**2) Nível operacional****Data de publicação**

- A data de publicação do texto está visível?
- A data de eventuais atualizações ou correções está visível, havendo diferenciação em relação à data original da publicação?
- Quando se trata de um conteúdo que entrou em recirculação, há destaque visual referente à data original, alertando para que o leitor observe este ponto?

**Documentos**

- Os documentos utilizados como fonte estão claramente identificados?
- Há acesso à íntegra desses documentos, com possibilidade de consulta aos materiais originais?
- Os documentos são disponibilizados junto ao corpo do texto, para pleno acesso do leitor, ou por meio de hiperlink, direcionando a outra página?

**Fontes de dados**

- As fontes de dados estão claramente identificadas?
- O texto oferece acesso à fonte original, como as planilhas utilizadas, incluindo indicação de hiperlink externo?
- Há materiais complementares à fonte original (a exemplo dos dicionários de dados), oferecendo ao leitor o contexto de produção e um guia para compreensão das variáveis adotadas?

**Graus de incerteza**

- O texto explicita as lacunas e o caráter provisório das informações consultadas?
- Apresenta as limitações dos dados e dos métodos empregados?
- Há clareza do que ainda precisa ser apurado ou explicado?

**Marcas de apuração**

- O texto torna visível o trabalho de apuração e verificação do jornalista?
- Há menção à dificuldade de acesso aos dados ou ao caminho percorrido para encontrá-los?
- Explicita as tentativas frustradas de contato com uma fonte pessoal importante?
- Diz claramente quando precisou recorrer à Lei de Acesso à Informação e o tempo de espera pela resposta?

### **Metodologia**

- O texto apresenta uma nota metodológica?
- Os procedimentos de seleção, coleta e análise de dados estão descritos (ex.: filtros aplicados às planilhas)?
- É possível que, por meio da descrição metodológica, a audiência seja capaz de replicar a análise empreendida?

### **Reprodutibilidade e código aberto**

- Quando existe uso de linguagem de programação, há divulgação do código-fonte, permitindo sua reprodução?
- O código-fonte permite acesso ao passo a passo empregado na seleção e análise, incluindo os arquivos utilizados?
- O código-fonte está disponível em repositório aberto e apresenta possibilidade de reprodução via *creative commons* (em que o uso ocorre livremente mediante menção à fonte original)?

As perguntas anteriormente descritas foram sistematizadas no **Quadro 5**, a seguir, e servem de parâmetro para que o pesquisador seja capaz de observar os elementos de transparência adotados por empresas jornalísticas e em reportagens de Jornalismo Guiado por Dados. Para aplicar o protocolo, reforço que é necessário se colocar no lugar de um leitor médio, que deve ter acesso fácil a cada um dos elementos, encontrando-os no site e na reportagem que se deseja analisar. Sugiro, ainda, que sejam definidos parâmetros antes da investigação, indicando quais são os tópicos principais e que poderão apresentar pontuação diferenciada, quando for adotada algum tipo de pontuação.

**Quadro 5 – Indicadores do Protocolo de Transparência no Jornalismo Guiado por Dados**

Etapa	Elemento	Indicadores
1) Nível <b>organizacio- nal</b>	Assinatura, perfil e expertise	O texto está assinado? Permite acessar um perfil com informações biográficas do autor? Permite acessar outros trabalhos do autor? Há indicação do nível de conhecimento do autor sobre o tema?
	Contato	O site do jornal exhibe uma lista de contatos de fácil localização e acesso? Há modos variados de contato (e-mail, aplicativos de mensagem, telefone etc.)? Há ombudsman ou ouvidor? Há uma editoria de audiência? Há uma seção de comentários do leitor? Essa seção é mediada pelo veículo ou apenas remete a uma rede social digital?
	Correção de erros	Há uma política clara sobre a correção de erros? A descrição dessa política editorial é de fácil ou difícil acesso aos leitores? Há uma seção específica para as retificações? Há indicação, no texto original, de que houve algum tipo de equívoco? A correção de erros ocorre no texto original de forma clara (indicando, por exemplo, data e horário em que foi feita)? Há um botão que indica a possibilidade de o leitor comunicar erros? Existe preocupação do veículo em fornecer espaço adequado à comunicação de erros, de modo que a correção receba um espaço significativo para ganhar visibilidade e de fato funcionar como uma correção?
	Fontes de financiamento	O site do jornal explicita as fontes de financiamento ou patrocinadores? Há uma seção de prestação de contas com a indicação de receitas e despesas? É possível acessar o balanço patrimonial detalhado? A receita gerada pelo pagamento da audiência é monitorada e disponibilizada? Quando integra uma corporação ou conglomerado, o jornal permite o amplo acesso do leitor, indicando as demais empresas que o compõem?
	Manual da redação e princípios editoriais	O veículo apresenta claramente seus princípios editoriais? A transparência é um dos princípios citados pelo jornal? Há um manual da redação disponível aos leitores? Neste manual, há preocupação em explicitar a conduta ética do jornalista e suas relações com as fontes?

	Política de diversidade	O jornal apresenta uma política de diversidade de jornalistas, temas e fontes? Há uma editoria de diversidade, de modo a tratar dessa política? Há subsídios ou ferramentas para o leitor monitorar o cumprimento da política de diversidade?
	Política de privacidade	O leitor tem acesso a informações claras sobre o tratamento de seus dados pessoais ao acessar o site do jornal? O veículo detalha o tipo de informação coletada pelo site? O jornal apresenta os dados sobre o perfil da audiência?
	Tipo de conteúdo	Há marcadores que permitam identificar se um conteúdo é informativo ou opinativo? No caso do conteúdo opinativo, existe clareza de que o texto é uma análise feita por um especialista? Há marcadores de texto indicando conteúdo editorial? O veículo investe em projetos ou iniciativas para o letramento do leitor, de modo que ele possa identificar diferenças entre informação e opinião? O veículo investe em projetos ou iniciativas para o letramento do leitor, de modo que compreenda o processo de produção de notícias e reportagens?
2) Nível <b>operacional</b>	Data de publicação	A data de publicação do texto está visível? A data de eventuais atualizações ou correções está visível, havendo diferenciação em relação à data original da publicação? Quando se trata de um conteúdo que entrou em recirculação, há destaque visual referente à data original, alertando para que observe este ponto?
	Documentos	Os documentos utilizados como fonte estão claramente identificados? Há acesso à íntegra desses documentos, com possibilidade de consulta aos materiais originais? Os documentos são disponibilizados junto ao corpo do texto, para pleno acesso do leitor, ou por meio de hiperlink, direcionando a outra página?
	Fontes de dados	As fontes de dados estão claramente identificadas? O texto oferece acesso à fonte original, incluindo indicação de hiperlink externo? Há materiais complementares à fonte original (a exemplo dos dicionários de dados), oferecendo ao leitor o contexto de produção e um guia para compreensão das variáveis adotadas?
	Graus de incerteza	O texto explicita as lacunas e o caráter provisório das informações consultadas? Apresenta as limitações dos dados e dos métodos empregados? Há clareza no que ainda precisa ser apurado ou explicado?

	Marcas de apuração	O texto torna visível o trabalho de apuração e verificação do jornalista? Há menção à dificuldade de acesso aos dados ou ao caminho percorrido para encontrá-los? Explicita as tentativas frustradas de contato com uma fonte pessoal importante? Diz claramente quando precisou recorrer à Lei de Acesso à Informação e o tempo de espera pela resposta?
	Metodologia	O texto apresenta uma nota metodológica? Os procedimentos de seleção, coleta e análise de dados estão descritos? É possível que, por meio da descrição metodológica, a audiência seja capaz de replicar a análise empreendida?
	Reprodutibilidade e código aberto	Quando existe uso de linguagem de programação, há divulgação do código-fonte, permitindo sua reprodução? O código-fonte permite acesso ao passo a passo empregado na seleção e análise, incluindo os arquivos utilizados? O código-fonte está disponível em repositório aberto e apresenta possibilidade de reprodução via <i>creative commons</i> (em que o uso ocorre livremente mediante menção à fonte original)?

Fonte: A autora (2021)

Os níveis de transparência – organizacional e operacional – devem funcionar em consonância. Um jornal que tem a preocupação de apresentar práticas mais transparentes deve fazê-lo não apenas na publicação de reportagens e projetos, mas na demonstração de como se dá sua sustentabilidade financeira. A seguir, apresento uma proposta de método com ênfase na transparência e na reprodutibilidade, compreendendo esses aspectos como norteadores da prática jornalística baseada em conhecimento.

## 6.2 Proposta de método com foco na transparência e na reprodutibilidade

No início de 2017, sob a coordenação dos professores Luciana Mielniczuk (*in memoriam*), Ilza Girardi e Marcelo Träsel, testamos na disciplina de Jornalismo Ambiental em Dados (PPGCOM/UFRGS) o que intuíamos ser uma peça fundamental neste tipo de prática: a elaboração de uma hipótese como ponto de partida para a investigação jornalística e a descrição metodológica do processo de apuração em bases de dados para posterior verificação e reprodução por parte dos pares e dos leitores.

A disciplina teve início com o aporte teórico sobre jornalismo digital e de dados e jornalismo ambiental, etapa que antecedeu a prática jornalística e a geração de produto – o portal *Guaíba Dados*<sup>127</sup>, que além de repositório tornou-se um projeto de extensão na UFRGS. Os pós-graduandos, entre os quais eu me incluo, investigaram à época diferentes segmentos relacionados ao Lago Guaíba (como política, indústrias, agrotóxicos etc.) que pudessem ser observados a partir de dados públicos. O lago é fundamental para o abastecimento da Região Metropolitana de Porto Alegre, e por isso foi escolhido como tópico principal.

Todas as seções do site, disponibilizado ao final da disciplina, apresentam uma estrutura textual comum: *proposta, dados encontrados, caminhos, descaminhos e caminhos futuros*. Este fio condutor é, portanto, um método desenvolvido para produções orientadas por dados. A *proposta* está ligada à ideia de hipótese em que se especula o desdobramento de determinado tópico. Os *dados encontrados*, por sua vez, correspondem ao processo de busca, análise e contextualização das informações quantificadas. Após a comunicação dos dados

---

<sup>127</sup> Disponível em: <https://www.ufrgs.br/guaibadados/>



encontrados, o texto de cada segmento indica os *caminhos*, ou seja, a descrição metodológica referente ao passo a passo adotado para obter os resultados indicados, e os *descaminhos*, mostrando as limitações dos dados ou pontos que a reportagem não pôde abranger. Por fim, há ainda os *caminhos futuros*, sugerindo propostas para a continuidade da investigação.

O estabelecimento de um método comum a todos foi fundamental para que houvesse uniformidade de apuração entre os próprios pós-graduandos, visto que o trabalho foi desenvolvido em duplas e trios dentro de uma mesma editoria. Se um de nós tivesse seguido seu próprio método em vez de um passo a passo pré-estabelecido, não teríamos obtido os mesmos resultados. Dessa forma, uma documentação exaustivamente descritiva auxiliou a evitar dúvidas e percalços.

Constituído e explorado em caráter laboratorial, o *Guaíba Dados* buscava estruturar e oferecer melhores práticas para o Jornalismo Guiado por Dados. À época de seu desenvolvimento, o JGD era menos disseminado no Brasil do que atualmente. Portanto, havia conhecimento limitado em relação às possibilidades de trabalho com os bancos de dados, incluindo menor adesão às linguagens de programação. A título de ensaio, acreditamos que o uso de softwares de planilhas seria suficiente para aquela investigação. Posteriormente, observamos que passou a existir maior oferta de cursos e oportunidades que uniam jornalismo e programação. O *Centro Knight para o Jornalismo nas Américas*<sup>128</sup>, referência em cursos gratuitos e amplamente difundidos, por exemplo, ofereceu o curso *Introdução à Programação Python para Jornalistas*<sup>129</sup> em língua portuguesa ao final de 2017.

Levando em conta o cenário atual, a proposta empreendida dentro do projeto *Guaíba Dados* poderia ser aprimorada a partir da inserção de novos elementos de transparência metodológica, a exemplo do uso de código aberto e da publicação da íntegra de alguns documentos. O emprego de editores de planilhas como *Excel* e *Google Sheets* continua sendo útil para executar análises pontuais e, embora esses editores também permitam o uso de formatos

---

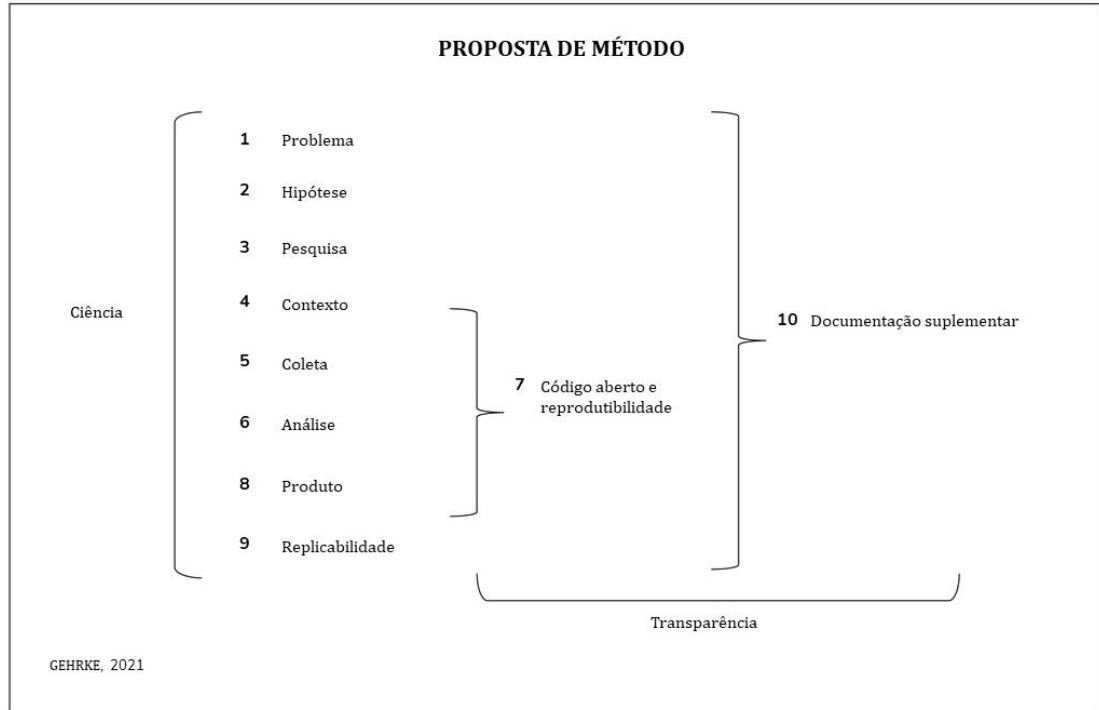
<sup>128</sup> Os cursos oferecidos pelo *Centro Knight*, vinculado à Universidade do Texas, nos Estados Unidos, acompanham o desenvolvimento de novas práticas no jornalismo feito nas Américas, visto que as aulas são oferecidas em inglês, espanhol e português. Esses cursos em geral são gratuitos e aceitam uma quantidade ilimitada de pessoas, e por isso de certa forma suprem a necessidade de conteúdo muitas vezes ainda não contemplados pelos cursos de ensino superior nas universidades ou mesmo em tradicionais congressos e eventos sobre jornalismo digital e investigativo.

<sup>129</sup> Disponível em <https://journalismcourses.org/pt-br/news/com-apoio-do-google-news-lab-centro-knight-oferece-curso-online-de-programacao-em-python-para-jornalistas/>

abertos e replicáveis como o .csv<sup>130</sup>, apresentam certa opacidade inerente ao software. Gerações mais recentes de jornalistas de dados são normalmente chamadas também de *journalcoders* e jornalistas programadores, evidenciando o papel de competências híbridas na prática de reportagem e, assim, a combinação de metodologias e de tecnologias inovadoras (HAHN; STALPH, 2019).

Apresento, a seguir, uma proposta metodológica que enfatiza a transparência no Jornalismo Guiado por Dados. Ao fazê-lo, busco suprir uma lacuna teórica e prática detectada em meus estudos e em métodos previamente desenvolvidos. Minha proposta desloca a transparência para o eixo central e tem o uso de código como ponto fundamental para o escrutínio entre pares e a replicabilidade ou reprodutibilidade do método. Em minha percepção, esta é uma das principais formas de fazer cumprir a expectativa de Meyer (2002) em relação ao jornalismo, aproximando as práticas orientadas por dados da ciência.

**Figura 32 – Proposta metodológica com ênfase na transparência**



Fonte: A autora (2021)

<sup>130</sup> O .csv é um dos formatos de arquivo identificados como aberto pela Open Knowledge Foundation e presentes no ranking do criador da Web, Tim Berners-Lee. Disponível em: <https://5stardata.info/pt-BR/>

A proposta retratada pela **Figura 32** apresenta uma série de passos para o trabalho jornalístico. Parte-se de um **problema** (1) e de uma **hipótese** (2) para a qual devem ser buscadas informações por meio de **pesquisa** (3) e cujos bancos de dados apresentam um **contexto** (4) sobre como foram produzidos e onde estão inseridos. Após examinar a viabilidade de uso das informações encontradas, parte-se para as etapas de **coleta** (5) e **análise das informações** (6), que demandam a elaboração de linhas de **código aberto** (7) para posterior **reprodutibilidade**. O código está relacionado ao **produto** (8) que se deseja criar – pode ser uma visualização de dados ou algum conteúdo interativo que evidencie os dados dos leitores, por exemplo – e à **replicabilidade** (9), que diz respeito às possibilidades de uso e reuso do material desenvolvido, incluindo informações sobre eventuais conflitos de interesse entre a empresa jornalística e o tema abordado. Por fim, todas as etapas do método estão sujeitas à **documentação suplementar** (10), ou seja, ao detalhamento do passo a passo, oferecendo aos leitores informações que podem ajudar na compreensão sobre a tomada de decisões dos jornalistas em uma investigação. Pode-se fazer uma analogia com a bula que acompanha as medicações: é por meio dela que o público tem acesso a recomendações e ressalvas inerentes ao produto.

Esta proposta de método traz dois conceitos importantes em seu horizonte: a ciência é o fio condutor do passo a passo metodológico (disposta no eixo vertical), enquanto a transparência (disposta no eixo horizontal) acompanha a consolidação de duas etapas principais: **o código aberto e a reprodutibilidade** e a **documentação suplementar**. No caso da transparência, pode-se dizer que aciona as duas categorias discutidas no **Capítulo 3**. Concatena aspectos como **operadora de uma prática ética**, à medida que revela seus métodos de trabalho, e da transparência como **recurso discursivo de legitimação**.

Antes de apresentar cada etapa do método, explico alguns conceitos fundamentais para a compreensão da proposta como um todo. O **código aberto** (*open source*) diz respeito à livre consulta, distribuição e modificação de um produto, sem necessidade de pagamento para utilização. Atualmente, as linguagens de programação *Python* e *R* estão entre as mais utilizadas no jornalismo e na ciência de dados porque podem ser empregadas livremente e porque apresentam uma grande comunidade de usuários. Quanto mais usuários se dedicam a aprender e a utilizar uma linguagem de programação, mais bibliotecas ou pacotes – série

de instruções que buscam resolver um problema específico, segundo Machlis (2019) – são disponibilizados gratuitamente. Além disso, o espírito de coletividade permite a troca de informações e ideias sobre a solução de problemas em códigos<sup>131</sup>, e tanto a linguagem *Python* quanto *R* apresentam interfaces amigáveis para elaboração de códigos, como o *Jupyter Notebook* e o *R Studio*, respectivamente.

O código aberto é uma funcionalidade essencial para a **replicabilidade** ou **reproduzibilidade** do método. Para explicar as diferenças entre os termos, tomo como base a explicação de Trielli e Diakopoulos (2019): enquanto a *replicabilidade* pressupõe o aproveitamento de um código para novas análises, ou mesmo poderia se referir a uma metodologia explicada narrativamente, a *reproduzibilidade* envolve a execução de etapa por etapa, permitindo a verificação completa dos resultados e sua validação, semelhante às premissas adotadas na ciência aberta – a qual será discutida no final deste capítulo.

O método proposto nesta seção é destinado à produção de conteúdo informativo guiado por dados e é executável nos casos em que a equipe de reportagem dispõe de tempo prolongado para investigar um tema. Cabe observar que um jornalista dificilmente conseguirá desenvolver todas as etapas de trabalho sozinho (BRADSHAW, 2017), embora seja recomendável que o profissional conheça toda a cadeia de produção noticiosa. No final dos anos de 1960, quando implementou uma metodologia de teste de hipóteses e análise para a cobertura dos protestos de Detroit, o então repórter Philip Meyer contou com o auxílio de acadêmicos para elaborar o desenho da pesquisa.

A proposta que apresento aqui tem o objetivo de servir como um guia para a análise crítica que envolve o emprego de dados e documentos no JGD em suas diferentes etapas, ressaltando que os dados funcionam como uma fonte e, por isso, devem ser consultados com cautela. Assim, não é mandatório que todos os jornalistas aprendam uma linguagem de programação – visto que a obrigatoriedade inclusive poderia gerar precarização de trabalho e acúmulo de tarefas –, mas é fato que este é um caminho que aproxima o Jornalismo Guiado por Dados da ciência aberta, e que é uma vantagem conhecer as possibilidades e os limites inerentes ao desenvolvimento de código. Embora muitas situações de trabalho sejam

---

<sup>131</sup> Neste aspecto, um dos principais repositórios destinados à discussão de dúvidas e questões relacionadas ao código aberto é o *Stack Overflow*, disponível em: <https://stackoverflow.com/>

resolvidas a partir do uso de planilhas, essa é uma ferramenta limitada em termos de reprodutibilidade e transparência – questões que esta proposta de método pretende contornar.

### 6.2.1 Problema

Problema é o termo que caracteriza o cenário de investigação jornalística sobre um tema de interesse público. Segundo Guerra (2008), a detecção de um problema representa o início do percurso interpretativo<sup>132</sup> do jornalista, ou seja, quando se estabelece o objeto de investigação. Caracterizado o cenário que se pretende cobrir, o profissional está apto a criar hipóteses relacionadas ao problema. A solução desse problema, então, é o entendimento final que se apresenta no jornalismo em forma de notícia, reportagem ou projeto investigativo.

A escolha da palavra *problema* nesta proposta de método é intencional à medida que se aproxima da ciência. Na pesquisa acadêmica, Gil (2019a) associa a existência de um problema a um tópico ou recorte pouco explorado e que precisa ser investigado em determinado campo do conhecimento. Além disso, na ciência, é preciso encontrar formas de testar um problema – que pode ser estruturado na forma de pergunta – por meio de verificação empírica, havendo procedimentos rígidos e sistemáticos. Uma estratégia para pensar a respeito do problema é indicar, inicialmente, um tema que se pretende abordar. Assim, em um segundo momento, torna-se mais fácil estabelecer o problema derivado do tema.

O problema deve ter algum tipo de delimitação – temporal ou espacial, por exemplo – para que seja viável desdobrá-lo em hipóteses. A abordagem que estabelece um problema é frequente nos processos de empreendedorismo. Neste caso, a ideia é encontrar soluções para problemas relevantes para o público-alvo e que geram impacto na vida em sociedade; utiliza-se de estratégias de pesquisa para medir esse alcance. O mapeamento de problemas, neste contexto, também faz com que empreendedores possam errar e corrigir rapidamente o rumo de sua atuação ao refletir e alterar o problema e as hipóteses que dele se desdobram.

---

<sup>132</sup> O percurso interpretativo do jornalista, diz Guerra (2008, p. 189), corresponde ao conhecimento do jornalista sobre um fato. Trata-se do movimento conduzido pelo profissional para “[...] costurar e estabelecer um conjunto e nexos entre coisas, pessoas, textos e expectativas da audiência [...]”. Além do próprio percurso, classificado pelo autor como o ato de interpretação, os jornalistas lançam mão do que ele chama de “técnicas cognitivas”, ou seja, os saberes acionados por eles ao realizar a interpretação dos fatos.

Ao jornalismo, pensar a partir de um problema é útil também para conhecer melhor a audiência a quem o produto se destina. O objetivo de iniciar um processo com tal raciocínio traz uma abordagem que se difere da elaboração da tradicional pauta jornalística, que em geral é pré-definida com ângulos e fontes, sem que o jornalista possa, de fato, entender o panorama de uma situação a ser investigada. Em minha perspectiva, após a indicação do problema, torna-se viável criar e testar hipóteses.

### 6.2.2 Hipótese

A elaboração de uma **hipótese** é central na investigação científica. Funciona como uma espécie de guia para o pesquisador, que direciona o foco para questões específicas a serem respondidas, e para o leitor, que tem em mãos uma espécie de planos e premissas para compreender os resultados apresentados (CRESWELL, 2010). No cotidiano do jornalismo, a preparação de uma pauta normalmente substitui a hipótese, visto que surge de uma observação não controlada da realidade em vez de um sistema teórico previamente estabelecido (MEDITSCH, 1992).

O Jornalismo Guiado por Dados, porém, apresenta algumas diferenças em relação ao jornalismo praticado diariamente. Produz notícias-sobre-relações (NIELSEN, 2017), ou seja, conteúdo contextualizado que buscará um caráter mais duradouro, diferentemente das efêmeras notícias-como-impressões (ou seja, os alertas de um novo acontecimento, normalmente divulgados nas redes sociais digitais) e das notícias-como-itens, que tem a durabilidade aproximada de 24 horas.

A hipótese vem sendo sugerida por pesquisadores para guiar projetos jornalísticos investigativos (HUNTER *et al.*, 2013) que naturalmente demandam maior tempo de apuração. Em linhas gerais, acredita-se que o jornalismo poderia se beneficiar de um pensamento científico estruturado, partindo de uma afirmação norteadora, para testar a existência de um fenômeno em um recorte da realidade. Segundo Gil (2019b), hipóteses são afirmações que são testadas e aceitas ou testadas e refutadas de acordo com os resultados obtidos, e podem tanto relacionar causa e efeito quanto possuir um caráter explicativo. Observação, consulta a outros estudos e conhecimento teórico estão nas bases de sua formulação.

No jornalismo, segundo Hunter *et al.* (2013), o uso de hipóteses é benéfico porque permite que um aspecto bastante específico seja verificado. Desta forma, em vez de trabalhar com um emaranhado de dados sobre os quais precisa gerar sentido sem ter um ponto de partida, o repórter realiza algumas suposições que deverão ser testadas. Caso se conclua que não há um fato noticiável, pode-se recomeçar e examinar outras hipóteses e possibilidades. Representa, ainda, uma maneira de mudar rapidamente o curso de uma investigação.

Em pesquisa anterior, sugerimos que a formulação de hipóteses é basilar para investigações em bancos de dados públicos (GEHRKE; FREITAS, 2021), chamadas de fontes abertas por Hunter *et al.* (2013) e também conhecidas como fontes documentais (GEHRKE, 2018a). Mesmo nos casos em que a fonte principal é composta por documentos vazados por terceiros, os jornalistas têm condições de formular hipóteses iniciais e modificá-las posteriormente.

A adoção de hipóteses no Jornalismo Guiado por Dados pode ser considerada uma herança do Jornalismo de Precisão, pois integra os primeiros passos para o cumprimento do método científico na prática jornalística, permitindo que os acontecimentos sejam mensurados (MEYER, 2002). Relatado no **Capítulo 3** da tese, a cobertura dos protestos em Detroit no ano de 1967, por Philip Meyer, envolveu o teste de hipóteses e a coleta de dados a partir de pesquisa feita com os moradores em vez de apenas ouvir estudiosos ou mesmo de coletar impressões pontuais de moradores. Para elaborar o questionário e desenhar a amostra a ser consultada, Meyer teve o apoio de acadêmicos da Universidade de Michigan e da Universidade Wayne State (DAVIS, 2021). Como dissemos na seção anterior, não é esperado que um único jornalista ou grupo de jornalistas execute todos os procedimentos esperados. A colaboração, característica do JGD, pode ser um ponto positivo a ser explorado com o auxílio de cientistas habituados a esses processos.

Wihbey (2019) tem uma visão otimista da aproximação do jornalismo com a ciência: para ele, o uso de hipóteses no jornalismo significa que esta prática está apta a receber e conhecer novas evidências e, desta forma, operar em uma perspectiva científica – aberta e mais próxima da realidade das pessoas, estabelecendo espaços de diálogo com a audiência e admitindo, sobretudo, a provisoriade de dados e informações sobre um acontecimento, eventualmente mostrando suas incertezas. Em consonância com a ideia de uma prática científica, a etapa de busca das informações, a seguir explorada, é chamada de pesquisa.

### 6.2.3 Pesquisa

O teste de hipóteses, como explicado na etapa anterior, depende da disponibilidade de dados – o que geralmente é suprido a partir das fontes públicas documentais (GEHRKE, 2018a; GEHRKE; FREITAS, 2021) ou abertas (HUNTER *et al.*, 2013), incluindo a vertente passiva da Lei de Acesso à Informação.

A **pesquisa**, nesta etapa do método, está relacionada à busca de informações, ou seja, ao entendimento do tópico a ser abordado e de que forma os dados podem ser obtidos. Há, portanto, uma fase inicial de abstração e análise crítica, em que é necessário identificar quais são as fontes em potencial que poderão auxiliar no teste da hipótese levantada. Em uma segunda etapa, pode-se pensar sobre as formas da extração propriamente dita, conforme veremos nas próximas seções.

Quando a hipótese que o jornalista deseja testar não está amparada por dados públicos já existentes, uma das alternativas é a criação de um banco de dados próprio, estruturado nas organizações jornalísticas. Para tanto, é recomendável que os jornalistas pensem, sistematicamente, sobre como coletar informações sobre sua área de especialização, a exemplo de política, economia, saúde e esporte. Bradshaw (2013) sugere que os jornalistas reflitam sobre sua prática como se fossem bibliotecários, criando o hábito de sistematizar e armazenar informações que possam ser úteis posteriormente.

A título de ilustração, podemos pensar em como essa sugestão poderia ser cumprida na área de segurança pública. É factível, por exemplo, que os repórteres registrem sistematicamente dados sobre incêndios, como sugerido por Holovaty (2006) nos primórdios do Jornalismo Guiado por Dados, ou como realizado pela equipe do jornal *Diário Gaúcho* em 2016 para contornar a demora na liberação de dados públicos por parte da Secretaria de Segurança Pública do Rio Grande do Sul sobre roubos, homicídios e latrocínios (GEHRKE, 2016b). A criação de uma planilha colaborativa desenvolvida entre os membros da própria equipe permitiu que o jornal cobrisse não só os casos pontuais de violência, mas gerou um dossiê após cinco anos de coleta de dados.

Criar o próprio banco de dados para posterior consulta e teste de novas hipóteses é viável em diversas áreas. A depender da situação, os jornalistas podem inclusive envolver a audiência em um processo de apuração coletiva (*crowdsourcing*), a exemplo da mobilização



realizada pelo jornal *La Nación*, na Argentina, com relação aos 40 mil áudios interceptados que contribuíram para a investigação jornalística sobre a morte do promotor Alberto Nisman<sup>133</sup> em 2015. Além de viabilizar o teste de hipóteses independentemente de haver dados oficiais, este tipo de coleta permite, sobretudo, que o jornalista seja capaz de identificar padrões – ação fundamental na análise de dados.

Seja público, seja um produto desenvolvido dentro de uma organização jornalística, um conjunto de dados deve vir acompanhado de ampla documentação para que a audiência seja capaz de compreender o modo como as informações foram coletadas e como são apresentadas. As fontes, por sua vez, apresentam o próprio **contexto** de produção e limitações inerentes ao processo.

#### 6.2.4 Contexto

Nesta proposta metodológica, o **contexto** está associado ao cenário de produção e de disponibilidade dos dados. Como dizem Lückman e Fonseca (2017, p. 163), o contexto remete “[...] à ideia de inter-relação de circunstâncias que acompanham um fato ou uma situação”. Por opção ou devido às limitações na estruturação de bancos de dados próprios, jornalistas de dados utilizam fontes oficiais com frequência. Em geral, apontam as estatísticas oficiais como confiáveis, raramente verificando os números utilizados (LAWSON, 2021). Ao analisar a cobertura sobre a crise humanitária, por exemplo, o autor descobriu que, no lugar da verificação dos números, os jornalistas tendem a hierarquizar as fontes que fornecem os dados, acionando, por fim, a fonte institucional que consideram mais confiável.

Em minha percepção, a falta de verificação e adição de contexto aos dados é um movimento prejudicial ao jornalismo – ao mesmo tempo em que o Jornalismo Guiado por Dados busca fugir de uma prática declaratória, acaba reproduzindo essa falha por outras vias quando se atém apenas às fontes oficiais, principalmente se desconsiderar seu contexto de produção. Como apontado por Lawson (2021) no título do artigo que discute este tópico, ao fazer isso os profissionais “se escondem atrás das fontes”. Décadas antes, Tuchman (1972) já

---

<sup>133</sup> Disponível em: <https://www.niemanlab.org/2017/04/how-the-argentinian-daily-la-nacion-became-a-data-journalism-powerhouse-in-latin-america/>

havia detectado este tipo de procedimento nas redações jornalísticas, chamando-o de “objetividade como ritual estratégico” – ou seja, quando os profissionais buscam se resguardar de eventuais erros ao utilizar citações diretas às fontes ouvidas, de forma acrítica.

A inserção de informações em uma planilha ou em outro tipo de documento envolve um esforço inicial humano e, por isso, está sujeita a equívocos. Assim, os dados devem ser encarados como uma fonte documental qualquer e, por isso, precisam ser avaliados em sua totalidade. Defendo que o contexto dos dados deve ser estudado tão logo o jornalista encontre as evidências com as quais pretende trabalhar. Uma das alternativas para evitar equívocos é criar uma espécie de “biografia dos dados”, como sugerido pela cientista de dados Heather Krause (BELISÁRIO *et al.*, 2020). Isso significa que, além de localizar as estatísticas que pretende analisar, o profissional deve ficar atento à documentação que acompanha esses dados (geralmente chamados de dicionários de dados) para compreender o uso de variáveis e da metodologia adotada em uma pesquisa. É preciso registrar, portanto, aspectos sobre a origem, a captação, o recorte e as limitações dos números.

Belisário *et al.* (2020) recomendam uma abordagem semelhante a que seria realizada com fontes humanas, em que o profissional deve se perguntar: quais são os interesses da fonte, suas áreas de especialidade e atuação? Qual é o histórico dessa fonte como informante? Há uma conexão com pessoas ou grupos capazes de gerar conflito de interesse ou algum tipo de controvérsia? Essa fonte costuma ser ouvida por outros jornalistas? Rusbridger (2020) adiciona: de onde vêm os dados, quem os coletou e por quê? A resposta para essas e outras questões compõem a biografia de dados, que representa o detalhamento contextual das informações estruturadas. Trata-se de uma abordagem semelhante ao princípio de crítica às fontes (STEENSEN, 2019), premissa segundo a qual toda fonte tem uma tendência – e descobrir qual é essa tendência, confrontando-a com outras fontes, é essencial para o processo de apuração e verificação no jornalismo.

Em minha dissertação (GEHRKE, 2018a), analisei as fontes de 60 notícias produzidas a partir de técnicas de JGD. Do total de 315 fontes acionadas, 213 eram fontes documentais – isto é, que pressupõem um tipo diferente de abordagem das fontes humanas tradicionais, uma vez que não podem ser consultadas por meio de entrevista, mas devem ser observadas sem serem submetidas a perguntas e respostas diretas. Esse dado sugere que há uso recorrente de bancos de dados oficiais como matéria-prima no JGD, em especial dos repositórios

com dados já disponíveis. Séries históricas, estatística descritiva, indicadores, taxas e relatórios representavam 67,61% das fontes encontradas no levantamento.

Considerar o contexto de produção dos dados é um dos sete princípios apontados por D'Ignazio e Klein (2020) na obra *Data Feminism*, cuja proposta tem como objetivo discutir os usos e as limitações de dados, a começar por reconhecer que o poder não é distribuído de forma igualitária no mundo. Ponderar o contexto de produção, então, tem relação com admitir que os dados muitas vezes são produtos de relações sociais não igualitárias, e reconhecer esse cenário tem a ver com a condução de uma análise acurada e ética. As autoras relatam, por exemplo, que muitas vezes há sub-representação de minorias em dados demográficos. Quando esses dados são utilizados de forma acrítica, sem o reconhecimento das limitações, acabam sendo tomados como um retrato fiel da realidade, podendo gerar notícias, produtos e outras formas de conhecimento que não correspondem à realidade. Nestes casos, é necessário, por exemplo, não apenas olhar para os dados estruturados, mas para suas formas de coleta – as limitações por vezes se encontram em uma etapa anterior à publicação dessas informações em repositórios públicos.

Para evitar que os dados sejam utilizados sem uma análise criteriosa ou fora de contexto, pesquisadores como Diakopoulos (2019, p. 11) têm discutido o termo “*algorithm accountability*”, isto é, conhecer e reconhecer o modo como os dados são coletados, estruturados, e que tipo de informação podem gerar. Em processos altamente automatizados, em que há o uso de aprendizado de máquina, esta espécie de auditoria de algoritmos torna-se ainda mais fundamental para detectar vieses de gênero e raça, por exemplo. A presença de algoritmos em diversos processos sociais altera também a relação das pessoas com o processo de tomada de decisões – que inclui a justiça criminal e a educação, por exemplo.

O esforço analítico é essencial para dar continuidade às etapas seguintes desta proposta metodológica, evitando erros e problemas na coleta de dados.

### 6.2.5 Coleta

Observado o contexto de produção e disponibilização de dados, bem como a documentação que os acompanha, pode-se, finalmente, passar para a etapa da **coleta** de informações estruturadas. A extração de dados pode ocorrer de forma automatizada – em que há o auxílio de linguagem de programação para a criação de uma sequência que faça a raspagem de dados – e não automatizada, como o simples download de uma planilha (ou de um conjunto de planilhas).

Sobre este tópico, Bradshaw (2013) sugere que o jornalista de dados tente *pensar como um computador* na hora de coletar as informações. Significa que, se o banco de dados provavelmente será utilizado em consultas futuras, cabe ao jornalista pensar em formas de automatização da extração, de modo que não haja retrabalho. Em situações como essa, pode-se trabalhar com formas como *Application Programming Interface* (API) presentes em extensas bases de dados.

A coleta de dados deve estar relacionada ao problema e à hipótese proposta no início do método, e por isso pode envolver a combinação de diferentes fontes. Nem sempre a coleta de dados é simples e direta – em alguns portais governamentais, por exemplo, os dados são disponibilizados em arquivos fechados (como .pdfs, que dificultam a extração das informações). Nestes casos, dizemos que os dados estão disponíveis, mas não estão acessíveis.

Nesta etapa de obtenção das informações, o jornalista também deve ficar atento ao modo como as informações estão organizadas para prevenir-se de eventuais erros de digitação ou ausência de informações. Exige, portanto, a limpeza do banco de dados, envolvendo a detecção de erros antes de passar para a etapa analítica. Para tanto, existem manuais técnicos que auxiliam nesses detalhes.

Conforme explicado na **Figura 32**, é a partir desta etapa que se pode pensar na estruturação de código aberto de trabalho, fornecendo o detalhamento de como os dados foram coletados e como serão analisados posteriormente.

### 6.2.6 Análise

A etapa de **análise** diz respeito aos resultados que serão obtidos a partir do cruzamento de informações estruturadas. O problema e a hipótese elaborados no início deste método é definidora, e a ela pode estar associado o uso de algumas estratégias estatísticas simples, como o cálculo da média, moda ou mediana, ou mesmo o uso de correlações e de modelos complexos, como alguns tipos de regressão.

A análise de dados requer múltiplos cuidados. É necessário verificar, primeiramente, se os números integram um conjunto de dados que apresenta distribuição normal. Os números que não pertencem a essa distribuição ordenada são chamados de *outliers* e, embora o jornalismo tenha a tradição de reportar casos singulares, dificilmente um número absoluto ou isolado é capaz de explicar um fenômeno. Assim, como descrevemos nas etapas anteriores, há de se considerar que os números estão inseridos em um contexto mais amplo, e para fazer sentido por vezes precisam ser submetidos ao emprego de taxas, índices e outras estratégias de uniformização de resultados.

Quando há o uso de softwares de planilhas para análise de dados, em geral são utilizadas fórmulas, filtros e tabelas dinâmicas, processos que costumam ficar ocultos nos resultados. No caso de haver estruturação das etapas analíticas em linguagem de programação, as ferramentas permitem que o passo a passo seja comentado e documentado, fazendo com que os procedimentos fiquem claros e acessíveis para que sejam replicados ou reproduzidos posteriormente. Por isso, o uso de código de programação é indicado nesta proposta de método, como veremos na próxima seção.

A análise de dados pressupõe, ainda, que haja conhecimento teórico específico sobre um tema de investigação, pois o cruzamento de variáveis só faz sentido quando há informações prévias sobre um tema. Por isso, sugere-se que os jornalistas reflitam sobre suas etapas de trabalho com o auxílio de especialistas na área que desejam cobrir.

### 6.2.7 Código aberto e reprodutibilidade

Softwares de planilha e linguagens de programação em código aberto, a exemplo de *Python* e *R*, estão entre as ferramentas mais recorrentes de trabalho entre os jornalistas de dados (BELISÁRIO *et al.*, 2020). Entretanto, como vimos no tópico anterior, as planilhas, mesmo que mais simples de manusear, apresentam limitações em termos de reprodutibilidade, visto que o passo a passo de filtros e fórmulas adotadas não costuma ser documentado. “Fórmulas – que representam o principal trabalho analítico em uma planilha – só ficam visíveis quando há o clique em uma célula. Isso significa que é mais difícil revisar os verdadeiros passos que levaram a uma conclusão”<sup>134</sup> (LEON, 2021, p. 149).

Nesta proposta de método, o uso de **código aberto e de reprodutibilidade** compõe o sexto passo metodológico porque congrega diferentes etapas: **coleta, análise, produto e replicabilidade**. Para que o processo seja reproduzido na íntegra, é preciso que todo o código de programação seja acessível e esteja comentado etapa por etapa. Dependendo da complexidade do projeto, é comum a utilização de diferentes repositórios para o código-fonte, de modo que todas as frentes de trabalho fiquem disponíveis para consulta.

---

<sup>134</sup> No original “Formulas – which do most of the analytical work in a spreadsheet – are only visible when clicking on a cell. This means that it is harder to review the actual steps taken to reaching a given conclusion.”

**Figura 33 – Exemplo de repositório e código-fonte utilizado em *No Epicentro***

**Replicação passo-a-passo**

Para reproduzir os resultados em sua própria máquina, você precisa configurar seu ambiente de desenvolvimento da seguinte maneira:

1. Clone o repositório
2. Baixe este arquivo compactado e extraia no diretório `data`.
3. Instale os requerimentos em um ambiente virtual do Anaconda com os seguinte comandos:

```
conda create -n gpd_0.8
conda activate gpd_0.8
conda config --env --add channels conda-forge
conda config --env --set channel_priority strict
conda install python=3 geopandas
conda install pygeos --channel conda-forge
conda install feather-format
conda install requests
```

Ou, caso prefira instalar com outro método, veja a sessão [Using the optional PyGEOS dependency](#) da documentação do GeoPandas.

4. Caso queira gerar também os pontos que aparecem no mapa, execute `python generate_points.py`. Isso deve demorar um bocadinho e é opcional.
5. Use `python prepare.py` para pré-processar diversos dados, o que vai otimizar o processo de cálculo.
6. Use `python run_query.py lat lon` para obter um objeto JSON-like com as informações necessárias para gerar a visualização de dados personalizada.
7. O arquivo `update.py` deve ser executado repetidamente em um intervalo fixo de tempo via cron ou mecanismo semelhante. Ele é responsável por atualizar a contagem de casos de covid-19 no país, além de já calcular previamente o raio de mortes nas principais cidades do país.

Tenha em mente que o aplicativo roda em um ambiente Anaconda, criado exatamente seguindo as especificações desse tutorial, em um servidor Ubuntu 18.04. Não fizemos testes em outras configurações, mas sintam-se a vontade para abrir um issue caso encontre algum problema.

Fonte: *No Epicentro – GitHub*. Captura de tela em 28 set. 2021.

Em uma perspectiva educacional, Leon (2021, p. 149) sugere o uso de “*literate programming environments*”, ou seja, uma espécie de ambiente de programação que cumpre um papel didático, semelhante à alfabetização, em que a linguagem humana se mistura à linguagem de computação. Na **Figura 33**, elaborada a partir do repositório<sup>135</sup> do projeto *No Epicentro*<sup>136</sup>, no *GitHub*, isso ocorre por meio do uso de blocos de código de programação junto aos blocos de texto, que servem de texto de apoio e como explicação. À esquerda na imagem, por exemplo, é possível visualizar uma lista com três eixos principais ligados ao desenvolvimento do projeto, com informações de como o mapa principal foi elaborado. À direita, observa-se o documento intitulado *read me*<sup>137</sup> referente ao primeiro eixo (*back*, destinado à organização das informações que serão fornecidas pelos usuários). Nele, os autores do projeto explicam o passo a passo de como é possível rodar as mesmas informações de qualquer computador. Adicionalmente, detalhes de como as decisões do projeto foram tomadas estão explicitadas

<sup>135</sup> Disponível em: <https://github.com/orgs/noepicentro/repositories>

<sup>136</sup> Disponível em: <https://piaui.folha.uol.com.br/lupa/epicentro/>

<sup>137</sup> Disponível em: <https://github.com/noepicentro/back/blob/master/README.md>

em uma publicação no *Medium*<sup>138</sup>. Para Leon (2021, p. 151), fornecer um material suplementar que funcione de forma didática para o público é uma das vantagens de apresentar o código-fonte nesse formato: “Embora o leitor não programador possa não ser capaz de entender ou executar o código em si, comentários e explicações no documento podem ser capazes de assegurá-los de que as medidas apropriadas foram tomadas para mitigar o erro”.

O compartilhamento de código-fonte apresenta relação com a cultura *open source*, ou seja, de fonte e código aberto, que emerge da interação entre jornalistas e programadores (CODDINGTON, 2014) e tem como inspiração o movimento de software livre. Por princípio, esse movimento permite que o usuário faça modificações, melhorias, reuso e distribuição em códigos de programação, como explicamos no início deste capítulo. Em geral, isso ocorre por meio de publicação das informações no repositório chamado de *GitHub*, que concentra códigos e documentação aberta de projetos.

Como detectou Träsel (2014) em sua investigação de doutorado, existe uma espécie de ética *hacker* entre jornalistas de dados, que buscam compartilhar conhecimentos e soluções acerca de um problema, caracterizando uma postura cooperativa dentro da comunidade – assim como ocorre no grupo de desenvolvedores e outros profissionais da ciência da computação. Exemplo disso é o fórum online *Jornalismo de Dados*<sup>139</sup>, utilizado para tirar dúvidas e estimular a busca por soluções. A partir do código aberto, a audiência tem acesso ao “DNA” do produto jornalístico, visto que pode conferir todas as etapas de formatação e execução. Por isso, torna-se ponto fundamental na transparência.

---

<sup>138</sup> Disponível em: <https://medium.com/datavizbr/como-fizemos-o-mapa-interativo-que-te-coloca-no-epicentro-da-epidemia-de-covid-19-no-brasil-4ce949a9183b>

<sup>139</sup> Disponível em: <https://jornalismodedados.org/>



### 6.2.8 Produto

A etapa intitulada **produto** está situada ao final desta proposta, e também pressupõe o uso de código aberto e detalhamento metodológico. Embora a visualização (produção de gráficos e mapas) seja um dos resultados mais comuns no Jornalismo Guiado por Dados, não é a única alternativa para ilustrar o conteúdo informativo desta natureza, que também abrange a criação de outros recursos interativos. Por isso, acreditamos que o termo é mais adequado para esta etapa do método, uma vez que a visualização pode ser empreendida em etapas anteriores, quando se deseja consultar a distribuição dos dados em uma amostra.

Independente do produto que se pretende gerar, os dados continuam sendo sua matéria-prima. Por isso, as etapas apresentadas anteriormente nesta proposta de método sugerem cautela e desempenho analítico para garantir a confiabilidade das análises e dos produtos gerados. Para que uma visualização ou outro material não sugira padrões equivocados, Cairo (2019) aponta que o rigor na verificação dos dados e o cuidado na organização das informações são essenciais nesse processo. As escolhas do que se pretende mostrar ou omitir, então, vão muito além da estética.

A criação e divulgação metodológica de um produto, como dissemos, está relacionada à disposição da organização jornalística em abrir seus processos. Esta “carta de intenções” pode ser identificada na etapa de replicabilidade.

### 6.2.9 Replicabilidade

A **replicabilidade** nesta proposta de método refere-se à descrição narrativa de decisões tomadas no Jornalismo Guiado por Dados, incluindo aspectos sobre a transparência editorial de um veículo. O percurso interpretativo (GUERRA, 2008) anteriormente mencionado, que corresponde ao trabalho do jornalista para a obtenção de informações, também aparece aqui: o estabelecimento de relações entre dados e fontes, bem como a hierarquização do conteúdo, são operacionalizadas pelos profissionais de forma racional. Dessa forma, as escolhas adotadas em uma cobertura podem ser justificadas – e inclusive comunicadas ao público quando há a defesa de procedimentos mais transparentes.

Quando o jornal possui a política de utilizar código aberto e disponibilizá-lo em repositório público, naturalmente está de acordo com as regras de compartilhamento de informações – ou seja, permite que o código seja verificado e reutilizado por terceiros. Sob licença *Creative Commons*, os códigos abertos normalmente publicados no *GitHub* podem ser reaproveitados desde que citadas a fonte original da coleta e/ou análise de dados. Dessa forma, trata-se de uma oportunidade para que os veículos de comunicação adicionem informações sobre a natureza da coleta de dados, incluindo eventuais conflitos de interesse diante de um tema ou investigação, como costuma ser apresentado na pesquisa científica.

Tanto a replicabilidade quanto as demais etapas do processo devem ser detalhadas e indicadas à audiência como documentação suplementar, abrindo a possibilidade de consulta a todo processo jornalístico.

### 6.2.10 Documentação suplementar

Nesta proposta de método, o passo a passo de todo o processo deve ser escrito e disponibilizado em código aberto juntamente com uma **documentação suplementar** que mostre as escolhas dos jornalistas de dados frente a um tópico de cobertura.

Além de incluir uma abordagem educacional levantada por Leon (2021), de modo que não especialistas possam fazer a leitura técnica dos procedimentos ocorridos, torna-se interessante incluir premissas da interação entre humanos e computadores. As práticas pertencentes ao *Human-Computer Interaction* (HCI) são usadas principalmente na pesquisa da área da computação e são identificadas por Trielli e Diakopoulos (2019) como benéficas ao jornalismo. Para que a reprodutibilidade seja possível, os autores listam alguns pontos que devem estar explicados no código e/ou no material que o acompanha: os métodos estão explicados? Os dados e os códigos usados pelos autores estão disponíveis online? Há explicação sobre como rodar o código ou replicar os resultados no futuro?

Bancos de dados públicos em geral vêm acompanhados de dicionários de dados com o detalhamento de variáveis, e proponho que os tópicos contemplados neste capítulo possam ser reunidos em formato de documentação suplementar para facilitar a reprodutibilidade do método, com uma descrição suficientemente detalhada das variáveis. Normalmente,

repositórios apresentam somente um documento intitulado “*read me*” com informações limitadas sobre as variáveis utilizadas na análise. A ideia, neste caso, é fornecer o máximo de subsídios.

Uma documentação minuciosa é chave para que todos possam seguir o mesmo método, o que facilita, por exemplo, o processo de investigação em rede, frequentemente realizado no Jornalismo Guiado por Dados – a exemplo de casos como *Panama Papers*, *Paradise Papers* e *Implant Files*, coordenados pelo Consórcio Internacional de Jornalistas Investigativos (ICIJ, na sigla em inglês).

### 6.3 Limites e perspectivas para a reprodução

O Jornalismo Guiado por Dados é apontado por Coddington (2014) como a especialidade que tem o maior potencial para a abertura metodológica se comparado a práticas anteriores, como a Reportagem Assistida por Computador (RAC), e até mesmo a práticas contemporâneas, como o Jornalismo Computacional.

Enquanto a RAC ainda está fortemente associada ao jornalismo investigativo, em que algumas fontes dificilmente são reveladas, o Jornalismo Computacional se encontra naturalmente atrelado ao uso de algoritmos e softwares muitas vezes opacos. Em geral, são utilizados para ao menos quatro grandes processos: priorizar, classificar, associar e filtrar (DIAKOPOULOS, 2019), e a automatização de procedimentos ocorre muitas vezes sobre os mesmos dados, envolvendo apenas as informações já disponíveis e evitando certos tópicos de cobertura para o qual não existem informações.

O JGD apresenta a análise e a publicação de dados no núcleo de suas práticas – ainda que uma parte da imprensa continue guardando materiais detalhados apenas para si, desviando de princípios básicos da transparência. Em sua tese, Träsel (2014) menciona que, ao não apresentar essas informações, os profissionais estariam deixando de incentivar a auditoria de seu trabalho pelo público, o que ocorreria por meio do movimento *Open Source*, e de revisão entre pares, como ocorre na ciência. Mazotte (2021), por sua vez, aponta que, ao omitir o passo a passo, os jornalistas privam os leitores dos meios de compreender uma notícia – compreensão, essa, ainda mais desejável em um momento de abundância de informações.

Como vimos na seção sobre código aberto, uma das saídas para que haja entendimento técnico do público é a adesão a um método que una linguagem de computação e linguagem humana, possibilitando que o código de programação seja comentado e explicado passo a passo, incluindo a documentação dos pacotes e dos dados utilizados. Uma audiência qualificada pode ser decisiva para apontar, inclusive tecnicamente, problemas de apuração jornalística (DALMASO; MIELNICZUK, 2011). Assim, a transparência jornalística vai além de procedimentos da comunidade jornalística voltados apenas para seus pares.

A implementação de melhores práticas de transparência exige mudança de mentalidade em toda a cadeia noticiosa e demanda, ainda, a reflexão sobre as formas de financiamento do jornalismo. Enquanto a abertura metodológica parece conquistar a simpatia de parte dos jornalistas de dados (BEILER; IRMER; BREDA, 2020; GEHRKE, 2020), não necessariamente é prioridade nas organizações jornalísticas, especialmente na imprensa de referência, por questões de concorrência entre empresas. Observa-se, ainda, que os jornalistas enxergam a si próprios como profissionais transparentes (CHRISTOFOLETTI, 2021), acionando a transparência como **operadora de uma prática ética**, mas que isso não necessariamente se materializa em seus ambientes de trabalho.

No método proposto neste capítulo, busco abrir espaço para que sejam demonstradas eventuais limitações ou conflitos de interesse por parte dos jornais, visto que uma empresa jornalística pode, ao mesmo tempo, estimular a transparência de sua equipe e agir de forma opaca sobre as relações mantidas com grupos financeiros e políticos, como constatado por Christofolletti (2021).

Embora qualquer tipo de mudança encontre resistência em estruturas organizacionais tradicionais, nativos digitais têm demonstrado que estabelecer uma relação honesta com a audiência, inclusive abrindo suas fontes de financiamento, é parte significativa da transparência e do estímulo à confiança. Ainda assim, é preciso admitir que as empresas só irão compartilhar as informações que lhes interessa, em especial aquelas que reforçam sua aparência de organização preocupada com a transparência, neste caso acionada como **recurso discursivo de legitimação**.

Apesar dos problemas identificados com relação à transparência, disponibilizar informações, ainda que parcialmente, é um passo fundamental para aproximar o Jornalismo Guiado por Dados do fazer científico porque envolve oferecer à audiência as mesmas evidências

às quais os jornalistas tiveram acesso para escrever as notícias. Conforme explorado por pesquisadores como Kovach e Rosenstiel (2014) e O'Connor e Weatherall (2019), o público dificilmente está em uma posição na qual consegue acessar, em primeira mão, as mesmas evidências utilizadas na prática da reportagem.

Sob este aspecto, uma citação de O'Connor e Weatherall (2019, p. 44-45) é bastante ilustrativa:

A razão para acreditar que houve um buraco na camada de ozônio não foi porque os cientistas disseram que existia; foi porque vários dispositivos, projetados e cuidadosamente calibrados para medir os níveis de ozônio na alta atmosfera, localizados em diferentes lugares e fazendo suas medições de maneiras diferentes, detectaram níveis de ozônio substancialmente reduzidos<sup>140</sup>

A transparência metodológica no jornalismo produz efeito semelhante: mostra aos leitores como se aferiu ou como se chegou a uma determinada evidência. A autoridade jornalística que explica por que isso ocorre ainda é válida – caso contrário, nem mesmo a apresentação de elementos de transparência faria sentido –, mas o fato de que o público pode verificar, de forma autônoma, a validade das evidências apresentadas pode ser visto como um fator importante para a relação de confiança.

A apresentação de evidências evoca os sentidos de reprodutibilidade e replicabilidade científica, das quais o Jornalismo Guiado por Dados se aproxima, e que sempre foram o objetivo do Jornalismo de Precisão. Os dois conceitos estão vinculados à transparência e trabalham com a ideia de código aberto. Enquanto a replicabilidade permite que haja a consulta a um código ou narrativa para novos experimentos, a reprodutibilidade busca garantir a reprodução de tais procedimentos na íntegra. “Enquanto publicar o código usado em uma análise computacional pode fornecer um grau de transparência, a falta de dados subjacentes, por

---

<sup>140</sup> No original “The reason to believe that there was a hole in the ozone layer was not because scientists said there was such a hole; it was because multiple devices, designed and carefully calibrated to measure ozone levels in the upper atmosphere, located in different places and making their measurements in different ways, detected substantially reduced ozone levels”.

exemplo, impede a reprodutibilidade, uma vez que a transparência algorítmica tem várias camadas”<sup>141</sup> (TRIELLI; DIAKOPOULOS, 2019, p. 2).

O Jornalismo Guiado por Dados se aproxima da perspectiva de ciência aberta pela proximidade de princípios e por meio da reprodutibilidade metodológica. Como resposta ao que chamam de “crise de replicabilidade” na pesquisa quantitativa, Dienlin *et al.* (2021) buscam implementar uma agenda desse tipo de pesquisa na área da Comunicação, estabelecendo algumas premissas para a prática da ciência aberta: 1) *publicação de materiais, dados e códigos*, de modo que seja viável uma verificação independente por parte de outros cientistas; 2) *pré-registro de estudos e submissão de relatórios*, fazendo com que as perguntas de pesquisa sejam registradas antes da coleta de dados, buscando evitar vieses; 3) *condução de replicações*, ou seja, a realização de novos testes a partir dos métodos adotados em pesquisas anteriores; 4) *colaboração*, que pressupõe maior troca entre os pesquisadores, incluindo troca proativa de dados e materiais de estudo; 5) *promoção de habilidades de ciência aberta*, encorajando o registro de estudos e uso de dados replicáveis; 6) *implementação de normas de transparência e abertura* para guiar autores e editores; e 7) *incentivo às práticas open science*, que podem estar associadas à promoção e ao financiamento em nível profissional.

Além do temor pela exposição de dados anônimos usados nas pesquisas que envolvem o uso de códigos e do receio de que possa haver um aumento significativo de trabalho para os estudiosos, limitações já elencadas na pesquisa de Dienlin *et al.* (2021), o emprego de premissas associadas à ciência aberta deve ser considerado dentro de um contexto mais amplo. Sob a perspectiva de implementação de protocolos de ciência aberta na América Latina, Oliveira *et al.* (2021) problematizam os parâmetros elencados por Dienlin *et al.* (2021), afirmando que aspectos como a replicabilidade, a aceleração e a eficiência estão associadas à padronização da ciência, e incluem apenas o contexto que historicamente beneficia o Norte Global, desconsiderando regiões menos privilegiadas, em que a distribuição de conhecimento é assimétrica.

Assim, Oliveira *et al.* (2021) propõem que a ciência aberta esteja amparada na participação, na justiça social e na democratização do conhecimento. Por isso, defendem a

---

<sup>141</sup> No original “While publishing the code used in a computational analysis might provide a degree of transparency, the lack of underlying data, for instance, impedes reproducibility, since algorithmic transparency has various layers”.

aproximação desse conceito com a abertura da produção científica, oferecendo acesso universal à informação e ao conhecimento, que entendem como sendo um direito básico – se a sociedade financia a ciência, deve ter acesso livre a ela. Por fim, os autores defendem que haja maior inserção da sociedade nesse tipo de prática, estimulando uma relação de construção de confiança entre o público e o pesquisador. Isso se relaciona com o que dizem Fox *et al.* (2021), para quem os indivíduos, as comunidades e os grupos sociais deveriam participar ativamente do processo científico.

Embora reconheçam que a replicabilidade e a transparência metodológica sejam princípios importantes das práticas científicas, Oliveira *et al.* (2021) temem que a necessidade de fornecer material complementar, a existência de novos repositórios e ferramentas para publicação de artigos em *pré-print* possam estimular uma espécie de plataformização da ciência, gerando custos os mais diversos, o que deixaria países subdesenvolvidos em desvantagem. Parece-me, assim, que cabe uma análise pontual de cada contexto para que ocorra a prática da ciência aberta.

Diante das breves leituras feitas até o momento, é possível perceber que o Jornalismo Guiado por Dados pode se aproximar das premissas da ciência aberta por meio da reprodutibilidade metodológica. Segundo Trielli e Diakopoulos (2019), são três os principais pontos para que isso ocorra: 1) validação de descobertas; 2) prevenção de erros; e 3) maximização da circulação de conteúdo. Os três eixos estão conectados à ideia de verdade verificável com a qual o jornalismo tem seu principal compromisso, abrindo informações a serem verificadas pelos pares e pelo público geral. Além disso, o próprio fato de disponibilizar dados potencialmente faz com que os jornalistas redobrem os cuidados com a verificação antes de sua publicação, evitando provocar novos ruídos no saturado espaço digital.

## 7 CONSIDERAÇÕES FINAIS

A pandemia de Covid-19, que ao final 2021 havia provocado a morte de mais de 600 mil brasileiros e de 5,4 milhões de pessoas no mundo, é um acontecimento global em curso. Conseqüentemente, um acontecimento jornalístico de amplas proporções. Além de provocar mudanças sociais, entendo que a pandemia acelerou processos que começavam a ser ensaiados no jornalismo e se manifestavam até então no discurso de profissionais, a exemplo da **transparência**.

Como valor para o jornalismo e com vistas à abertura de procedimentos de reportagem e decisões editoriais que se manifestam na esfera do produto, a transparência é um fundamento que faz sentido principalmente em iniciativas que têm a ciência como inspiração metodológica, caso do Jornalismo Guiado por Dados, que se propõe a gerar um tipo de conhecimento explicativo e duradouro – como as *notícias-sobre-relações*, conforme classificação de Nielsen (2017). Mesmo que tomem como base um acontecimento factual, as notícias vinculadas ao modo de produção do JGD são diferentes no que diz respeito ao conteúdo jornalístico efêmero corriqueiramente produzido para o jornalismo digital. Pressupõem o cumprimento de um método que parte de um problema e culmina com a geração de um produto cujos resultados preconizam a abertura e o compartilhamento de saberes com a divulgação da metodologia ou do código-fonte adotado, mantendo o modo de operação colaborativo (TRÄSEL, 2014). A transparência também é parte fundamental de práticas como o *fact-checking*, sendo um valor normativo e integrante da metodologia de apuração (SEIBT, 2019).

Derivado do Jornalismo de Precisão, o JGD está inserido em um contexto de informação e desinformação abundante no ambiente digital (BOCZKOWSKI, 2021), em que diversos tipos de conteúdo circulam principalmente por meio das plataformas (BELL; OWEN, 2017), dentro de um cenário de jornalismo em transformação (DEUZE; WITSCHGE, 2020) e jornalismo em rede (HEINRICH, 2011). No Brasil, autores como Mielniczuk (2003), Machado (2006) e Barbosa (2007) vêm registrando as transformações do jornalismo digital e o protagonismo dos dados, que perpassam desde o processo de produção até a circulação das notícias nesse ambiente. No JGD, os bancos de dados são vistos também como fonte (GEHRKE, 2018a), buscando monitorar e garantir a manutenção da transparência de instituições públicas na sociedade e ampliando o poder do repórter em abordagens junto às fontes pessoais



(MEYER, 2002). Ainda que conjuntos de dados públicas sejam oficiais, há mais condições para que sejam examinadas previamente, diferente do que ocorre com o discurso público, geralmente analisado *a posteriori* em iniciativas como o *fact-checking*.

Observando o desenvolvimento do Jornalismo Guiado por Dados desde 2016, quando ingressei no mestrado na UFRGS, percebo o aumento da procura por cursos, disciplinas, congressos e prêmios relacionados à área. Uma das principais expectativas relacionadas a essas buscas já haviam sido detectadas por Träsel (2014): no início dos anos de 2010, o JGD foi a resposta eleita por um grupo de profissionais brasileiros para contornar a crise identitária e econômica que permeava a profissão. À época, o pesquisador indicou a existência da interseção entre o jornalismo e outras áreas do conhecimento – o *ethos* do jornalista, então, era composto por aspectos também associados ao que se observava na computação, como a cooperação e a busca por dados abertos. Este tipo de conduta, ligado à disponibilidade de dados abertos e de códigos de programação, bem como auxílio e troca de informações para solução de problemas, indicava que a transparência poderia se desenvolver nesses espaços.

Além da própria pandemia, cujo rápido desenvolvimento exigiu o emprego de dados públicos com maior rapidez e assertividade no jornalismo, a desconfiança na imprensa e o excesso de desinformação em países como o Brasil despertou o meu interesse em observar como o Jornalismo Guiado por Dados, enquanto conjunto de saberes e práticas aplicado à reportagem e inspirado nas ciências sociais, estava tratando a questão. A Covid-19, como acontecimento jornalístico, parecia representar um período propício para analisar como o JGD se comportava diante da possibilidade de apresentar uma conduta mais aberta ao público – afinal, especialmente no começo da pandemia, havia limitações e incertezas quanto à disponibilidade de dados e até mesmo quanto à postura de instituições, como o Ministério da Saúde, na divulgação das informações. Por isso, decidimos investigar quais elementos de transparência apareciam na cobertura da imprensa brasileira sobre a Covid-19 – era definitivamente um momento de dúvidas para o Jornalismo Guiado por Dados, mas também de possibilidades.

Desde que ingressei na pesquisa em jornalismo, me alio à vertente que o considera uma forma de conhecimento que revela a realidade de forma diferente de outras práticas sociais (MEDITSCH, 1992; 1997), a meio caminho do senso comum e da ciência (PARK, 1940). Em trabalhos anteriores (GEHRKE, 2017), resgatei a teoria construída por Meyer (2002)

sobre o Jornalismo de Precisão para aproximar o Jornalismo Guiado por Dados de um conhecimento bastante próximo da ciência, que apresenta método, sistematização de resultados e transparência. Recentemente, revisitamos essa definição para acrescentar a discussão sobre pós-verdade e desinformação, defendendo que, enquanto modo de jornalismo informativo pertencente à categoria das notícias-sobre-relações (NIELSEN, 2017), o JGD tem potencial para gerar conhecimento a partir do acesso aos dados públicas, e a partir de novas formas de engajamento com o público (TRÄSEL; GEHRKE, 2021). Para tanto, não apenas narra os fatos, mas os põe em contexto, expondo aspectos sobre os fenômenos com os quais dialoga.

Defendo, nesta tese, que o JGD está alicerçado sobre três pilares fundamentais: a **verdade**, a **objetividade** e a **transparência**. Em resumo, significa dizer que se parte de uma aproximação com os fatos, por meio de técnicas e procedimentos estruturados de apuração – tal forma de acesso à realidade corresponde à objetividade como método – para então apresentar seus modos de pensar e fazer, ou seja, a transparência. A verdade verificável e provisória com a qual trabalha o jornalismo, e que a Covid-19 mostrou ser mais provisória do que nunca, pode ser apreendida e comunicada abertamente por meio do discurso.

Investigar a teoria associada à transparência nos fez classificá-la em dois eixos principais: como **operadora de uma prática ética** e **recurso discursivo de legitimação**. Enquanto o primeiro está associado a uma postura deontológica do jornalista, integrando a lista de princípios básicos com os quais o profissional deve se comprometer para executar suas práticas de rotina e manter o tratado de confiança com o público, o segundo discute a transparência como estratégia organizacional para que a empresa *demonstre* que é transparente, abrindo parte de suas políticas editoriais e fontes de financiamento. Neste caso, é claro, a empresa jornalística mostra somente aquilo que deseja, escolhendo as informações a que o público terá acesso e produzindo um *efeito de transparência*.

O aporte teórico somado ao percurso metodológico desta tese permitiu que eu confirmasse minha hipótese de que, à luz da verdade e da objetividade, **a transparência torna-se um valor fundamental para a prática do Jornalismo Guiado por Dados no Brasil**. O que se manifestava até então discursivamente nos últimos anos (TRÄSEL, 2014; GEHRKE, 2020), em uma esfera conectada ao *ethos* profissional, pôde ser verificado em termos de produto. A cobertura da Covid-19, por sua vez, serviu de mola propulsora para um processo que já estava em andamento. Em minhas pesquisas, observo que essa aceleração tomou forma por

iniciativa dos jornalistas, e não necessariamente das empresas. Isso porque, entre outros fatores, a concorrência entre jornais ainda é vista como um temor, especialmente entre as publicações na imprensa de referência. Em outras palavras, a colaboração que se constitui pela abertura das informações nem sempre é de interesse quando se trata de negócios, ainda que a transparência seja um valor defendido pelos jornalistas de dados.

Na etapa metodológica desta tese, para identificar os elementos de transparência empregados no Jornalismo Guiado por Dados, analisei materiais informativos (notícias, reportagens e projetos) publicados de março de 2020 a agosto de 2021 por 11 jornais de referência e nativos digitais, utilizando como metodologia a combinação do estudo de casos múltiplos (STAKE, 2006) e o estudo de caso como ilustração (MACHADO; PALACIOS, 2007) para acessar a cobertura da Covid-19. Para estudar como o fenômeno da transparência se manifesta na esfera do produto, busquei contemplar uma diversidade de veículos de comunicação e de datas de publicação dos materiais. Essa estratégia metodológica se mostrou bastante profícuca para conseguir uma fotografia ampla da cobertura do Jornalismo Guiado por Dados na pandemia, quando os conhecimentos abarcados por essas práticas foram tão exigidos.

Como resultados da tese e cumprimento dos objetivos específicos, que envolviam a discussão e o aprofundamento teórico da transparência, assim como o mapeamento de seus parâmetros, apresento os 15 elementos de transparência detectados na análise empírica, distribuídos nos níveis **organizacional e operacional**, e aqui dispostos em ordem alfabética: **assinatura, perfil e expertise; contato; correção de erros; data de publicação; documentos; fontes de dados; fontes de financiamento; graus de incerteza; marcas de apuração; manual da redação e princípios editoriais; metodologia; política de diversidade; política de privacidade; tipo de conteúdo; reprodutibilidade e código aberto.**

Embora o Jornalismo Guiado por Dados tenha uma abordagem metodológica significativa, não é possível descolar sua esfera operacional da organizacional. A apresentação de práticas transparentes em notícias e reportagens pressupõe uma estrutura empresarial minimamente disposta a compartilhar com os leitores como ocorre a tomada de decisões dentro da organização – seja pela apresentação do manual da redação, seja estabelecendo e demonstrando sua preocupação com a diversidade. Idealmente, espera-se que ocorra a abertura de informações sobre sua sustentabilidade financeira, permitindo que o leitor avalie eventuais conflitos de interesse, por exemplo.

Por isso, os níveis operacional e organizacional configuram a parte submersa da metáfora do iceberg. A **Figura 7** mostra que, normalmente, apenas parte das informações estão disponíveis à audiência em formato de notícia, mas os processos jornalísticos são mais complexos e profundos do que o produto, sendo permeados por decisões comerciais e editoriais. Mesmo que um repórter seja o responsável pela apuração de um conteúdo – e por isso assine o texto, requisito básico de práticas transparentes – os meandros relacionados à sua produção vão muito além dos desejos de um profissional, envolvendo um corpo de editores e diretores.

Os 15 parâmetros supracitados servem como guia para o que chamo de **Protocolo de Transparência no Jornalismo Guiado por Dados**, destinado à análise de produtos jornalísticos já publicados. Como se trata de uma primeira sistematização e proposição, busca auxiliar pesquisadores, jornalistas e quem mais estiver disposto a refletir sobre a implementação ou melhor aproveitamento do tema. Para servir como um guia, proponho uma série de perguntas para cada item, de modo que oriente seu uso. Além disso, alguns elementos são estruturantes e por isso podem ter maior peso em uma avaliação. Esse tipo de sistematização, conforme defendido por Karlsson (2022), aproxima a transparência do cumprimento de seu propósito principal – a aproximação entre jornal e audiência –, visto que permite a implementação de testes e quantificações para compreender que valores são priorizados e podem ser implementados ou aprimorados.

Em outro resultado da tese, ofereço a proposta de **Método para o Jornalismo Guiado por Dados com foco na transparência e na reprodutibilidade** para implementação em iniciativas jornalísticas orientadas por dados. Nesta proposta, sintetizada na **Figura 32**, o método científico (eixo vertical) é o fio condutor do passo a passo metodológico, enquanto a transparência (eixo horizontal) acompanha a consolidação de duas etapas fundamentais: o código aberto e a reprodutibilidade, além da documentação suplementar. Assim como o Protocolo, a Proposta de Método para o JGD aciona aspectos da transparência como **operadora de uma prática ética** e como **recurso discursivo de legitimação**. A proposta de método é compreendida por dez pontos: **1) Problema; 2) Hipótese; 3) Pesquisa; 4) Contexto; 5) Coleta; 6) Análise; 7) Código aberto e reprodutibilidade; 8) Produto; 9) Replicabilidade; e 10) Documentação suplementar**. Essa proposição busca suprir uma lacuna teórico-experimental: como identificado na tese, existem algumas tipologias de métodos que servem

como guia para a prática do Jornalismo Guiado por Dados em redações, mas as fórmulas existentes até o momento – embora também gerem contribuições para a área – não têm o compromisso ou a proposta de se guiar pelo método científico ou pela transparência. A formulação de hipótese, por exemplo, é abertamente defendida por poucos autores, a exemplo de Hunter *et al.* (2013) em um manual para a elaboração de projetos de jornalismo investigativo.

Entendo que os resultados da tese, estruturados pelas contribuições teóricas e empíricas que se materializaram na formulação do Protocolo e da Proposta de Método, geram contribuições inéditas e factíveis para que ocorra o avanço da transparência no Jornalismo Guiado por Dados. Ao cumprir premissas como as listadas em ambas as contribuições, compreendo que o JGD finalmente tem sólidas condições de se aproximar do que era preconizado por Meyer da metade para o final dos anos de 1960, com o Jornalismo de Precisão: uma forte e significativa aproximação com as premissas da ciência, utilizando hipótese, pesquisa, análise de dados, senso crítico e comunicação clara de resultados.

Por essa razão, busco, ao final, fazer uma última contribuição teórica pela via da aproximação do Jornalismo Guiado por Dados com os estudos de ciência aberta. Por um lado, os objetivos de ambos os conjuntos de práticas estão alinhados e dependem de protocolos para sua implementação e monitoramento de replicabilidade ou reprodutibilidade. Por outro lado, como alertado por Oliveira *et al.* (2021), todo tipo de padronização gera um risco de acentuar desigualdades, visto a aplicação depende das condições não apenas da empresa ou do microcenário onde se desenvolve, mas da situação de países ou regiões onde há limitado acesso ao conhecimento ou à tecnologia. No entanto, pensando na realidade brasileira e de tantos outros países em desenvolvimento, chego às minhas propostas com a consciência de que, a depender das circunstâncias, sua reprodução, na íntegra, dificilmente será possível. E isso não é um problema: a apropriação de teorias e metodologias requer maleabilidade e adaptação. Seu uso e transformação faz parte da construção de conhecimento inerente à ciência.

É inegável a existência de críticas à abordagem da transparência no jornalismo, como discutido ao longo da tese. Uma delas, desenvolvida por Han (2017), diz respeito aos perigos da dupla face transparência/vigilância, em que a privacidade das pessoas está sob risco em uma sociedade que cobra transparência excessiva de suas instituições. Outra crítica frequente tem relação com a dependência tecnológica de práticas jornalísticas, o que se acentua

no Jornalismo Guiado por Dados. Karlsson (2022), por exemplo, critica a opacidade algorítmica ligada ao uso de softwares sobre os quais pouco se tem conhecimento.

O Jornalismo Guiado por Dados busca contornar essas questões dando preferência ao uso de ferramentas de software livre e compartilhamento de código-aberto. Isso também se reflete na pesquisa: Diakopoulos (2019) propõe uma discussão sobre transparência algorítmica. Entretanto, justamente por questões de desigualdade de implementação do JGD, isso nem sempre é possível. O JGD criado e mantido nas favelas brasileiras (DE LIMA SANTOS; MESQUITA, 2021), por exemplo, é viabilizado apenas porque existem ferramentas gratuitas e de acesso massivo, como formulários online do *Google* e o aplicativo de mensagens *WhatsApp*. Por detrás dessas práticas, existem grandes *players* de tecnologia. Se não houvesse, porém, comunidades inteiras não teriam acesso aos dados de sua própria realidade para conhecer o seu entorno. Enquanto o jornalismo hiperlocal e local buscam formas de financiamento para garantir sua sustentabilidade, essa parece uma forma de começar a produzir reportagens guiadas por dados.

Assim, entendo que a abertura do que for possível ao escrutínio público e a apresentação do contexto de produção ainda parece ser uma alternativa mais viável do que simplesmente não produzir Jornalismo Guiado por Dados e argumentar que, dadas todas as limitações, tornar-se-ia inviável. A postura essencialmente crítica com relação à transparência não pode ser uma forma de escapar de sua aplicabilidade, ou de adotar uma conduta de que, se não é totalmente possível, não deveria sequer haver a tentativa de implementação. Em alguma medida, sempre haverá um tipo de dependência – ao software ou às fontes oficiais, e o Jornalismo Guiado por Dados não está livre disso. Mesmo que não seja possível que um jornal utilize linguagem de programação para criar produtos jornalísticos, há meios simples de começar a abrir seus processos, como apresentar a metodologia (por escrito) do que foi feito e indicar, por meio de hiperlinks externos, a origem das informações consultadas. Além disso, o detalhamento na correção de erros e a facilidade de comunicação por parte dos leitores tende a ser um recurso valorizado pela audiência (KARLSSON, 2022), e pode ser implementado pelos jornais para a manutenção da confiança e da credibilidade. É preciso que se pense em práticas viáveis no dia a dia, visto que não são apenas as grandes investigações jornalísticas que ganham atenção do público. É no jornalismo praticado diariamente que se estabelece a relação de confiança e honestidade com os leitores.

Como dito anteriormente, as contribuições que faço nesta tese integram um cenário pandêmico que abre precedentes para expectativa e internalização de processos de transparência no jornalismo. Pude observar, empiricamente, que alguns nativos digitais estão mais avançados nesse sentido, assentando suas práticas a partir de premissas de abertura metodológica e organizacional. Para outros veículos, porém, a continuidade desse processo dependerá de uma mudança de mentalidade em relação ao jornalismo digital, seus modos de narração e financiamento. Minha expectativa é de que tanto o Protocolo quanto a Proposta de Método possam ser explorados nesses espaços. Também espero que as contribuições sejam empregadas em sala de aula por docentes que queiram discutir criticamente os modos de fazer e de apresentar análises a partir de dados como fonte –o que não precisa ocorrer necessariamente em disciplinas específicas sobre Jornalismo Guiado por Dados; é possível inserir discussões sobre transparência e seus elementos em disciplinas de caráter teórico e prático. Espero, portanto, que esta tese gere contribuições também para o ensino do jornalismo. A partir de minha experiência docente desenvolvida durante o mestrado, o doutorado e mais recentemente como professora de MBA de Jornalismo de Dados, consegui observar que os alunos se sentem curiosos e confiantes para tratar deste tema e refletir sobre os seus próprios processos profissionais.

O emprego de dados por vias científicas no jornalismo, além de contribuir para cidadãos mais bem-informados, munidos de informações estatísticas sobre sua própria realidade e podendo praticar a democracia de forma mais consciente, serve de refúgio em tempos de desinformação. A transparência é a busca honestidade em tempos de todos os tipos de *fake*. Além disso, entendo que a abertura de métodos e processos pode reforçar a importância da verificação jornalística, uma vez que parte dos jornalistas checam com baixa frequência – ou mesmo utilizam de forma acrítica – suas fontes numéricas (LAWSON, 2021). A abertura dos dados e de informações adicionais, dessa forma, pode gerar uma maior preocupação dos profissionais com a qualidade e acurácia do conteúdo divulgado.

Naturalmente, minha tese não alcança todas as discussões sobre transparência. Uma vez que esse fundamento visa conceder à audiência acesso a informações normalmente opacas, entendo que estudos futuros poderão se debruçar justamente sobre questões relacionadas à percepção do público sobre os elementos de transparência citados nesta tese ou mesmo novos, que este trabalho não pôde abranger. A exemplo do que foi observado por

pesquisadores como Johnson e St. John (2021), alguns desses itens tendem a ser mais relevantes para manutenção da credibilidade percebida pelo público.

Finalmente, entendo que, em uma atmosfera constituída pelo excesso de informação e desinformação, é preciso enxergar a existência de formas especializadas de audiências – ou seja, públicos de nicho que se constituem em *audiências*, no plural, incluindo assinantes ou pessoas que contribuem esporadicamente para o financiamento de uma publicação. Considero, então, que minha tese serve como constituição e organização de um conhecimento basilar e com possibilidade de aplicação para novas pesquisas, em especial relacionadas às audiências que fazem parte do jornalismo digital.



## REFERÊNCIAS

ALLEN, David S. The trouble with transparency. **Journalism Studies**, Londres, v. 9, n. 3, p. 323-340, 2008.

ANGWIN, Júlia. Show your (computational) work. **NiemanLab**, 2021. Disponível em: <https://www.niemanlab.org/2020/12/show-your-computational-work/>. Acesso em 19 maio 2021.

AMAZEEN, Michelle. Revisiting the epistemology of fact-checking. **Critical Review: A Journal of Politics and Society**, v. 27, n. 1, p. 1-22, 2015.

ANDERSON, C. W.; BELL, Emily; SHIRKY, Clay. Jornalismo Pós-Industrial: adaptação aos novos tempos. **Revista de Jornalismo ESPM**, São Paulo, ano 2, n. 5, p. 30-89, 2013.

ANDERSON, C.W. **Apostles of certainty: data journalism and the politics of doubt**. Oxford: Oxford University Press, 2018.

ARQUEMBOURG, Jocelyne. Nascimento do acontecimento mediático moderno: o desastre de Lisboa. Dimensões do acontecimento: configuração, mediação, tempo e experiências. **Caleidoscópio Revista de Comunicação e Cultura**, Lisboa, v. 10, p. 41-58, 2011.

BABO-LANÇA, Isabel. Configuração mediática dos acontecimentos do ano. Dimensões do acontecimento: configuração, mediação, tempo e experiências. **Caleidoscópio Revista de Comunicação e Cultura**, Lisboa, v. 10, p. 73-84, 2011.

BACCIN, Alciane. A gramática da reportagem multimídia: organização e composição da informação. In: BACCIN, Alciane; SILVEIRA, Stefanie; BELOCHIO, Vivian (Org.). **25 anos de jornalismo digital no Brasil: a contribuição da pesquisadora Luciana Mielniczuk para os estudos no país**. Florianópolis: Insular, 2021. p. 119-139.

BARBON, Júlia et al. Cidades mais desenvolvidas do país vacinam mais rápido contra a Covid. **Folha de S.Paulo**, São Paulo, 4 jul. 2021. Disponível em: <https://www1.folha.uol.com.br/equilibrioesaude/2021/07/cidades-mais-desenvolvidas-do-pais-vacinam-mais-rapido-contr-a-covid.shtml>. Acesso em 20 out. 2021.

BARBOSA, Suzana. Jornalismo convergente e continuum multimídia na quinta geração do jornalismo nas redes digitais. In: CANAVILHAS, João (Org.). **Notícias e mobilidade: o jornalismo na era dos dispositivos móveis**. Covilhã: LabCom, 2013. p. 33-54.

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

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

BARNHURST, Kevin. The interpretative turn in news. In: SCHREIBER, Martin; ZIMMERMANN, Clemens (Org.). **Journalism and technological change: historical perspectives, contemporary trends**. Frankfurt: Campus Verlag, 2014. p. 111-141.

BAUER, Martin W.; AARTS, Bas. A construção do corpus: um princípio para a coleta de dados qualitativos. In: BAUER, Martin W.; GASKELL, George (Org.). **Pesquisa qualitativa com texto, imagem e som**. 13. ed. Petrópolis: Vozes, 2015. p. 39-63.

BEILER, Markus; IRMER, Felix; BREDA, Adrian. Data journalism at german newspapers and public broadcasters: a quantitative survey of structures, contents and perceptions. **Journalism Studies**, v. 21, n. 11, p. 1571-1589, 2020.

BELISÁRIO, Adriano *et al.* **O fluxo de trabalho com dados: do zero à prática**. São Paulo: Open Knowledge Brasil, 2020.

BELL, Emily; OWEN, Taylor. A imprensa nas plataformas: como o Vale do Silício reestruturou o jornalismo. **Revista de Jornalismo ESPM**, São Paulo, v. 6, n. 20, p. 48-83, 2017.

BENETTI, Marcia. O jornalismo como acontecimento. In: BENETTI, Marcia; FONSECA, Virgínia (Org.). **Jornalismo e acontecimento: mapeamentos críticos**. Florianópolis: Insular, 2010. p. 143-164.

BENETTI, Marcia. O jornalismo como gênero discursivo. **Galáxia**, São Paulo, n. 15, p. 13-28, 2008.

BENETTI, Marcia; STORCH, Laura; FINATTO, Paulo. Jornalismo de revista, meta-acontecimento e dispositivo de autoridade. In: LEAL, Bruno Souza; ANTUNES, Elton; VAZ, Paulo Bernardo (Org.). **Jornalismo e acontecimento: percursos metodológicos**. Florianópolis: Insular, 2011. p. 55-78.

BERGER, Peter; LUCKMANN, Thomas. **A construção social da realidade: tratado de sociologia do conhecimento**. 36. ed. Petrópolis: Vozes, 2014.

BUCCI, Eugênio. **Sobre ética e imprensa**. São Paulo: Companhia das Letras, 2000.

BULLA, Olívia. **A evidência dos números no discurso jornalístico através das relações de comunicação e trabalho**. Dissertação (Mestrado em Ciências da Comunicação), São Paulo, Universidade de São Paulo, 2015.

BOCZKOWSKI, Pablo J. **Abundance: on the experience of living in a world of information plenty**. Oxford: Oxford University Press, 2021.

BRADSHAW, Paul. Data journalism workflow: confronting the myths. In: MAIR, John; KEEBLE, Richard Lance (Org.). **Data journalism**: mapping the future. Bury St Edmunds: Abramis, 2013. p. 37-42.

BRADSHAW, Paul. The “second wave” of data journalism. In: MAIR, John et. Al (Ed.). **Data Journalism**: past, present and future. [S.l.]: Abramis, 2017. p. 244-253.

BRANDI, Daniel Pereira. **Agenda do meio ambiente no Jornal Nacional**: avaliação da qualidade da relevância temática. Dissertação (Mestrado em Comunicação), Universidade Federal de Sergipe, Aracaju, 2018.

BRENOL, Marlise Viegas. **Transparência digital e jornalismo**: modalidades comunicativas com uso de dados públicos. Tese (Doutorado em Comunicação e Informação), Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.

BRENOL, Marlise Viegas; TRÄSEL, Marcelo. Uso e apropriação de dados de saúde pelos jornais na pandemia Covid-19 no Brasil. In: VII SEMINÁRIO DE PESQUISA EM JORNALISMO INVESTIGATIVO. **Anais...** São Paulo: Abraji, 2020.

BROERSMA, Marcel. A refractured paradigm: journalism, hoaxes and the challenge of trust. In: BROERSMA, Marcel; PETERS, Chris (Ed.). **Rethinking Journalism**: trust and participation in a transformed news landscape. Nova York: Routledge, 2013. p. 28-44.

CAIRO, Alberto. **How charts lie**: getting smarter about visual information. Nova York: Norton & Company, 2019.

CANAVILHAS, João (Org.). **Webjornalismo**: 7 características que marcam a diferença. Covilhã: LabCom, 2014.

CAPELLA, Joseph N.; OPHIR, Yotam; SUTTON, Jazmyne. The importance of measuring knowledge in the age of misinformation and challenges in the tobacco domain. In: SOUTHWELL, Brian; THORSON, Emily; SHEBLE, Laura (Org.). **Misinformation and mass audiences**. Austin: University of Texas Press, 2018. p. 50-66.

CARNIELLI, Walter A.; EPSTEIN, Richard L. **Pensamento crítico**: o poder da lógica e da argumentação. São Paulo: Rideel, 2019.

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

CHRISTOFOLETTI, Rogério. O que pensam os jornalistas brasileiros sobre a transparência das suas práticas? **Revista Famecos**, Porto Alegre, v. 28, p.1-12, 2021.

CHRISTOFOLETTI, Rogério. Percepções de jornalistas brasileiros sobre privacidade. **Matrizes**, São Paulo, v. 13, n. 2, p. 179-197, 2019.

CODDINGTON, Mark. Clarifying journalism's quantitative turn: a typology for evaluating data journalism, computational journalism, and computer-assisted reporting. **Digital Journalism**, v. 3, n. 3, p. 331-348, 2014.

CONCONI, Augusto *et al.* Eventos de Bolsonaro durante a pandemia. **O Estado de S.Paulo**, São Paulo, 6 jun. 2021. Política. Disponível em: <https://arte.estadao.com.br/politica/2021/06/deslocamentos-jair-bolsonaro-pandemia/>. Acesso em 20 jul. 2021.

COULDRY, Nick; HEPP, Andreas. **The mediated construction of reality**. Cambridge: Polity Press, 2017.

CORNU, Daniel. **Ética da informação**. Bauru: EDUSC, 1998.

CORNU, Daniel. **Jornalismo e verdade**: para uma ética da informação. Lisboa: Instituto Piaget, 1994.

COSENZA, Ramon. **Por que não somos racionais**. Porto Alegre: ArtMed, 2016.

COSTA, Ana Beatriz Lemos da. **Transparência e seus avessos**: direito de acesso à informação no México. Dissertação (Mestrado em Comunicação), Universidade de Brasília, Brasília, 2017.

COSTERA-MEIJER, Irene; GROOT KORMELINK, Tim. Checking, sharing, clicking and linking. **Digital Journalism**, v. 3, n. 5, p. 664-679, 2015.

CRESWELL, John W. **Projeto de pesquisa**: métodos qualitativo, quantitativo e misto. Porto Alegre: Artmed, 2010.

DALMASO, Silvana; MIELNICZUK, Luciana. A participação cidadã em blogs como fator de ampliação da resolução semântica das notícias. **Interin**, Curitiba, v. 11, p. 1-12, 2011.

DAVIS, Chase. How academics can work with journalists (hint: they already have). In: BÉ-LAIR-GAGNON, Valérie; USHER, Nikki (Org.). **Journalism research that matters**. Oxford: Oxford University Press, 2021. p. 211-216.

D'IGNAZIO, Catherine; KLEIN, Lauren F. **Data Feminism**. Cambridge: The MIT Press, 2020.

DE LIMA SANTOS, Mathias Felipe. Tem #DDJBR aqui? Mapeando a presença do jornalismo de dados no Brasil. In: 42º CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO. **Anais...** Belém: Intercom, 2019.

DE LIMA SANTOS, Mathias Felipe; MESQUITA, Lucia. Data Journalism in favela: made by, for, and about forgotten and marginalized communities. **Journalism Practice**, p. 1-20, 2021.

DEMENECK, Ben-hur. **Jornalismo transnacional**: prática, método e conceito. Tese (Doutorado em Ciências da Comunicação), Universidade de São Paulo, São Paulo, 2016.

DEUZE, Mark; WITSCHGE, Tamara. **Beyond Journalism**. Cambridge: Polity Press, 2020.

DIAKOPOULOS, Nicholas. **Automating the news**: how algorithms are rewriting the media. Cambridge: Harvard University Press, 2019.

DIEGUES, Leonardo; FARIA, Flávia; TAKAHASHI, Fábio. Base de dados de cartórios traz falhas que impedem calcular efeito real do coronavírus no Brasil. **Folha de S.Paulo**, São Paulo, 13 maio 2020. Coronavírus. Disponível em: <https://www1.folha.uol.com.br/equilibrioesaude/2020/05/base-de-dados-de-cartorios-traz-falhas-que-impedem-calcular-efeito-real-do-coronavirus-no-brasil.shtml#>. Acesso em 1 jun. 2020.

DIENLIN, Tobias *et al.* An agenda for open science in Communication. **Journal of Communication**, n.71, p.1-26, 2021.

DONSBACH, Wolfgang. Journalism as the new knowledge profession and consequences for journalism education. **Journalism**, v. 15, n. 6, p. 661-677, 2014.

DOURADO, Tatiana M. S. G. **Fake news na eleição presidencial de 2018 no Brasil**. Tese (Doutorado em Comunicação e Cultura Contemporâneas), Universidade Federal da Bahia, Salvador, 2020.

EIDE, Martin. Accounting for Journalism. **Journalism Studies**, v. 15, n. 5, p. 679-688, 2014.

EGELHOFER, Jana *et al.* From novelty to normalization? How journalists use the term “fake news” in their reporting. **Journalism Studies**, v. 21, n. 10, p. 1323-1343, 2020.

EMPOLI, Giuliano Da. **Os engenheiros do caos**. São Paulo: Vestígio, 2020.

ESTEVANIM, Mayanna. **Processos no jornalismo digital**: do Big Data à visualização de dados. Dissertação (Mestrado em Comunicação), Universidade de São Paulo, São Paulo, 2016.

FENGLER, Susanne. Accountability in journalism. In: VOS, Tim; HANUSCH, Folker (Org.). **The International Encyclopedia of Journalism Studies**. [S. l.]: John Wiley & Sons, 2019.

FINK, Katherine; SCHUDSON, Michael. The rise of contextual journalism, 1950s-2000s. **Journalism**, Londres, v. 15, n. 1, p. 3-20, 2014.

FLECK, Giovana; MARTINS, Laís. Influenciadores digitais receberam R\$ 23 mil do governo Bolsonaro para propagandear “atendimento precoce” contra Covid-19. **Agência Pública**, São Paulo, 31 de março de 2021. Reportagem. Disponível em: <https://apublica.org/2021/03/influenciadores-digitais-receberam-r-23-mil-do-governo-bolsonaro-para-propagandear-atendimento-precoce-contracovid-19/>. Acesso em 18 maio 2021.

FOX, Jesse *et al.* Open science, closed doors? Countering marginalization through an agenda for ethical, inclusive research in Communication. **Journal of Communication**, online first, p.1-21, 2021.

FRANCISCATO, Carlos Eduardo. **A fabricação do presente**: como o jornalismo reformulou a experiência do tempo nas sociedades ocidentais. Aracaju: Editora UFS; Fundação Oviêdo Teixeira, 2005.

FRANCISCATO, Carlos Eduardo. Jornalismo, ciência e senso comum: contribuições do método científico para a reportagem jornalística. In: XV ENCONTRO DA COMPÓS, Grupo de Trabalho Estudos de Jornalismo. **Anais...** Bauru: Compós, 2006.

FREITAS, Fabiana Rossi da Rocha. **A visualização guiada por dados na TV**: o infográfico como efeito de realidade e elemento de articulação da narrativa telejornalística. Dissertação (Mestrado em Comunicação e Informação), Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018.

GADE, Peter *et al.* Management of Journalism transparency: journalists' perceptions of organizational leaders' management of an emerging professional norm. **International Journal on Media Management**, v. 20, n. 3, p.157-173, 2018.

GALISON, Peter. The journalist, the scientist, and objectivity. In: PADOVANI, Flavia; RICHARDSON, Alan; TSOU, Jonathan (Org.). **Objectivity in science**: new perspectives from science and technology studies. Springer, 2015. p. 57-75.

GAMBA, Estêvão; RIGHETTI, Sabine. Registros indicam que milhares no Brasil tomaram vacina vencida contra Covid; veja se você é um deles. **Folha de S.Paulo**, São Paulo, 2 jul. 2021. Coronavírus. Disponível em: <https://www1.folha.uol.com.br/equilibrioe-saude/2021/07/milhares-no-brasil-tomaram-vacina-vencida-contracovid-veja-se-voce-e-um-deles.shtml#erramos>. Acesso em 20 jul. 2021.

GANDOUR, Ricardo. **Jornalismo em retração, poder em expansão**: a segunda morte da opinião pública. São Paulo: Summus Editorial, 2020.

GEHLEN, Marco Antônio. **Jornalismo de (im)precisão**: o conhecimento matemático e a apuração de números. Tese (Doutorado em Comunicação Social), Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2016.

GEHRKE, Marília; BENETTI, Marcia. A desinformação no Brasil durante a pandemia de Covid-19: temas, plataformas e atores. **Fronteiras – estudos midiáticos**, v. 23, n. 2, p. 14-28, 2021.

GEHRKE, Marília; FREITAS, Fabiana. Mapa do contexto: um instrumento didático para investigações jornalísticas baseadas em hipótese. In: PÔRTO Jr, Gilson (Org.). **Comunicação e jornalismo**: metodologias para se pensar a docência, o ensinar e o pesquisar. Palmas: Editora EdUFT, 2021. p. 101-121.

GEHRKE, Marília. Jornalismo de dados e conhecimento científico: uma aproximação possível. In: 14º ENCONTRO NACIONAL DE PESQUISADORES EM JORNALISMO. **Anais...** Palhoça: Unisul, 2016a.

GEHRKE, Marília. O uso de dados na perspectiva do jornalismo como forma de conhecimento: a experiência do Diário Gaúcho. In: XXXIX CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO. **Anais....** São Paulo: Intercom, 2016b.

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

GEHRKE, Marília. O resgate da objetividade como método aplicado ao jornalismo guiado por dados. In: 15º ENCONTRO NACIONAL DE PESQUISADORES EM JORNALISMO. **Anais...** São Paulo: Universidade de São Paulo, 2017.

GEHRKE, Marília. **O uso de fontes documentais no jornalismo guiado por dados**. Dissertação (Mestrado em Comunicação e Informação), Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018a.

GEHRKE, Marília. Transparência no método como valor para o jornalismo. In: 16º ENCONTRO NACIONAL DE PESQUISADORES EM JORNALISMO. **Anais...** São Paulo: FIAM-FAAM/Universidade Anhembi-Morumbi, 2018b.

GEHRKE, Marília. Transparency as a key element of data journalism: perceptions of Brazilian professionals. In: COMPUTATION + JOURNALISM SYMPOSIUM 2020. **Anais...** Boston: Northeastern University, 2020.

GELAPE, Lucas; MARTINS, Laís. Ministério da Saúde camuflou dezenas de tweets sobre 'tratamento precoce'. **Núcleo**, São Paulo, 2 ago. 2021. Reportagens. Disponível em: <https://www.nucleo.jor.br/reportagem/2021-08-02-ms-tweets-tratamento-covid/>. Acesso em 18 out. 2021.

GENRO FILHO, Adelmo. **O segredo da pirâmide**: para uma teoria marxista do jornalismo. Florianópolis: Insular, 2012.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 2019a.

GIL, Antônio Carlos. **Métodos e técnicas de pesquisa social**. São Paulo: Atlas, 2019b.

GITELMAN, Lisa; JACKSON, Virginia. Introduction. In: GITELMAN, Lisa (Org.). **“Raw data” is an oxymoron**. Cambridge: The MIT Press, 2013. p. 1-14.

GOMES, Lucas; MAIA, Gabriel; SOUZA, Caroline. O cálculo de uma tragédia. **Nexo**, Rio de Janeiro, 8 ago. 2020. Reportagem especial. Disponível em:

<https://www.nexojornal.com.br/especial/2020/08/08/100-mil-mortes-no-Brasil-oc%C3%A1culo-de-uma-trag%C3%A9dia>. Acesso em 20 jul. 2021.

GOMES, Wilson. **Jornalismo, fatos e interesses**: ensaios de teoria do jornalismo. Florianópolis: Insular, 2009.

GRAVES, Lucas. **Deciding what's true**: the rise of political fact-checking in American journalism. Nova York: Columbia University Press, 2016.

GREEN, Melanie; DONAHUE, John. The effects of false information in news stories. In: SOUTHWELL, Brian; THORSON, Emily; SHEBLE, Laura (Org.). **Misinformation and mass audiences**. Austin: University of Texas Press, 2018. p. 102-115.

GUERRA, Josenildo. **O percurso interpretativo na produção da notícia**: verdade e relevância como parâmetros de qualidade jornalístico. Aracaju: Editora UFS; Fundação Oviêdo Teixeira, 2008.

GUERRA, Josenildo. Transparência editorial: a credibilidade jornalística à luz dos sistemas de gestão da qualidade. **Revista Latinoamericana de Ciencias de La Comunicación**, São Paulo, v. 11, n. 20, 2014.

GUIMARÃES, Hellen; BUONO, Renata. O 7x1 da Covid. **Revista Piauí**, Rio de Janeiro, 21 jun. 2021. Igualdades. Disponível em: <https://piaui.folha.uol.com.br/o-7x1-da-covid/>. Acesso em 20 jul. 2021.

HAHN, Oliver; STALPH, Florian. **Digital investigative journalism**: data, visual analytics and innovative methodologies in international reporting. Cham: Palgrave, 2018.

HAN, Byung-Chul. **A sociedade da transparência**. Petrópolis: Vozes, 2017.

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

HENN, Ronaldo. O ciberacontecimento. In: VOGEL; Deisi; MEDITSCH, Eduardo; SILVA, Gislene. **Jornalismo e acontecimento**: tramas conceituais. v. 4. Florianópolis: Insular, 2013. p. 31-48.

HOLOVATY, Adrian. **A fundamental way newspapers sites need to change**. [S. l. ], 2006. Disponível em: <<http://www.holovaty.com/writing/fundamental-change/>>. Acesso em: 30 abr. 2020.

HUNTER, Mark Lee et. al. **A investigação a partir de histórias**: um manual para jornalistas investigativos. Montevideú: Unesco, 2013.



JOHNSON, Kirsten A.; ST. JOHN, Burton. Transparency in the news: the impact of self-disclosure and process disclosure on the perceived credibility of the journalist, the story, and the organization. **Journalism Studies**, v. 22, n. 7, p. 953-970, 2021.

JOHNSON, Steven. **O mapa fantasma**: como a luta de dois homens contra o cólera mudou o destino de nossas metrópoles. Rio de Janeiro: Zahar, 2008.

KAKUTANI, Michiko. **A morte da verdade**: notas sobre a mentira na era Trump. Rio de Janeiro: Intrínseca, 2018.

KARLSSON, Michael. Dispersing the opacity of transparency in journalism on the appeal of different forms of transparency to the general public. **Journalism Studies**, v. 21, n. 13, p. 1795-1841, 2020.

KARLSSON, Michael; CLERWALL, Christer. Transparency to the rescue? Evaluating citizens' views on transparency tools in journalism. **Journalism Studies**, v. 19, n. 13, p. 1923-1933, 2018.

KARLSSON, Michael. Rituals of transparency: Evaluating online news outlets' uses of transparency rituals in the United States, United Kingdom and Sweden. **Journalism Studies**, v. 11, n. 4, p. 535-545, 2010.

KARLSSON, Michael. **Transparency and journalism**: a critical appraisal of a disruptive norm. Nova York: Routledge Focus, 2022.

KOLISKA, Michael; CHADHA, Kalyani. Transparency in German newsrooms. **Journalism Studies**, Londres, v. 19, n. 16, p. 2400-2416, 2018.

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

KOVACH, Bill; ROSENSTIEL, Tom. **The elements of journalism**: what newspeople should know and the public should expect. 3. ed. New York: Three Rivers Press, 2014.

LACY, Stephen; ROSENSTIEL, Tom. Defining and measuring quality journalism. **Media + the public interest initiative - news measures research project**. New Jersey: Rutgers School of Communication and Information, 2015.

LaFLEUR, Jennifer. Evangelizing for CAR. In: PAUL, Nora (Org.). **When nerds and words collide**: reflections on the development of Computer-Assisted Reporting. St. Petersburg: Poynter Institute, 1999. p. 25-27.

LANDERS, James. The National Observer, 1962-77: interpretive Journalism pioneer. **Journalism History**, v. 31, n. 1, 2005.

LAWSON, B.T. Hiding Behind Databases, Institutions and Actors: How Journalists Use Statistics in Reporting Humanitarian Crises. **Journalism Practice**, 2021.

LEON, Sam. Accounting for methods: spreadsheets, scripts and programming notebooks. In: Data Journalism Handbook. 2021. In: GRAY, Jonathan; BOUNEGRU, Liliana. **The data journalism handbook: towards a critical data practice**. Amsterdã: Amsterdam University Press, 2021. p. 146-158.

LISBOA, Sílvia; BENETTI, Marcia. O jornalismo como crença verdadeira justificada. **Brazilian Journalism Research**, v. 11, n. 2, p. 10-29, 2015.

LISBOA, Sílvia. **Jornalismo e a credibilidade percebida pelo leitor: independência, imparcialidade, honestidade, objetividade e coerência**. Dissertação (Mestrado em Comunicação e Informação), Universidade Federal do Rio Grande do Sul, Porto Alegre, 2012.

LISBOA, Sílvia; FRANZONI, Sabrina. Jornalismo e efeito de transparência: definições na teoria e na práxis. In: 16º ENCONTRO NACIONAL DE PESQUISADORES EM JORNALISMO, FIAM-FAAM/Anhembi-Morumbi. **Anais...** São Paulo: SBPJOR, 2018.

LIPPMANN, Walter. **Liberty and the news**. Nova York: Harcourt; Brace and Howe, 1920.

LOOSEN, Wiebke et. al. X Journalism: exploring journalism's diverse meanings through the names we give it. ECREA JOURNALISM STUDIES SECTION. **Anais...** Viena: ECREA, 2019.

LORENZI, Rodrigo Henrique Leite. **A visualização de informação no jornalismo nativo digital: uma análise do Nexa e do FiveThirtyEight**. Dissertação (Mestrado em Comunicação e Informação), Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.

LÜCKMAN, Ana Paula. **A noção de contexto no jornalismo: uma proposição a partir da teoria da complexidade**. Tese (Doutorado em Comunicação e Informação), Universidade Federal do Rio Grande do Sul, Porto Alegre, 2020.

LÜCKMAN, Ana Paula; FONSECA, Virgínia Pradelina da Silveira. Contexto e contextualização no Jornalismo: uma proposta conceitual. **Estudos em Jornalismo e Mídia**, Florianópolis, v. 14, n. 2, p. 162-174, 2017.

MACHADO, Caio *et al.* Ciência contaminada: analisando o contágio de desinformação sobre coronavírus via Youtube. [S.l.]: **Centro de Análise da Liberdade e do Autoritarismo (Laut)**, 2020a. Disponível em: <https://laut.org.br/ciencia-contaminada.pdf>.

MACHADO, Caio *et al.* Scientific isolation: international trends in misinformation and the departure from the scientific debate. [S.l.]: **Centro de Análise da Liberdade e do Autoritarismo (Laut)**, 2020b. Disponível em: <https://laut.org.br/wp-content/uploads/2020/11/Political-Self-Isolation-vF.pdf>

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

MACHADO, Elias; PALACIOS, Marcos. Um modelo híbrido de pesquisa: a metodologia aplicada pelo GJOL (2007). In: BARBOSA, Suzana; MACHADO, Elias; PALACIOS, Marcos (Org.). **GJOL 20 anos de percurso**. Textos fundadores e metodológicos. Salvador: EDUFBA, 2018. p. 341-363.

MACHLIS, Sharon. **Practical R for mass communication and journalism**. Boca Raton: Taylor and Francis Group, 2019.

MANDETTA, Luiz Henrique. **Um paciente chamado Brasil**: os bastidores da luta contra o coronavírus. Rio de Janeiro: Objetiva, 2020.

MARAS, Steven. **Objectivity in Journalism**: key concepts in journalism. Cambridge: Polity Press, 2013.

MARCELINO, Daniel. TrateCov indicava cloroquina e outros remédios sem eficácia a quase todos pacientes. **Jota**, Brasília, 21 jan. 2021. Dados. Disponível em: <https://www.jota.info/dados/tratecov-indicava-cloroquina-e-outros-remedios-sem-eficacia-a-quase-todos-pacientes-22012021>. Acesso em 18 maio 2021.

MARCONDES FILHO, Ciro. **Comunicação e jornalismo**: a saga dos cães perdidos. São Paulo: Hacker Editores, 2000.

MASSARANI, Luisa et. al. **O que os jovens brasileiros pensam da ciência e da tecnologia?** Pesquisa realizada pelo Instituto Nacional de Ciência e Tecnologia em Comunicação Pública da Ciência e Tecnologia (INCT-CPCT). Rio de Janeiro: Fiocruz, 2021. Disponível em: [https://www.inct-cpct.ufpa.br/wp-content/uploads/2021/02/LIVRO\\_final\\_web\\_2pag.pdf](https://www.inct-cpct.ufpa.br/wp-content/uploads/2021/02/LIVRO_final_web_2pag.pdf). Acesso em 30 mar. 2021.

MAZOTTE, Natália. Working Openly in Data Journalism. In: GRAY, Jonathan; BOUNEGRU, Liliana. **The data journalism handbook**: towards a critical data practice. Amsterdã: Amsterdam University Press, 2021. p.159-164.

McNAIR, Brian. Trust, truth and objectivity: sustaining quality journalism in the era of the content-generating user. In: BROERSMA, Marcel; PETERS, Chris (Org.). **Rethinking Journalism**: trust and participation in a transformed news landscape. Nova York: Routledge, 2013. p. 75-88.

MEDITSCH, Eduardo. **O conhecimento do jornalismo**. Florianópolis: Editora da UFSC, 1992.

MEDITSCH, Eduardo. O jornalismo é uma forma de conhecimento? **Biblioteca Online de Ciências da Comunicação**. Covilhã: UBI, 1997.

MEIER, Klaus. Transparency in journalism: credibility and trustworthiness in the digital future. In: THE FUTURE OF JOURNALISM. **Anais...** Cardiff: Cardiff University, 2009.

MELLO, Patrícia Campos. **A máquina do ódio**: notas de uma repórter sobre fake news e violência digital. São Paulo: Companhia das Letras, 2020.

MENEGAT, Rodrigo. Como a matemática pode ajudar a entender (e combater) epidemias. **O Estado de S. Paulo**, São Paulo, 13 mar. 2020. Especial Coronavírus. Disponível em: <https://www.estadao.com.br/infograficos/saude,como-a-matematica-pode-ajudar-a-entender-e-combater-epidemias,1082298>. Acesso em 20 abr. 2020.

MESQUITA, Ana Paula Soares. **A construção do sucesso e os contratos comunicacionais presentes nos portais Folha de S. Paulo e Estadão**. Dissertação (Mestrado em Comunicação e Semiótica), Pontifícia Universidade Católica de São Paulo, São Paulo, 2018.

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

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

MEYER, Philip. 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. St. Petersburg: Poynter Institute, 1999. p. 4-5.

MIELNICZUK, Luciana. **Jornalismo na Web**: uma contribuição para o estudo do formato da notícia na escrita hipertextual. Tese (Doutorado em Comunicação e Cultura Contemporâneas), Universidade Federal da Bahia, Salvador, 2003.

MIGUEL, Luiz Felipe. O jornalismo como sistema perito. **Tempo Social**, São Paulo, v. 11, n. 1, p. 197-208, 1999.

MORAN, Rachel. Subscribing to transparency: trust-building within virtual newsrooms on Slack. **Journalism Practice**, 2020.

NEWMAN, Nic. et al. **Digital News Report 2021**. Oxford: Reuters Institute for the Study of Journalism, 2021.

NIELSEN, Rasmus Kleis. Digital news as forms of knowledge: a new chapter in the sociology of knowledge. In: BOCZKOWSKI, Pablo J.; ANDERSON, C. W. (Org.). **Remaking the news**: essays on the future of journalism scholarship in the digital age. Cambridge: The MIT Press, 2017.

NIELSEN, Rasmus; GRAVES, Lucas. **“News you don’t believe”**: audience perspectives on fake news. Oxford: Reuters Institute for the Study of Journalism, 2017.

NOGUEIRA, Leidyanne Viana. **Interesse público em mídias legislativas**: um estudo da produção noticiosa da Agência Câmara Notícias. Dissertação (Mestrado em Comunicação), Universidade Federal do Ceará, Fortaleza, 2017.

O'CONNOR, Cailin; WEATHERALL, James Owen. **The misinformation age: how false beliefs spread.** London: Yale University Press, 2019.

O'NEILL, Onora. **A question of trust.** Cambridge: Cambridge University Press, 2002.

OLIVEIRA, André Rosa de. **Metadados como atributos da informação estruturada em bases de dados jornalísticas na Web.** Tese (Doutorado em Comunicação Social), Universidade Metodista de São Paulo, São Bernardo do Campo, 2016.

OLIVEIRA, Thaianie *et al.* Towards an Inclusive Agenda Science for Communication Research a Latin American approach. **Journal of Communication**, online first, 2021.

ORESQUES, Naomi; CONWAY, Erik. **Merchants of Doubt: how a handful of scientists obscured the truth on issues from tobacco smoke to global warming.** New York: Bloomsbury, 2010.

ORESQUES, Naomi. **Why trust science?** Princeton: Princeton University Press, 2019.

PARK, Robert. News as a form of knowledge: a chapter in the sociology of knowledge. **American Journal of Sociology**, Chicago, v. 45, n. 5, 1940. p. 669-686.

PATTERSON, Thomas. **Informing the news: the need for knowledge-based journalism.** Nova York: Vintage Books, 2013.

PENNYCOOK, Gordon *et al.* Fighting COVID-19 Misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. **Association for Psychological Science**, v. 1, n. 11, p. 1-11, 2020.

PEREIRA, Tathiany do Nascimento. **Governos (IN)visíveis e jornalismo mediador no Ceará: análise do uso de ferramentas de transparência pública como fontes de informação jornalística.** Dissertação (Mestrado em Comunicação), Universidade do Ceará, Fortaleza, 2019.

PHILLIPS, Angela. Transparency and the new ethics of journalism. **Journalism Practice**, Londres, v. 4, n. 3, p. 373-382, 2010.

PICKARD, Victor. **Democracy without journalism? Confronting the misinformation society.** Oxford: Oxford University Press, 2020.

PLAISANCE, Patrick Lee. Transparency: an assessment of the Kantian roots of a key element in media ethics practice. **Journal of Mass Media Ethics**, v. 22, n. 23, p. 187-207, 2007.

PONTES, Felipe Simão; SILVA, Gislene. Acontecimento jornalístico e história. In: BENETTI, Marcia; FONSECA, Virgínia (Org.). **Jornalismo e acontecimento: mapeamentos críticos.** Florianópolis: Insular, 2010. p. 43-61.

PORLEZZA, Colin; SPLENDORE, Sergio. Accountability and transparency of entrepreneurial journalism: unresolved ethical issues in crowdfunded journalism projects. **Journalism Practice**, v. 10, n. 2, p. 196-216, 2016.

POZEN; David E. SCHUDSON, Michael. Introduction. In: POZEN; David E.; SCHUDSON, Michael (Org.). **Troubling transparency: the history and future of freedom of information**. Nova York: Columbia University Press, 2018.

POSETTI, Julie; BONTCHEVA, Kalina. **Disinfodemic: deciphering Covid-19 disinformation**. United Nations Educational, Scientific and Cultural Organization, 2020. Disponível em: <https://en.unesco.org/covid19/disinfodemic/brief1>

QUÉRÉ, Louis. Entre o fato e o sentido: a dualidade do acontecimento. **Trajectos**, Lisboa, n. 6, 2005.

RAMOS, Ruama Lorena Ferraz. **E-gov em Cachoeira Paulista - SP: canais de relação entre governo municipal e cidadãos em um município de pequeno porte**. Dissertação (Mestrado em Desenvolvimento Regional), Universidade de Taubaté, Taubaté, 2017.

REGINATO, Gisele. **As finalidades do jornalismo**. Florianópolis: Insular, 2019.

ROCHA, Camilo. Por que tanta gente ignora regras no pior momento da pandemia. **Nexo**, São Paulo, 29 de março de 2021. Expresso. Disponível em <https://www.nexojornal.com.br/expresso/2021/03/29/Por-que-tanta-gente-ignora-regras-no-pior-momento-da-pandemia>. Acesso em 18 ago. 2021.

RODRIGUES, Adriano Duarte. O acontecimento. In: TRAQUINA, Nelson (Org.). **Jornalismo: questões, teorias e “estórias”**. Florianópolis: Insular, 2016. p. 51-59.

RODRIGUES, Kelly de Conti. **Jornalismo de dados na web: estudo da produção de sentido na infografia do Blog do Estadão Dados e do La Nación Data Blog**. Dissertação (Mestrado em Comunicação) – Faculdade de Arquitetura, Artes, Comunicação e Design, Universidade Estadual Paulista (Unesp), Bauru, 2016.

ROGERS, Simon. **Facts are sacred**. Londres: Faber & Faber, 2013.

ROSENBERG, Daniel. Data before the fact. In: GITELMAN, Lisa; JACKSON, Virginia (Org.). **“Raw data” is an oxymoron**. Cambridge: The MIT Press, 2013. p. 15-40.

ROSSINI, Patrícia *et al.* Dysfunctional information sharing on WhatsApp and Facebook: The role of political talk, cross-cutting exposure and social corrections. **New Media & Society**, 2020.

ROSSINI, Patricia; KALOGEROPOULOS, Antonis. **News and Mis(information) about Covid-19 in Brazil**. Liverpool: University of Liverpool, 2021. Disponível em:

<https://www.liverpool.ac.uk/communication-and-media/research/groups/news-and-misinformation-covid-19-brazil/>. Acesso em 2 mar. 2021.

RUSBRIDGER, Alan. **News and how to use it: what to believe in a fake news world**. Edimburgo: Canongate, 2020.

SANTOS, Egicyane Lisboa Farias. **Qualidade jornalística: uma avaliação dos indicadores pluralidade e confiabilidade no jornalismo literário da Revista Piauí**. Dissertação (Mestrado em Comunicação), Universidade Federal de Sergipe, Aracaju, 2017.

SANTOS, Kassia Nobre dos. **Em busca da credibilidade perdida: a rede de investigação jornalística na era das Fake News**. Tese (Doutorado em Comunicação e Semiótica), Universidade Católica de São Paulo, São Paulo, 2018.

SANTOS JÚNIOR, Edson Gil. **O jornalismo da fonte como parte necessária de um jornalismo plural democrático: as rotinas da divulgação da Assembleia Legislativa do Paraná**. Dissertação (Mestrado em Jornalismo), Universidade Estadual de Ponta Grossa, Ponta Grossa, 2015.

SARTOR, Basilio Alberto. A noção de interesse público no jornalismo: dimensões conceituais. In: 16º ENCONTRO NACIONAL DE PESQUISADORES EM JORNALISMO. **Anais...** São Paulo: FIAM-FAAM/Universidade Anhembi-Morumbi, 2018.

SCHUDSON, Michael. **Descobrimo a notícia: uma história social dos jornais nos Estados Unidos**. Petrópolis: Vozes, 2010.

SEIBT, Taís; FONSECA, Virgínia. Transparência como princípio normativo do jornalismo: a prática de fact-checking no Brasil. In: VI Seminário de Pesquisa em Jornalismo Investigativo. **Anais...** São Paulo: Universidade Anhembi-Morumbi, 2019.

SEIBT, Taís. **Jornalismo de verificação como tipo ideal: a prática de fact-checking no Brasil**. Porto Alegre, RS. Tese (Doutorado em Comunicação e Informação). Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.

SHAPIRO, Ivor *et al.* Verification as a strategic ritual. **Journalism Practice**, v. 7, n. 6, p. 657-673, 2013.

SILVA, Gislene; MAIA, Flávia Dourado. Análise de cobertura jornalística: um protocolo metodológico. **Rumores**, São Paulo, n. 10, p. 18-36, 2011.

SIMON, Felix M. CAMARGO, Chico Q. Autopsy of a metaphor: the origins, use and blind spots of the infodemic. **New Media & Society**, p. 1-22, 2021.

SMITH, Rory; CUBBON, Seb; WARDLE, Claire. **Under the surface: Covid-19 vaccine narratives, misinformation and data deficits on social media**. [S. l.]: First Draft, 2020. Disponível

em: [https://firstdraftnews.org/wp-content/uploads/2020/11/FirstDraft\\_Underthesurface\\_Fullreport\\_Final.pdf?x76851](https://firstdraftnews.org/wp-content/uploads/2020/11/FirstDraft_Underthesurface_Fullreport_Final.pdf?x76851). Acesso em 1 nov. 2021.

SOARES, Felipe; RECUERO, Raquel. How the mainstream media help to spread disinformation about Covid-19. **M/C Journal**, [S. l.], v. 24, n. 1, 2021. Disponível em: <https://journal.media-culture.org.au/index.php/mcjjournal/article/view/2735>. Acesso em: 16 mar. 2021.

SOARES, Marcelo. Mortes entre caixas, frentistas e motoristas de ônibus aumentaram 60% no Brasil no auge da pandemia. **El País**, São Paulo, 5 abr. 2021. Pandemia de coronavírus. Disponível em: <https://brasil.elpais.com/brasil/2021-04-05/caixas-frentistas-e-motoristas-de-onibus-registram-60-a-mais-de-mortes-no-brasil-em-meio-ao-auge-da-pandemia.html>. Acesso em 18 maio 2021.

SOCIETY OF PROFESSIONAL JOURNALISTS. **SPJ Code of Ethics**. Indianapolis: Society of Professional Journalists, 2014.

SOUTHWELL, Brian G.; THORSON, Emily A.; SHEBLE, Laura (Org.). **Misinformation and mass audiences**. Austin: University of Texas Press, 2018.

SOUZA, Rose Mara Vidal de. **Lei de Acesso à Informação: um canal à fonte de notícia**. Tese (Doutorado em Comunicação Social), Universidade Metodista de São Paulo, São Bernardo do Campo, 2016.

SPAGNUOLO, Sérgio. Na verdade, a qualidade da conexão à internet parece ter melhorado na quarentena. **Núcleo**, São Paulo, 13 maio 2020. Disponível em: <https://nucleo.jor.br/urbe/2020-05-13-qualidade-internet-conexao-quarentena>. Acesso em 19 maio 2021.

SPONHOLZ, Liriam. **Jornalismo, conhecimento e objetividade: além do espelho e das construções**. Florianópolis: Insular, 2009.

STAKE, Robert E. **Multiple case study analysis**. New York: The Guilford Press, 2006.

STEENSEN, Steen. Journalism's epistemic crisis and its solution: disinformation, datafication and source criticism. **Journalism**, v. 20, n. 1, p. 185-189, 2019.

STEENSEN, Steen; WESTLUND, Oscar. **What is digital journalism studies?** Londres: Routledge, 2021.

TANDOC JR., Edson; LIM, Zheng Wei; LING, Richard. Defining "Fake news": a typology of scholarly definitions. **Digital Journalism**, v.6, n.2, p.137-153, 2018.

TOFF, Benjamin *et al.* **Listening to what trust in news means to users: qualitative evidence from four countries**. Trust in news project. Oxford: Reuters Institute, 2021.



THURMAN, Neil. Computational journalism. In: WAHL-JORGENSEN, Karin; HANITZSCH, Thomas (Org.). **The handbook of journalism studies**. 2nd edition. New York: Routledge, 2020. p. 180-195.

TRAQUINA, Nelson. **Teorias do jornalismo**: porque as notícias são como são. 2. ed. Florianópolis: Insular, 2005.

TRÄSEL, Marcelo. Continuidades e rupturas: relendo um texto fundamental sobre ciberjornalismo. **Esferas**, n. 17, 2020. p. 26-37.

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), Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2014.

TRÄSEL, Marcelo. Pseudojornalismo: proposta de um termo para substituir “notícias falsas”. In: XII Simpósio Nacional da ABCiber. **Anais...** Porto Alegre, 2019. Disponível em: [https://www.abciber2019.sinteseeventos.com.br/trabalho/view?ID\\_TRABALHO=53](https://www.abciber2019.sinteseeventos.com.br/trabalho/view?ID_TRABALHO=53). Acesso em 17 mar. 2021.

TRÄSEL, Marcelo; GEHRKE, Marília. Que tipo de conhecimento é gerado pelo jornalismo guiado por dados? In: BACCIN; Alciane; SILVEIRA, Stefanie; BELOCHIO, Vivian (Org.). **25 anos de jornalismo digital no Brasil**: a contribuição da pesquisadora Luciana Mielniczuk para os estudos no país. Florianópolis: Insular, 2021. p. 339-357.

TRÄSEL, Marcelo; LISBOA, Sílvia; REIS, Giulia. Indicadores de credibilidade no jornalismo: uma análise dos produtores de conteúdo político brasileiros. In: XXVIII ENCONTRO ANUAL DA COMPÓS. **Anais...** Belo Horizonte: Compós, 2018.

TRIELLI, Daniel; Diakopoulos, Nicholas. How HCI can enable reproducibility in Computational Journalism. HCI for Accurate, Impartial and Transparent Journalism: Challenges and Solutions workshop. [S.l.]: **CHI Conference on Human Factors in Computing Systems**, 2019.

TUCHMAN, Gaye. Objectivity as Strategic Ritual: An Examination of Newsmen's Notions of Objectivity. **American Journal of Sociology**, v. 77, n. 4, 1972, p. 660-679.

TUCHMAN, Gaye. **Making news**: a study in the construction of reality. Nova York: The Free Press, 1978.

UTH, Bernadette; BADURA, Laura; BLÖBAUM, Bernd. Perceptions of trustworthiness and risk: how transparency can influence trust in journalism. **Trust and Communication**, Springer Nature Switzerland, p. 61-81, 2021.

VIEIRA, Livia de Souza. **Parâmetros éticos para uma política de correção de erros no jornalismo online**. Dissertação (Mestrado em Jornalismo). Universidade Federal de Santa Catarina, Florianópolis, 2014.

VU, Hong Tien; SALDAÑA, Magdalena. Chilling effects of Fake News: changes in practices related to accountability and transparency in American newsrooms under the influence of misinformation and accusations against the news media. **Journalism & Mass Communication Quarterly**, v. 1, n. 21, 2021.

WAISBORD, Silvio. Fake health news in the new regime of truth and (mis)information. **Reciis**, v. 14, n. 1, p.6-11, 2020.

WAISBORD, Silvio. The 5Ws and 1H of digital journalism. **Digital Journalism**, v.7, n.3, p.351-358, 2019.

WARD, Stephen J. A. **Ethics and the media**. Cambridge: Cambridge University Press, 2011.

WARDLE, Claire; DERAKHSHAN, Hossein. **Information disorder: toward an interdisciplinary framework for research and policymaking**. Estrasburgo: Conselho Europeu, 2017.

WARDLE, Claire. **Understanding Information Disorder**. [S. l.]: First Draft News, 2019. Disponível em: [https://firstdraftnews.org/wp-content/uploads/2019/10/Information\\_Disorder\\_Digital\\_AW.pdf?x76701](https://firstdraftnews.org/wp-content/uploads/2019/10/Information_Disorder_Digital_AW.pdf?x76701).

WEINBERGER, David. Transparency: the new objectivity. **KM World**, Camden [s.n.], 28 ago. 2009. Disponível em: <http://www.kmworld.com/Articles/Column/DavidWeinberger/Transparency-the-new-objectivity-55785.aspx>. Acesso em: 9 abr. 2020.

WIHBEY, John P. **The social fact: news and knowledge in a networked world**. Cambridge: The MIT Press, 2019.

ZAMITH, Rodrigo. Transparency, interactivity, diversity, and information provenance in everyday data journalism. **Digital Journalism**, v. 7, n. 4, p. 470-489, 2019.

ZELIZER, Barbie. Why journalism is about more than digital technology. **Digital Journalism**, Londres, v. 7, n. 3, p. 343-350, 2019.

## APÊNDICE A – LEVANTAMENTO DO ESTADO DA ARTE

O levantamento do estado da arte tem o propósito de mapear e detectar teses e dissertações produzidas no Brasil e com as quais este projeto dialoga. As buscas efetuadas na Biblioteca Digital Brasileira de Teses e Dissertações<sup>142</sup> do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) mostram que existe uma lacuna teórica nos estudos que associam a transparência ao método jornalístico.

Quanto à especificidade do objeto desta tese, o Jornalismo Guiado por Dados, os resultados são limitados. Ao combinar os termos **transparência** e **jornalismo guiado por dados** no campo de buscas, há o retorno de apenas um resultado que se aproxima com a proposta deste estudo. A tese escrita por Brenol (2019) e defendida na UFRGS traz contribuições importantes sobre como os jornalistas que usam rotineiramente bancos de dados públicos percebem seu próprio trabalho na fiscalização do poder público. A transparência da qual fala a autora, porém, não está relacionada à clareza do método jornalístico, como neste trabalho, mas aos governos digitais e às informações que o poder público oferece como matéria-prima para a investigação.

A tese de Brenol (2019) dialoga com os portais de transparência de órgãos públicos e informações fornecidas via LAI. Como resultado de sua pesquisa, a autora listou quatro modalidades comunicativas estruturadas a partir de dados públicos: estática, interativa, multimídia e personalizada, sendo que essa última é a que mais utiliza recursos digitais e fornece uma experiência completa ao usuário, possibilitando o uso de dados para fazer suas próprias consultas. Ainda que os jornalistas ouvidos na pesquisa utilizem com frequência a LAI e suas formas de solicitação de dados, o estudo indica que o uso de documentos “vazados” continua sendo uma prática recorrente.

---

<sup>142</sup> A Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) reúne textos completos das teses e dissertações defendidas nas instituições brasileiras de ensino e pesquisa. A definição das palavras-chave de busca e as primeiras consultas a esse repositório ocorreram nos dias 6 e 9 de março de 2020, época de execução do projeto de qualificação desta tese, e foram atualizados em 27 de outubro de 2021. Nessa data, o portal apresentava um total de 708,8 mil documentos (entre teses e dissertações). Interessa-nos observar os resultados de trabalhos produzidos na última década. A pesquisa nesta base de dados foi realizada por meio do seguinte endereço: <http://bdttd.ibict.br/vufind/>.

Por se tratar de uma palavra empregada em diferentes contextos, ainda que remeta à abertura e à clareza de procedimentos, em oposição à opacidade e obscuridade dos processos, a transparência aparece em teses e dissertações na área da comunicação com diferentes facetas. Ao consultar a combinação dos termos **jornalismo** e **transparência** na biblioteca do Ibict, encontrei 42 resultados. Após uma primeira triagem realizada com a leitura dos resumos, elenquei 16 trabalhos que apresentavam relação com meu objeto de estudo. A seleção desses casos indica a existência de dois eixos que se destacam nos trabalhos: 1) a transparência como ideal jornalístico, por vezes associado ao valor de *accountability*, que prevê a abertura de processos e correção de erros para o resgate da credibilidade e o estabelecimento de uma relação de respeito com a audiência; 2) a transparência como característica de governos digitais que geram matéria-prima a partir da Lei de Acesso à Informação (LAI) e, conseqüentemente, servem de fonte para jornalistas.

Destaco, no primeiro eixo, a investigação de Demeneck (2016). O autor discute a formação de redações pouco convencionais além-fronteiras, o que chama de jornalismo transnacional, estruturado em formato de rede colaborativa e configurando um conjunto de práticas adotado para a análise simultânea de dados e documentos vazados. Ao exemplificar tal construção por meio de casos como o *OffshoreLeaks* e o *SwissLeaks*, o pesquisador diagnostica elementos de objetividade, transparência e ética *hacker*, altamente associada à cooperação que se espera da comunidade de jornalistas de dados. Neste contexto, as técnicas do Jornalismo Guiado por Dados servem não apenas como base operacional, mas renovam as expectativas em torno da função social da imprensa.

Ainda em relação ao primeiro tópico de transparência, pesquisadoras como Vieira (2014) e Santos (2017) nos mostram que este valor está associado à política de correção de erros de veículos jornalísticos – partindo do pressuposto de que a precisão está no cerne desta prática – para a construção de credibilidade e também está relacionado à qualidade do conteúdo produzido. Ao monitorar 1,5 mil notícias pelo período de 100 dias, Vieira (2014) descobriu dez tipos de falhas na demonstração de retificação das informações, incluindo falta de transparência. Santos (2017), por sua vez, estudou a confiabilidade e a pluralidade de reportagens publicadas pela *Revista Piauí*. Para tanto, apropriou-se do conceito de

*accountability*, do qual a avaliação de qualidade<sup>143</sup> pode ser considerada um parâmetro. O termo, sem tradução fiel para o português, está relacionado à “capacidade de averiguar e cobrar responsabilidade e transparência no cumprimento de determinadas obrigações, sejam elas de cunho legal ou moral” (SANTOS, 2017, p. 44). Também sob o prisma de qualidade e *accountability*, Brandi (2018) examinou a cobertura do Jornal Nacional sobre meio ambiente, descobrindo que o conteúdo factual tem prioridade se comparado à cobertura de temas amplos e relevantes do ponto de vista ambiental, como planejamento urbano e seu impacto.

Em contexto semelhante, Santos (2018) traz a transparência metodológica do *fact-checking* como instrumento de clareza de procedimentos e de combate à desinformação. Um dos pontos levantados pela autora está na transparência das fontes – considero este o principal elemento de clareza que o jornalista pode oferecer aos leitores –, pré-requisito da prática de checagem de fatos e que se materializa por meio da disponibilização (online, via hyperlink) de planilhas e outros documentos de acesso à proveniência das informações. Fundamental, ainda, é a clareza da fonte de financiamento de um veículo – fator citado pela autora e que garante a abertura de dados sobre o modo como o veículo de comunicação mantém sua sustentabilidade financeira, permitindo à audiência avaliar se existe conflito de interesses.

O segundo eixo evidencia as práticas de transparência pública e os dados abertos como fonte para os jornalistas, a exemplo do que ocorre via LAI. Em sua dissertação, Pereira (2019) investigou como os repórteres de dois jornais do Ceará incorporam ferramentas de transparência em sua rotina. No universo investigado, a autora descobriu que, em cada dez repórteres, dois utilizam informações públicas para a produção de notícias, principalmente na área da política. Pereira não chega a explorar em profundidade os conceitos de JGD, embora relate a existência de núcleos especializados no tema.

Outros trabalhos mapeados no estado da arte também se apropriam da Lei de Acesso à Informação como tema central, enfatizando o protagonismo da legislação ao servir de fonte para jornalistas (SOUZA, 2016). Há, também, um olhar sobre legislações semelhantes no mundo (COSTA, 2017) e sobre as relações entre os canais de comunicação de transparência

---

<sup>143</sup> Conforme os conceitos desenvolvidos no trabalho de Santos (2017), a qualidade atribuída a um veículo jornalístico está associada à sua capacidade de desenvolver métodos claros para a avaliação de desempenho do conteúdo, assim como utilizar parâmetros para medir boas práticas e manter-se atualizado quanto às pesquisas de métodos de avaliação para produtos jornalísticos.

dos governos no diálogo com a comunidade e com os jornalistas (SANTOS JÚNIOR, 2015; RAMOS, 2017; NOGUEIRA, 2017). O ponto acerca das legislações, no entanto, interessa menos a esta tese.

Ao combinar os termos **fonte** e **transparência** no campo de pesquisas do banco de teses e dissertações do Ibict, considerando que os dados podem ser tratados como fonte no jornalismo, 340 resultados são exibidos. No entanto, elenquei apenas três relacionados à área da comunicação e ao tema aqui estudado. Trata-se dos trabalhos de Pereira (2019), Santos Júnior (2015) e Souza (2016), mencionados anteriormente. A soma de termos torna-se mais relevante à medida em que a busca é mais específica. Os termos **fonte** e **jornalismo guiado por dados**, associados, trazem somente duas pesquisas: minha dissertação (GEHRKE, 2018a) e a pesquisa de Oliveira (2016). Meu trabalho caracterizou uma primeira aproximação com esses dois tópicos, que seguem sendo primordiais na discussão da tese, ainda que sob outros aspectos. Analisei, à época, fontes documentais empregadas em 60 notícias, classificando-as em três tipos: arquivo documental, estatística e reprodução. Meu intuito era olhar para as especificidades das fontes acionadas no JGD, que vão além dos bancos de dados públicos. A pesquisa de Oliveira (2016), por sua vez, investiga de que forma a adoção de metadados contribui para o desenvolvimento da informação jornalística em bases de dados, aproximando as ciências da Comunicação e da Computação.

Substituindo o termo **jornalismo guiado por dados** por **jornalismo de dados** na combinação com **fonte**, outros quatro resultados são encontrados. Desses quatro, somente dois correspondem à área da Comunicação. Além da tese de Souza (2016) sobre a Lei de Acesso à Informação, há a pesquisa doutoral de Gehlen (2016), que investigou a forma como os números são abordados no jornalismo. O autor descobriu que os jornalistas de dados vão além do uso corriqueiro de operações matemáticas simples, pois buscam um maior raciocínio matemático-quantitativo diante da disponibilidade de bancos de dados, que demandam o uso de estatística para a geração de conhecimento.

Ao utilizar apenas o termo **jornalismo de dados** no mecanismo de busca, o portal do Ibict exibe 16 resultados. Mesmo que o termo tenha sido pesquisado entre aspas, nem todos os resultados retornam o termo exato. Somente seis trabalhos de fato correspondem ao tema, sendo que a metade deles já foi abordada neste estado da arte. Dois deles (ESTEVANIM, 2016;

RODRIGUES, K., 2016) discutem infografia e visualização, e a pesquisa de Mesquita (2018) traz informações sobre estratégias comunicacionais nas redações.

O termo **jornalismo guiado por dados**, por sua vez, retorna oito trabalhos, dos quais três já foram mencionados neste estado da arte. Dentre os três estudos restantes, dois (FREITAS, 2018; LORENZI, 2019) têm como tópico principal a infografia e a visualização de dados, área ainda mais específica dentro dos estudos de Jornalismo Guiado por Dados e que não possui relação direta com o tema desta tese. Importa, aqui, a pesquisa de Träsel (2014), cuja contribuição é uma das pioneiras no Brasil em relação ao JGD. Em sua tese, o autor identificou características do perfil do *ethos* do jornalista de dados, bem como situou esta prática como resposta à crise profissional e desenvolvida a partir de métodos que permitem aproximá-la da objetividade jornalística. Assim, essa investigação se conecta com os tópicos de minha pesquisa por aspectos como a verdade (provisória e verificável), a objetividade do método e a transparência.

A combinação dos termos **objetividade** e **jornalismo** apresentou 4,2 mil retornos, provocando a necessidade de afunilar o objeto para melhor compreender em que lacuna teórica esta tese se encaixa. Somados os termos **objetividade** e **jornalismo guiado por dados** – a variação jornalismo de dados não apresenta resultados consistentes na mesma pesquisa –, surgem as teses de Träsel (2014) e Brenol (2019), sendo que apenas a primeira de fato apresenta as duas palavras-chaves e traz tal discussão.

O estado da arte permitiu perceber, de maneira geral, que a pesquisa em jornalismo, mais especificamente em se tratando de **Jornalismo Guiado por Dados e transparência**, ainda é inicial no Brasil. Durante a elaboração da tese, tomei conhecimento de duas dissertações sobre o tema – uma delas especificamente no contexto do Jornalismo Guiado por Dados, porém analisando o discurso de profissionais – defendidas nos programas de pós-graduação em Jornalismo e Comunicação na Universidade Federal de Santa Catarina (UFSC) e na Universidade Federal de Santa Maria (UFSM). Até o fechamento desta tese, porém, os trabalhos ainda não haviam sido publicados no Ibict.