

UNIVERSIDADE FEDERAL DE MATO GROSSO
FACULDADE DE COMUNICAÇÃO E ARTES
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E PODER

SILVANA BAZANI

**CARACTERÍSTICAS DO JORNALISMO DE DADOS NO CENTRO-OESTE: UM
ESTUDO SOBRE OS WEBJORNAIS *GAZETA DIGITAL* E *CORREIO BRAZILIENSE***

CUIABÁ-MT

2025

SILVANA BAZANI

**CARACTERÍSTICAS DO JORNALISMO DE DADOS NO CENTRO-OESTE: Um
estudo sobre os webjornais *Gazeta Digital* e *Correio Braziliense***

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Comunicação e Poder da Universidade Federal de Mato Grosso, como parte dos requisitos necessários à obtenção do Título de Mestre em Comunicação, na Linha de Política e Cidadania.

Orientadora: Profa. Dra. Fernanda Vasques Ferreira

CUIABÁ-MT

2025

Dados Internacionais de Catalogação na Fonte.

B362c Bazani, Silvana.

Características do jornalismo de dados no centro-oeste: [recurso eletrônico] : um estudo sobre os webjornais gazeta digital e correio brasileiro / Silvana Bazani. -- Dados eletrônicos (1 arquivo : 144 f., pdf). -- 2025.

Orientadora: Fernanda Vasques Ferreira.

Dissertação (mestrado) - Universidade Federal de Mato Grosso, Faculdade de Comunicação e Artes, Programa de Pós-Graduação em Comunicação, Cuiabá, 2025.

Modo de acesso: World Wide Web: <https://ri.ufmt.br>.

Inclui bibliografia.

1. jornalismo digital. 2. jornalismo de dados. 3. algoritmo. 4. centro oeste. I. Ferreira, Fernanda Vasques, *orientador*. II. Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

Permitida a reprodução parcial ou total, desde que citada a fonte.



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE MATO GROSSO
PRÓ-REITORIA DE ENSINO DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM [NOME DO PPG]

FOLHA DE APROVAÇÃO

TÍTULO: Características do Jornalismo de dados no Centro-Oeste: um estudo sobre webjornais Gazeta Digital e Correio Braziliense

AUTORA: MESTRANDA **SILVANA BAZANI**

Dissertação defendida e aprovada em 22 de outubro de 2025

COMPOSIÇÃO DA BANCA EXAMINADORA

1. Professora Doutora Fernanda Vasques Ferreira (Presidente da Banca/Orientadora/PPGCOM/UFMT)
2. Professor Doutor Thiago Cury Luiz (Examinador Titular Interno/PPGCOM/UFMT)
3. Professora Doutora Marcelli Alves da Silva (Examinadora Titular Externa/UFMA)
4. Professor Doutor Bruno Bernardo Araújo (Examinador Suplente/PPGCOM/UFMT)

Cuiabá, 22 de Outubro de 2025



Documento assinado eletronicamente por **FERNANDA VASQUES FERREIRA**, Docente da Universidade Federal de Mato Grosso, em 23/10/2025, às 15:42, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **THIAGO CURY LUIZ**, Docente da Universidade Federal de Mato Grosso, em 23/10/2025, às 15:55, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

À minha filha, Adelle Athena Bazani Ribeiro,
companheirinha nesta longa e desafiadora
jornada.

Ao meu lado incondicionalmente – nas salas de
aula até altas horas, nos silêncios e nas
inquietações – foi e tem sido presença viva de
amor, paciência e força.

Com seu olhar curioso, sua alegria, doçura e
coragem, ensina-me tanto quanto qualquer
mestre.

Esta conquista é nossa.

Com todo meu amor, por tudo o que você foi, é
e sempre será na minha vida.

AGRADECIMENTOS

Este trabalho é fruto de uma travessia intensa – feita de noites longas, silêncios cheios de dúvida, mas também de encontros, afetos e descobertas que me transformaram profundamente. Hoje, ao olhar para trás, meu coração se enche de gratidão por todas as pessoas e instâncias que fizeram parte desse percurso. Ao Programa de Pós-Graduação em Comunicação e Poder (PPGCOM), meu sincero agradecimento. Em um país em que tantas portas se fecham para quem deseja estudar e conciliar o trabalho com a vida acadêmica, encontrar um espaço que compreende e acolhe as necessidades dos profissionais da comunicação é, acima de tudo, um alento. Obrigada à Universidade Federal de Mato Grosso (UFMT) e, principalmente, ao corpo docente do PPGCOM por tornar isso possível.

À professora Dr^a. Fernanda Vasques Ferreira, por ter me orientado no percurso de pesquisa, minha gratidão. Suas palavras firmes, sua escuta e seu conhecimento me acompanharam como uma bússola nos momentos de dúvidas e incertezas. Ao professor Dr. Thiago Cury Luiz, coordenador do PPGCOM, muito obrigada por me acolher em momentos de angústia, por responder às dúvidas com paciência e empatia, por estar presente de maneira tão humana. Suas atitudes fizeram diferença em momentos decisivos. Aos professores e às professoras que compartilharam seus saberes, com tanto zelo e dedicação: Dr^a. Tamires Ferreira Coêlho, Dr^a. Jociene Carla Bianchini F. Pedrini, Dr. Bruno Bernardo de Araújo, Dr. Luãn José Vaz Chagas, Dr. Rodrigo Levoti Portari, Dr. Pedro Pinto de Oliveira e Dr. Dôuglas Ferreira, meu reconhecimento e afeto. Em cada aula, em cada discussão, vocês deixaram marcas importantes no meu caminho. Aos amados amigos e amigas que estiveram comigo nessa jornada – do mestrado, da profissão e da vida –, obrigada por cada palavra de incentivo, por cada escuta atenta, por cada gesto de apoio nos momentos em que tudo parecia demais. Vocês foram minha rede e alento.

À minha família, meu abrigo e minha base. Ao meu pai, José Bazani, que já não está entre nós, mas permanece vivo em mim pelo exemplo de força, coragem e resiliência que deixou. À minha mãezinha, Laurita Prado Bazani, que me acompanhou com sua paciência e fé. Aos meus irmãos e irmãs, por estarem comigo em todos os momentos, mesmo à distância. À minha filha, Adelle Athena Bazani Ribeiro, que caminhou comigo de mãos dadas – literal e simbolicamente – durante toda essa jornada. Obrigada, meu maior amor da vida, por estar ao meu lado até tarde da noite nas aulas presenciais, por me apoiar, por me inspirar. Este trabalho é, também, para você. E à memória do seu pai, o jornalista e escritor Rodivaldo Araújo Ribeiro,

que nos deixou durante a pandemia de Covid-19, mas que sempre me incentivou a sonhar e acreditar que o mestrado era possível, minha eterna gratidão e homenagem. A todos que, de alguma forma, cruzaram meu caminho e deixaram um pouco de luz: muito obrigada. Ninguém caminha só. E eu sou profundamente grata por ter tido tanta gente especial ao meu lado.

RESUMO

Esta pesquisa objetiva identificar o uso de base de dados nas notícias e reportagens produzidas por dois webjornais mantidos por conglomerados de mídia na região Centro-Oeste do Brasil. Para isso, adotou-se uma abordagem conceitual que articula teorias no campo das novas mídias e do jornalismo contemporâneo, como o modelo Paradigma Jornalismo Digital em Base de Dados – JDBD (Barbosa, 2005; 2007; 2008; 2013) e suas potencialidades, sem perder de vista as transformações culturais, sociais e cognitivas trazidas pelas tecnologias digitais e pela internet (Lévy, 1998). Como procedimento metodológico aplicou-se a Análise de Conteúdo (Bardin, 2011; Sampaio e Lycarião, 2021). Para realizar a Análise de Conteúdo (AC), foram coletados e observados, por meio de pesquisa quali-quantitativa, publicações nos webjornais *Gazeta Digital*, do Grupo Gazeta de Comunicação, em Mato Grosso e *Correio Braziliense*, do Grupo Diários Associados S/A, no Distrito Federal. O *corpus* empírico foi constituído por conteúdos jornalísticos extraídos das capas dos veículos digitais mencionados durante 15 dias em 2023, caracterizando um recorte temporal aleatório. Considerando o levantamento realizado dentro desse período cronológico, foram reunidas 595 publicações do webjornal *Gazeta Digital* e analisadas 18 produções com viés informacional citando dados numéricos, caracterizadas como notícias e reportagens. Do *Correio Braziliense* foram identificadas 1.364 publicações informativas, sendo reunidas e analisadas 36 notícias e reportagens. A partir da coleta de 1.959 publicações e análise de 54 notícias e reportagens nos dois webjornais, foi identificado que, nas produções jornalísticas das versões digitais do *Gazeta Digital* e do *Correio Braziliense*, as bases de dados associadas a algoritmos não são elementos estruturantes da organização e narrativa noticiosa e não atuam como elemento central e direcionador de sistemas complexos para criação, manutenção, atualização, disponibilização e circulação de produtos jornalísticos dinâmicos, convergentes e multimídia. Portanto, partindo de pressupostos teóricos discutidos nesta pesquisa, a análise realizada evidencia que as 54 notícias e reportagens não incorporam plenamente as características e funcionalidades do Jornalismo Digital em Base de Dados (JDBD).

Palavras-chave: jornalismo digital; jornalismo de dados; algoritmo; Centro-Oeste.

ABSTRACT

This research aims to identify the use of databases in news and reports produced by two web newspapers maintained by media conglomerates in the Central-West region of Brazil. To achieve this, a conceptual approach was adopted that articulates theories in the field of new media and contemporary journalism, such as the Model Digital Journalism in Databases (DJDB) (Barbosa, 2005, 2007, 2008, 2013) and its potential, without losing sight of the cultural, social, and cognitive transformations brought about by digital technologies and the internet (Lévy, 1998). Content Analysis (Bardin, 2011; Sampaio and Lycarião, 2021) was applied as a methodological procedure. To conduct Content Analysis (CA), publications in the *Gazeta Digital* web newspapers, owned by Grupo Gazeta de Comunicação, in Mato Grosso, and *Correio Braziliense*, owned by Grupo Diários Associados S/A, in the Federal District, were collected and observed through qualitative and quantitative research. The empirical corpus consisted of journalistic content extracted from the front pages of these digital media outlets over a 15-day period in 2023, using a random timeframe. Considering the survey conducted within this chronological period, 595 publications from the *Gazeta Digital* web newspaper were collected, and 18 publications with an informational bias, citing numerical data, were analyzed, characterized as news and reports. From *Correio Braziliense*, 1,364 informative publications were identified, and 36 news and reports were collected and analyzed. Based on a collection of 1,959 publications and analysis of 54 news items and reports from the two online newspapers, we identified that, in the journalistic productions of the digital versions of *Gazeta Digital* and *Correio Braziliense*, databases associated with algorithms are not structuring elements of news organization and narrative, nor do they act as a central, guiding element of complex systems for the creation, maintenance, updating, provision, and circulation of dynamic, convergent, and multimedia journalistic products. Therefore, based on the theoretical assumptions discussed in this research, the analysis demonstrated that the 54 news items and reports do not fully incorporate the characteristics and functionalities of Digital Journalism in Databases (DJDB).

Keywords: digital journalism; data journalism; algorithm; Central-West.

LISTA DE ILUSTRAÇÕES

Figura 1 – Capa do <i>Gazeta Digital</i> em 01 de abril de 2001.....	23
Figura 2 – Capa do site <i>Gazeta Digital</i> em 21 de junho de 2023.....	23
Figura 3 – Foto da inauguração da atual sede do Grupo Gazeta de Comunicação, com a presença do ex-governador Dante Martins de Oliveira, em 23 de mai. 1998	24
Figura 4 – Página da primeira edição do jornal <i>A Gazeta</i> , de 23 de mai. de 1990.....	25
Figura 5 – Página do jornal <i>A Gazeta</i> em 13 de maio de 2024 alusiva aos 21 anos do site <i>Gazeta Digital</i>	28
Figura 6 – Primeira página da edição inaugural do <i>Correio Braziliense</i> ou <i>Armazém Literário</i> , em junho de 1808, fundado por Hipólito José da Costa	30
Figura 7 – Edição digitalizada da publicação impressa do diário <i>Correio Braziliense</i> em 21 de junho de 2023, disponível no site e pelo aplicativo <i>CB Digital</i>	43
Figura 8 – Edição digitalizada da publicação impressa do diário <i>A Gazeta</i> em 21 de junho de 2023, disponível no site e pelo aplicativo <i>Gazeta Digital</i>	44
Figura 9 – Cópia de mensagem eletrônica da <i>Internet Archive</i>	80
Figura 10 – Captura de tela de notícia publicada pelo <i>Gazeta Digital</i>	94
Figura 11 – Análise de conteúdo do <i>Gazeta Digital</i>	104
Figura 12 – Análise de conteúdo do <i>Correio Braziliense</i>	125
Figura 13 – Resultado da análise de conteúdo de <i>Gazeta Digital</i> e <i>Correio Braziliense</i>	129

LISTA DE QUADROS

Quadro 1 – Metáforas jornalísticas clássicas e metáforas futuras	39
Quadro 2 – Características geracionais do jornalismo digital	48
Quadro 3 – Diferenciações do jornalismo baseado em dados	72
Quadro 4 – Livro de códigos categorial – Unidades de registro	78
Quadro 5 – Livro de códigos categorial – Unidades de análise	80
Quadro 6 - Datas e dias de coletas	81
Quadro 7 – Períodos sem capturas pelo <i>Wayback Machine</i>	89
Quadro 8 – Análise de conteúdo do <i>Gazeta Digital</i>	101
Quadro 9 – Análise de conteúdo do <i>Correio Braziliense</i>	123

SUMÁRIO

INTRODUÇÃO	13
1 CONTEXTUALIZAÇÃO DO JORNALISMO REGIONAL	18
1.1 <i>Gazeta Digital</i>	22
1.2 <i>Correio Braziliense</i>	29
2 EVOLUÇÃO DO JORNALISMO DIGITAL.....	32
2.1 Jornalismo digital de primeira a quinta geração e o Jornalismo de Dados	34
2.2 Transposição e simbiose geracional	38
2.3 Jornalismo digital no Centro-Oeste: características gerais.....	47
3 DADOS NO JORNALISMO: ORIGENS E PERSPECTIVAS	49
3.1 Fundamentos conceituais do mundo orientado por dados	52
3.2 Jornalismo baseado em dados: na busca por conceitos e nomenclaturas	60
3.3 Rotinas produtivas do jornalismo baseado em dados.....	65
3.4 Jornalismo de dados e a Lei de Acesso à Informação (LAI)	70
4 PROCEDIMENTOS METODOLÓGICOS	73
5 ANÁLISE DO USO DE DADOS EM NOTÍCIAS E REPORTAGENS DE <i>GAZETA DIGITAL E CORREIO BRAZILIENSE</i>	83
5.1 Identificação do uso de dados em notícias e reportagens do <i>Gazeta Digital</i> (MT)..	88
5.2 Análise do uso de dados em notícias e reportagens do <i>Correio Braziliense</i> (DF)..	106
5.3 Análise comparativa e resultado geral do uso de dados em notícias e reportagens do <i>Gazeta Digital</i> (MT) e <i>Correio Braziliense</i> (DF).....	127
CONCLUSÃO	132
REFERÊNCIAS	136

INTRODUÇÃO

A informatização da sociedade impôs mudanças aos métodos de produção, distribuição, circulação e consumo de conteúdos noticiosos que convergem atualmente em meios multimodais, configurando um cenário de mutações e reconfigurações profissionais e da produção jornalística.

Ao longo da história da humanidade, a troca de informações entre indivíduos tem sido realizada com apoio de diferentes suportes. Enquanto os romanos fixavam folhas de notícias denominadas *Acta Diurna* em lugares públicos como forma de compartilhar as decisões tomadas, os chineses e coreanos utilizavam tipos móveis e papel de impressão centenas de anos antes do surgimento da prensa mecânica na Europa. Nos séculos XVII, XVIII e XIX, a imprensa foi o principal meio de informação, condição favorecida pelo uso da tipografia com tipos móveis, tecnologia criada por Johannes Gutenberg. Sua invenção provocou uma grande revolução na disseminação de informação através da reprodução de folhas impressas, ampliando, em certa medida, o acesso a notícias e, em partes, proporcionando condições para a alfabetização das massas. Em 1830, em Nova Iorque, surgiu o primeiro veículo de comunicação de massa, na forma do jornal de “um tostão”, batizado de *Penny Press*, que alcançou outros países em pouco tempo. No mesmo ano, o processo gráfico evoluiu para a impressão rápida. Na sequência, em meados do século XIX, foi inventado o telégrafo, propiciando um arranque na tecnologia dos veículos eletrônicos de comunicação de massa.

No século XX houve o mais significativo desenvolvimento de técnicas comunicacionais. Durante a primeira metade do século, o cinema se tornou uma forma de diversão familiar. Na década de 20, surgiu o rádio doméstico e, na década de 40, a televisão chegou aos lares. O rádio e a televisão atingiram o ponto de saturação de penetração nas residências norte-americanas nas décadas de 50 e 60, respectivamente (Araújo; Souza, 2003). No Brasil, a chegada da televisão na década de 1950 e sua popularização nas décadas seguintes colocou em xeque a sobrevivência do rádio. Como se dá com toda nova mídia ou suporte que desponta, sempre houve o temor de substituição ou de aniquilamento da mídia hegemônica anterior. Embora essas discussões estejam presentes na literatura, a trajetória dos meios de comunicação ao longo da história das civilizações revela constantes alterações e adaptações às transformações nos modos de criação, disseminação e recepção da informação jornalística.

Assim, a sociedade vem passando por transformações profundas no âmbito da comunicação. Com o surgimento da televisão, do rádio e do telefone, por exemplo, houve uma extensão da capacidade de entendimento dos seres humanos. No período, o filósofo canadense

Marshall McLuhan (1964) constatou que o progresso tecnológico estava transformando todo o planeta e previu uma revolução sendo delineada, motivada pelos novos meios eletrônicos de comunicação de massa, especialmente pela televisão, trazendo de volta a tribalização ou retribalização, com quase todos os seres humanos ao alcance de qualquer um, numa aldeia global.

Com a invenção dos computadores e, principalmente, da internet, tornou-se possível a conexão em rede e simultânea, em todo o planeta. A internet foi desenvolvida para proporcionar a extensão de capacidades naturais dos seres humanos, rompendo paradigmas até então consolidados. Uma contínua evolução de mecanismos tecnológicos trouxeram a humanidade à era do *Big Data*, iniciada com o advento dos computadores em 1946 e da *World Wide Web* (WWW) em 1969 nos Estados Unidos e, posteriormente, em 1980 no Brasil, originando um vasto conjunto de dados. A popularização de uma gama de aparelhos eletrônicos, como computadores e *smartphones*, interconectados a uma rede digital de alcance mundial que permite o compartilhamento instantâneo de dados e de conteúdos informacionais entre dispositivos, impactou o modo de vida da sociedade, criando um ambiente novo, reprocessando radicalmente o velho.

Práticas comunicacionais, políticas, relações sociais, culturais e econômicas vem sendo reconfiguradas de forma rápida e contínua pela circulação de informação no espaço virtual, que não existe em oposição ao real. O ciberespaço é o ambiente simbólico em que as comunidades virtuais se constituem. Por sua natureza plástica e elástica, o ciberespaço possibilita a combinação, articulação e integração de diferentes formatos, do textual ao não-textual, do imagético ao sonoro e vice-versa, em um fluxo contínuo de negociações intersemióticas (Ferrari, 2010).

O termo “ciberespaço”, de origem estadunidense, foi utilizado pela primeira vez em 1984 pelo escritor de ficção científica William Gibson, no romance *Neuromancer* (apud Lévy, 1998). Na obra, segundo Lévy (1998), o ciberespaço é descrito como o universo das redes digitais, um ambiente de encontros e aventuras, palco de conflitos globais e nova fronteira econômica e cultural. Trata-se de um campo vasto, aberto e ainda parcialmente indefinido, que não pode ser reduzido a apenas um de seus elementos. O ciberespaço é, por natureza, interconectado e capaz de se combinar com diversos dispositivos de criação, registro, comunicação e simulação. Ele transforma as formas de perceber, sentir, lembrar, trabalhar, jogar e conviver uns com os outros. Representa uma arquitetura interior, um sistema coletivo de inteligência ainda em construção. A gestão desse espaço, enquanto meio de comunicação e pensamento entre grupos humanos, será uma das grandes frentes de ação estética e política no

próximo século, conforme vislumbra Lévy (1998). A cibercultura, por sua vez, é um elemento da cultura contemporânea e consequência de uma sociedade influenciada pela tecnologia digital (Lemos, 2021; 2023).

Com a digitalização e, principalmente, a convergência de diferentes aparatos midiáticos, ampliaram-se as possibilidades de interação, reforçando liberdades individuais e coletivas, que são pilares da democracia. No entanto, essas transformações também trouxeram desafios à prática jornalística, como a propagação de desinformação e “misinformação”, resultado da descentralização dos agentes responsáveis pela produção de conteúdo nos ambientes digitais.

A desinformação no jornalismo consiste na elaboração e difusão deliberada de informações falsas ou enganosas, com o objetivo de manipular a opinião pública, prejudicar indivíduos, instituições ou grupos, ou ainda obter vantagens de ordem política, econômica ou social (Ito, 2024, p. 19). Suas principais características envolvem o uso de conteúdos falsos ou distorcidos, como mentiras explícitas, dados manipulados, informações fora de contexto e montagens audiovisuais, além da intencionalidade em enganar, o que a diferencia da misinformação (Wardle; Derakhshan; 2017), caracterizada pela veiculação de informações incorretas sem dolo, ou seja, de modo não intencional.

A desinformação é frequentemente propagada de forma estratégica por meio de redes sociais, *bots*, *sites* fraudulentos ou veículos de comunicação que desconsideram critérios éticos e editoriais. Seus efeitos comprometem a credibilidade da imprensa, desorientam o público e ameaçam a estabilidade democrática, sobretudo em períodos eleitorais ou de crise. Diante dessas circunstâncias, impõe-se a necessidade de respostas éticas e técnicas por parte do jornalismo, como o fortalecimento da checagem de fatos, a promoção da educação midiática e o compromisso com a transparência nos processos de apuração.

É nessa conjuntura que o jornalismo recobra seu estatuto, sua legitimidade e reforça a sua importância no sentido de informar com qualidade, sem censura, exercendo sua função social de assegurar o direito à informação como bem público e, ainda, gerir o conteúdo noticioso disponível na rede de forma responsável e ética (Bastos, 2015).

Para além disso, legislações recentes também reforçam o caráter público do jornalismo, ampliando a governança de dados, a transparência e o acesso de cidadãos às informações, consolidados pela Lei de Acesso à Informação (LAI), Lei nº 12.527/2011 (Brasil, 2011). No mesmo sentido, uma outra política pública também se fez necessária considerando o volume de dados produzidos e disponíveis. A Lei Geral de Proteção de Dados (LGPD), Lei nº 13.709/2018 (Brasil, 2018), foi criada como estratégia para proteger os dados pessoais dos cidadãos. Isso porque a cultura do computador potencializou mudanças de escala de processamento

computacional e o agigantamento da circulação global de informações (Kauffman; Santaella, 2021), impondo aumento vertiginoso de dados que alteram a comunicação e, consequentemente, reverberam na prática jornalística e também impactam a vida dos cidadãos.

Grandes conjuntos de dados que compõem o *Big Data* requerem adequações do jornalismo para extrair o máximo de informação e produzir narrativas em sinergia com as demandas da sociedade atual, caracterizada por uma audiência em constante mutação, mas ainda carente de políticas sociais mais justas e inclusivas. Esse cenário de ampliação de conjunto de dados culminou em pesquisas sobre Jornalismo de Dados (JD), Jornalismo Guiado por Dados (JGD) ou, ainda, Jornalismo Digital em Base de Dados (JDBD).

Diante da digitalização que se transmuta para uma dataficação, sendo faces de um amplo e ubíquo processo que reduz a complexidade e ocasionalidade da vida a códigos e dados digitais (Lemos, 2021; 2023; Segata; Rifiotis, 2021), afetando profundamente a comunicação, esta pesquisa busca identificar se e como webjornais da região Centro-Oeste estão se apropriando de bases de dados como fontes de informações e aplicando as técnicas do Jornalismo Digital em Base de Dados (JDBD).

Nesse contexto de incorporação de tecnologias emergentes, formula-se o problema de pesquisa: de que modo os webjornais *Gazeta Digital* e *Correio Braziliense* se apropriam de bancos ou bases de dados em suas práticas de produção jornalística digital, conforme o conceito de Jornalismo Digital em Base de Dados (Barbosa, 2005; 2007; 2008; 2013)?

Os dois veículos de comunicação selecionados para a composição do *corpus* como objeto empírico integram grupos midiáticos pioneiros na região central do Brasil. Com base em critérios de audiência, longevidade e mobilidade digital, definidos a partir das plataformas Guia de Mídia¹ e *SimilarWeb*², selecionamos os portais noticiosos *Gazeta Digital*, de Mato Grosso, e o *Correio Braziliense*, do Distrito Federal. A coleta dos conteúdos destacados nas capas destes dois portais informativos foi realizada mediante a utilização da ferramenta *Wayback Machine*³.

O objetivo geral é identificar como os bancos ou base de dados são utilizados pelos veículos jornalísticos selecionados para a produção de notícias e reportagens publicadas no ambiente digital.

Especificamente, busca-se: (1) revisar a literatura sobre Jornalismo de Dados e uso de bases de dados no jornalismo digital; (2) coletar e categorizar o *corpus*, verificando a presença

¹ Disponível em: <https://www.guiademidia.com.br>. Acesso em: 19 de set. 2025.

² Disponível em: <https://www.similarweb.com/>. Acesso em: 19 de set. 2025.

³ Wayback Machine é uma ferramenta digital criada pela organização Internet Archive que registra e preserva conteúdos de sites. Disponível em: <https://archive.org/web/>. Acesso em: 19 de set. 2025

de bases de dados, pluralidade de fontes e uso da Lei de Acesso à Informação (LAI); e (3) avaliar em que medida o uso dessas bases impacta a qualidade do jornalismo no Centro-Oeste.

Inicialmente, na etapa 1, buscamos compreender por meio de pesquisa exploratória e revisão bibliográfica o jornalismo digital, comunicação informatizada e Jornalismo de Dados. Na etapa 2 apresentamos e contextualizamos os veículos que compõem o *corpus*. Na etapa 3 foram realizadas coleta e análise de conteúdos, orientadas por categorias derivadas de Bardin (2011) e Sampaio e Lycarião (2021). Foram reunidas 595 publicações do *Gazeta Digital* e 1.364 do *Correio Braziliense*, ao longo de 15 dias distribuídos pela técnica de mês construído, no período de março a junho de 2023. O recorte temporal coincide com a ampliação do uso de inteligência artificial generativa e *machine learning*, bem como com o fortalecimento do debate sobre regulação das plataformas digitais, impulsionado pelo PL 2.630/2020. No cenário político, destaca-se ainda a mudança de governo em 2023, com o retorno à presidência do Brasil do candidato de esquerda, Luiz Inácio Lula da Silva (PT) que derrotou nas urnas o então candidato à reeleição presidencial pela direita, Jair Messias Bolsonaro (PL).

Para melhor organização e clareza do percurso investigativo, esta dissertação está estruturada em cinco capítulos, além desta introdução. O primeiro capítulo apresenta uma contextualização do jornalismo regional, com historicização do objeto empírico. O segundo capítulo discute os conceitos centrais relacionados ao jornalismo digital e sua evolução geracional. No terceiro capítulo, retomamos conceituações, origens e perspectivas do Jornalismo Digital em Base de Dados (JDBD). No quarto capítulo, descrevemos a metodologia adotada, detalhando o *corpus*, os procedimentos de coleta e análise, bem como as categorias utilizadas. O quinto e último capítulo traz a apresentação e análise dos resultados, articulando o material empírico às discussões teóricas. Nas considerações finais, apresentamos as contribuições do estudo, suas limitações e sugestões para pesquisas futuras, seguido das referências utilizadas.

1 CONTEXTUALIZAÇÃO DO JORNALISMO REGIONAL

O jornalismo vem se adaptando às mudanças tecnológicas, buscando sustentar sua função mediadora na sociedade (Barbero, 2008). Tais adequações são exigidas até mesmo dos veículos regionais, que operam fora dos grandes centros urbanos.

A cobertura midiática regionalizada tem se baseado em informação de proximidade, perpassada, às vezes, por distorções decorrentes da maneira como as relações de produção de notícias e de outros conteúdos midiáticos se processam, segundo Peruzzo (2005). Interesses comerciais, pressões políticas ou mesmo sociais podem influenciar no modo como as informações são selecionadas, apresentadas e interpretadas, levando a eventual cobertura tendenciosa, omissão de tópicos controversos ou agendamento de notícias que atendam a interesses de determinados grupos ou partidos.

Outra prática comum no jornalismo local, especialmente em cidades do interior, é o uso de *press-releases* produzidos por assessorias de comunicação dos poderes Executivo e Legislativo, bem como de instituições privadas. Esses comunicados fornecem informações pré-embaladas sobre eventos, questões ou iniciativas que podem ser de interesse público. Desde a escolha da pauta até a produção final da notícia, os *press-releases* podem influenciar o conteúdo, resultando em uma cobertura jornalística superficial, unidimensional e potencialmente tendenciosa. Ao receber um *press-release* sobre determinado assunto, os jornalistas podem basear a produção de notícia ou reportagem nas informações diretamente fornecidas pela assessoria de comunicação, em vez de realizar uma investigação mais aprofundada ou buscar fontes independentes. Isso resulta em uma cobertura que não aborda todas as nuances e os impactos de um evento, deixando de lado questões importantes ou perspectivas alternativas.

As assessorias de comunicação locais, como agentes diretos da notícia, fornecendo *releases* e informações mediadas pelos assessores, refletem uma realidade comum nas diversas áreas do jornalismo, especialmente em contextos em que recursos e acesso a múltiplas fontes podem ser limitados. Esse tipo de relação pode induzir a um “jornalismo declaratório”, decorrente da dependência excessiva de fontes oficiais, como assessorias de comunicação dos Poderes Constitucionais, bem como de instituições privadas.

Os jornalistas devem, contudo, estar alertas e cientes dessas dinâmicas, buscando realizar uma cobertura mais abrangente e apurada, diversificando suas fontes de informação, incluindo vozes da comunidade local, especialistas independentes e bases de dados. Desse modo, intentam a produção e divulgação de conteúdos noticiosos verossímeis e verídicos,

atendendo às necessidades e aos interesses do público, refletindo a complexidade dos acontecimentos e suas diferentes perspectivas, bem como atentando-se para a cobertura de temas complexos a partir de perspectivas complexas.

Nesse sentido, a coleta e análise de dados pelos jornalistas é uma prática que possibilita o acesso e a divulgação das informações de maneira mais aprofundada, ajudando a revelar padrões, tendências e anomalias que não seriam visíveis de outra forma. A prática do Jornalismo de Dados é interseccionada com o jornalismo investigativo. De acordo com Diakopoulos:

[...] um novo tipo de jornalismo computacional e de dados surge para investigar e responsabilizar a forma como o poder é exercido através de algoritmos. Chamo isso de “reportagem de prestação de contas algorítmica”, uma reorientação da tradicional função de sentinela do jornalismo, considerando exercícios de poder com algoritmos. Apesar de sua objetividade ostensiva, algoritmos não estão livres de erros e vieses que exigem maior escrutínio. Aos poucos, aprende-se a lidar com eles como algo inerente à prática jornalística, em conjunto com habilidades técnicas, de forma a prover o escrutínio mencionado acima (Diakopoulos, 2021, p. 220).

Apesar de distorções potenciais, Peruzzo (2005) argumenta que o jornalismo regional compreende melhor as nuances e os interesses locais e ajuda a fortalecer o tecido social, fornecendo um espaço para o diálogo e a troca de ideias. Ao destacar temas que afetam a realidade em microrregiões e dar voz às comunidades, o jornalismo contribui para criar um senso de identidade e coesão dentro de cada localidade, além de promover a transparência, a prestação de contas e a participação cívica. Atua como uma contraparte importante à grande mídia, fornecendo uma perspectiva única e detalhada sobre questões que podem não receber tanta atenção nos meios de comunicação de alcance nacional.

Embora seja importante reconhecer as limitações e desafios enfrentados pela mídia local, também é preciso valorizar sua contribuição na promoção do envolvimento comunitário e na disseminação de informações relevantes e atualizadas que são essenciais para a vida cotidiana das pessoas de um determinado território, sendo de base cultural, ideológica, idiomática, entre outras.

Com a facilidade de acesso a informações de todo o mundo decorrente do advento da internet⁴ e da *World Wide Web* (www)⁵, muitos acreditavam que a relevância das comunicações locais diminuiria. No entanto, o que se observou foi o oposto. Com o desenvolvimento da globalização da economia e das comunicações, num primeiro momento, chegou-se a pressupor

⁴ Estrutura que permite transferência de dados entre dispositivos interconectados em todo o mundo.

⁵ Conjunto de conteúdos que chegam até os dispositivos conectados à internet.

o fim da comunicação comunitária, para em seguida se constatar o contrário, ou seja, a revalorização, estimulando sua emergência em diferentes contextos e formas. Entre as razões que contribuem para ressignificar e valorizar a comunicação local estão a possibilidade de fortalecer a identidade e o senso de pertencimento, abordando questões que afetam o cotidiano das pessoas nas comunidades.

Como as questões locais têm maior impacto na vida das pessoas, informações detalhadas e análises específicas sobre assuntos delimitados a políticas municipais, eventos comunitários e infraestrutura são abordagens que se sobressaem no jornalismo regional. Além disso, jornalistas que atuam em conexão com as comunidades onde vivem têm um entendimento mais profundo das nuances e dinâmicas desses territórios, o que pode aumentar a confiança do público e a credibilidade no trabalho jornalístico. A comunicação regional também pode contribuir para o engajamento cívico e a representatividade das comunidades, fornecendo informações sobre oportunidades de participação na vida política e comunitária. Ao cobrir questões locais, o jornalismo pode inspirar ação e mudança dentro das comunidades sendo, portanto, um instrumento de fortalecimento da democracia e cidadania.

Apesar do potencial de revalorização e consolidação do jornalismo local em um mundo cada vez mais globalizado, o *Atlas da Notícia*, na sexta edição, divulgada em 2023, registrou queda na quantidade de iniciativas jornalísticas em todos os estados da região Centro-Oeste. No Distrito Federal foram 41 veículos fechados; em Mato Grosso, seis; em Mato Grosso do Sul, duas empresas; e em Goiás, uma. O total de veículos de comunicação que foram encerrados representam um por cento da base de dados da região no censo.

O principal motivo para o fechamento das empresas é a dificuldade em financiar e estabelecer modelo de negócio sustentável e lucrativo. No Centro-Oeste brasileiro ainda predominam alternativas tradicionais de venda de publicidade para órgãos públicos e privados. As empresas jornalísticas tiveram perda de lucratividade nesse modelo de negócio com a popularização de anúncios em redes sociais e *sites* de buscas, além da redução de repasse de verba pública para a imprensa nos últimos anos.

Desde os anos 1980, prevalece na região Centro-Oeste a cultura do consumo de informação noticiosa gratuita, inclusive no formato impresso. A prática segue com os *sites* de notícias locais em que o leitor não tem o hábito de contribuir financeiramente para a produção de conteúdo, segundo o censo *Atlas da Notícia* (2023).

Outra mudança detectada por meio do levantamento é que os jornalistas, a partir de suas realidades locais e como alternativa para a realização profissional, formam coletivos, associações, pequenas empresas e outras formas de organização para trabalhar. Além de

estabelecer novos formatos de iniciativas, os profissionais estão investindo em abordagens narrativas diferenciadas, como por exemplo a *Revista Badaró*⁶, criada com o objetivo de ser o primeiro veículo brasileiro focado na produção de jornalismo em quadrinhos e narrativas híbridas. Desde o início, a *Badaró* busca se manter de forma independente, financiada por pessoas que consomem o conteúdo produzido e por iniciativas parceiras. Atualmente, o conteúdo da revista é publicado de forma digital e impressa.

Um outro exemplo de projeto é o *TeatrineTV*⁷, site jornalístico independente dedicado a cobrir os diversos assuntos da arte e da cultura, tendo como objetivo principal investigar o uso da verba pública da cultura e oportunizar espaço para debate e promoção dos diversos gêneros artísticos produzidos na região Centro-Oeste. A proposta foge da cobertura convencional e da divulgação de agenda cultural.

Apesar da extinção de alguns veículos de comunicação, a região Centro-Oeste tem a menor porcentagem de desertos de notícias e o segmento que garante boa parte da cobertura é o rádio. São ao todo 729 rádios, 592 sites de notícias, 370 jornais impressos e 210 emissoras de televisão. As rádios estão pulverizadas e cobrem territórios como o Pantanal e o Cerrado, possibilitando que a informação chegue em municípios sem sinal de internet ou com acesso precário à rede.

O resultado geral do censo realizado em 2023 mostrou uma redução de 8,6% no total de desertos de notícias no Brasil. São 256 municípios a menos na conta – ainda assim, restam 2.712 cidades e 26,7 milhões de brasileiros que nelas habitam sem acesso a notícias sobre o lugar onde vivem. A redução dos desertos foi identificada num esforço coordenado com 303 colaboradores voluntários e estudantes de 80 organizações e universidades que se dedicaram a encontrar veículos de comunicação em lugares onde não havia registro de atividade jornalística e a atualizar um banco de dados de 14.444 organizações e iniciativas jornalísticas.

A expansão do digital e a identificação de rádios comunitárias que produzem conteúdo noticioso impulsionou a redução dos desertos. Em comparação com o mapeamento anterior, de 2022, o *Atlas da Notícia* acrescentou mais 575 iniciativas nativas digitais e 239 rádios à sua base de dados. Atualmente, há 5.245 veículos digitais e 4.836 rádios oferecendo notícias em seus canais. Juntos, os dois meios representam 70% do total de veículos mapeados pelo *Atlas* em 2023.

Considerando essas nuances e conforme apontamos anteriormente, esta pesquisa iniciou

⁶ <https://www.revistabadaro.com.br>

⁷ <https://www.teatrinetv.com.br>

comportando como objeto empírico quatro veículos de comunicação da região Centro-Oeste do Brasil. Dificuldades já mencionadas impuseram, porém, a necessidade de recortar o *corpus* e delimitá-lo a dois webjornais, *Gazeta Digital* e *Correio Braziliense*. Feito isso, finalmente foi possível que nos debruçássemos sobre os conteúdos noticiosos destacados nas capas dos webjornais apresentados a seguir.

1.1 *Gazeta Digital*

Sediado em Cuiabá, capital de Mato Grosso, o webjornal *Gazeta Digital* (GD) pontua em segundo lugar em audiência quando comparado com outros concorrentes potenciais ou emergentes, de acordo com a ferramenta *SimilarWeb*, em fevereiro de 2024. Contabilizando 1,4 milhão de visitas mensais, fica atrás apenas do *site* de notícias *Mídia News*, que alcançou 1,6 milhão de acessos e que já foi objeto de pesquisa⁸ relacionada ao Jornalismo Digital em Base de Dados (JDBD) no Programa de Pós-Graduação em Comunicação (PPGCOM) da Universidade Federal de Mato Grosso (UFMT).

Inaugurado no início dos anos 2000, com a primeira edição registrada pelo *Wayback Machine* em primeiro de abril de 2001, o portal *Gazeta Digital* (GD) nasceu como uma extensão do diário *Jornal A Gazeta*. Na transição até o modelo atual foram assimiladas algumas adaptações características da evolução geracional do jornalismo digital, conforme detalharemos no terceiro capítulo.

Atualmente, o GD mantém uma divisão editorial diferente da configuração do jornal *A Gazeta*. No *site* há, inclusive, uma editoria exclusiva sobre o tema coronavírus, totalizando 159 notícias desde o início das publicações, em 7 de março de 2021 até o momento, no ano de 2025.

⁸ FAVALESSA, M. **Jornalismo de Dados em veículos locais**: o uso de base de dados na cobertura da pandemia em Cuiabá (MT). 2023. Dissertação (Mestrado em Comunicação) Faculdade de Comunicação e Artes da Universidade Federal de Mato Grosso. Cuiabá, Mato Grosso, 2023.

Figura 1 – Capa da *Gazeta Digital* em 01 de abril de 2001



Fonte: captura de tela pela autora (Web Archive, 2001).

Figura 2 – Capa do site *Gazeta Digital* em 21 de junho de 2023



Fonte: captura de tela pela autora (Web Archive, 2023).

O portal de notícias GD pertence ao Grupo Gazeta de Comunicação, formado por outros meios como emissoras de rádios FM, emissoras de televisão, agência de publicidade, pelo instituto de pesquisa *Gazeta Dados*, gráfica e pelo jornal *A Gazeta*, que se tornou o periódico diário de maior circulação e com maior longevidade entre os atuais impressos de Mato Grosso, tendo sido fundado em 23 de maio de 1990 na capital mato-grossense Cuiabá. Na época, por ocasião de sua inauguração, reuniu diversas autoridades políticas, colocando-se como a primeira redação informatizada do Centro-Oeste.

Figura 3 – Foto da inauguração da atual sede do Grupo Gazeta de Comunicação, com a presença do ex-governador Dante Martins de Oliveira, em 23 de mai. 1998



Fonte: A Gazeta (1998).

Em sua primeira edição, o jornal *A Gazeta* destaca o fato de nascer informatizado. A “transformação e modernidade”, a “integração com a máquina”, “inovação em conceitos e critérios” e “independência jornalística” são ressaltadas nas páginas em preto e branco, atribuindo ao jornal *A Gazeta* a responsabilidade de treinar profissionais no exercício do jornalismo, no contexto de ausência de uma faculdade para tal em Mato Grosso (jornal *A Gazeta*, edição de 23 de maio de 1990, p. 2D, 2E e 3E). A Faculdade de Jornalismo, oferecendo a modalidade de bacharelado, da Universidade Federal de Mato Grosso (UFMT),

foi inaugurada em 5 de dezembro de 1990 como uma das habilitações do curso de Comunicação Social, seis meses depois da criação do jornal.

Figura 4 - Página da primeira edição do jornal *A Gazeta*, de 23 de mai. de 1990



Fonte: Barros (1990, p. 2E).

Costa (2019) registra que o processo de informatização das redações em Mato Grosso nos anos 1990 foi um marco significativo para o jornalismo regional. A chegada de computadores e sistemas informatizados revolucionou a produção de notícias, especialmente para os jornais impressos como *A Gazeta*. Essa modernização trouxe uma nova dinâmica para a produção jornalística, introduzindo tecnologia avançada e métodos mais rápidos e eficazes de impressão, como o sistema *offset*⁹.

Além da mudança tecnológica, essa fase também representou um período de transição para os jornalistas, tanto em termos de adaptação às novas ferramentas quanto em questões laborais. A implementação de tecnologia mais moderna demandava uma reestruturação nas redações e, por extensão, modificava a rotina dos jornalistas. As condições de trabalho eram uma questão importante, já que a velocidade e a eficiência trazidas pela informatização

⁹ Método comum de impressão que envolve a transferência de tinta de uma chapa para um cilindro de borracha (conhecido como *offset*) e então para o papel ou outro material de impressão.

pressionavam os trabalhadores por resultados mais rápidos e consistentes. Essa pressão, no entanto, não foi aceita sem resistência.

Jornalistas mais antigos, habituados ao sistema tradicional de linotipos e a métodos de produção mais artesanais, tiveram de se adaptar a esse novo ambiente tecnológico. A resistência à passagem geracional era compreensível, pois, para muitos profissionais, a transição para o ambiente informatizado representava não apenas um desafio técnico, mas também cultural, já que o “abraçar” dessas novas ferramentas implicava uma mudança significativa na maneira como o jornalismo era feito.

Para profissionais que cresceram em um ambiente dominado por tecnologias tradicionais, como a impressão por linotipo, a transição para um sistema informatizado significava mais do que aprender a usar computadores. Representava uma alteração na dinâmica das redações, com uma maior exigência por agilidade e a necessidade de dominar novas habilidades. Muitos jornalistas veteranos enfrentaram esse processo conflituoso de maneiras diferentes. Enquanto alguns profissionais mais antigos acolheram a mudança como uma evolução natural da indústria, outros resistiram, vendo nela uma ameaça à qualidade do jornalismo. A pressão para produzir conteúdo mais rapidamente e para estar em conformidade com os sistemas informatizados parecia contradizer as práticas jornalísticas tradicionais, que valorizavam o tempo para pesquisa e redação cuidadosa.

Essa resistência também estava relacionada ao medo de que a adoção irrestrita da tecnologia pudesse desvalorizar a experiência e a sabedoria acumuladas ao longo dos anos por esses jornalistas veteranos. Havia uma preocupação de que a informatização pudesse reduzir o jornalismo a uma produção rápida e superficial, em vez da realização de um trabalho artesanal que demandava tempo e dedicação. Em meio a conflitos e resistência, a informatização também trouxe oportunidades. Profissionais mais experientes que conseguiram se adaptar descobriram novas maneiras de aprimorar seu trabalho, aproveitando as ferramentas digitais para expandir a cobertura e alcançar públicos.

Apesar da crise econômica que o estado atravessava, tendo em vista o contexto do país com o Plano Collor, o jornal se mostrava financeiramente forte para abrir as portas do que iria se tornar um conglomerado midiático de impresso, rádio, TV e site, o Grupo Gazeta de Comunicação. O editorial que trata da comemoração de aniversário do veículo na edição de 23 de maio de 2021 destaca A Gazeta como “alicerce da verdade”, que “se dedica a absorver a explosão de conteúdos do dia anterior e a entregar ao leitor um material ‘redondo’ e tratado sob uma ótica mais analítica e profunda” (A GAZETA, 2021), reforçando valores jornalísticos considerados tradicionais, como “objetividade” e “neutralidade” (De Lima Jr.; Coêlho, 2023, p. 2).

No dia 13 de maio de 2024, em sua edição de número 11.457, o periódico jornal *A Gazeta* destaca as evoluções do portal *Gazeta Digital*, que completava 21 anos de existência. Vieira (2024) explica que, no início dos anos 2000, quando o acesso à internet ainda era majoritariamente discado e as primeiras redes sociais, como *MySpace*, *LinkedIn*, *Orkut* e *Facebook*, começavam a surgir, foi criado o webjornal *Gazeta Digital*, um dos pioneiros do jornalismo *online* em Mato Grosso. O portal integrou-se, desde sua origem, ao Grupo Gazeta de Comunicação, beneficiando-se da experiência de um conglomerado que reúne diversos veículos, entre eles o jornal *A Gazeta*, a TV Vila Real e a TV Pantanal, além de rádios como Gazeta FM, Vila Real FM, Cultura FM e afiliadas espalhadas por várias regiões do estado. Segundo o autor, o *site* se consolidou como um espaço de acesso integrado a esses meios, oferecendo versões digitais do jornal, conteúdos televisivos e transmissões radiofônicas ao vivo. A trajetória do portal, conforme afirma Vieira (2024), sustenta-se em uma cultura jornalística voltada à apuração rigorosa e ao compromisso com a veracidade, atributos que ganham ainda mais relevância em um contexto marcado pela disseminação de informações falsas. Além da cobertura de temas de interesse público, o *site* ampliou sua atuação para áreas como cultura, comportamento e atualidades, que hoje representam parte significativa de seu tráfego. O autor ressalta ainda que a permanência do *Gazeta Digital* como referência ao longo de 21 anos exigiu constantes adaptações tecnológicas, estéticas e editoriais. No início, os recursos eram bastante limitados, quando as publicações dependiam de conexão discada de 9.600 kbps e o *site* era construído apenas em linguagem HTML (*Hypertext Markup Language*), adequado a páginas estáticas. Com o tempo, foi necessário acompanhar o crescimento das redes sociais, expandindo o conteúdo para diferentes plataformas. Como exemplo dessa expansão, Vieira (2024) menciona que o portal reúne mais de 221 mil seguidores no *Facebook* e ultrapassa 88,7 mil seguidores orgânicos no *Instagram*.

Figura 5 – Página do jornal *A Gazeta* em 13 de maio de 2024 alusiva aos 21 anos do site *Gazeta Digital*

Escaneie o QR Code

Mais notícias aqui no *Gazeta Digital*

A GAZETA

Cuiabá, Segunda-feira, 13 de Maio de 2024

Vida

VIDA@GAZETADIGITAL.COM.BR

Redação do GD: perita profissional e gerações diversos garantem a comunicação eficiente com diferentes públicos

PIONEIRISMO E CREDIBILIDADE

GAZETA DIGITAL completa 21 anos em constante EVOLUÇÃO

LUIZ FERNANDO VIEIRA
EDITOR DO VIDA

No início dos anos 2000 a internet ainda era discada para a maioria das pessoas aquelas que tinham o privilégio de contar com uma conta de provedor -, comunidades como MySpace, LinkedIn, Orkut e Facebook estavam apenas engatinhando. Foi nesse contexto que nasceu um dos sites de notícias pioneiros de Mato Grosso, o *Gazeta Digital*, que completa 21 anos como um dos veículos de maior credibilidade do estado e conquista uma crescente legião de seguidores nas redes.

Integrante do Grupo *Gazeta de Comunicação*, o site já nasceu amparado pelo know how de um conglomerado que possui os mais importantes noticiários de Mato Grosso, como o jornal *A Gazeta*, *TV Vila Real* e *TV Pantanal*, além de rádios campeãs de audiência como *Gazeta FM*, *Vila Real FM*, *Cultura FM*, e afiliadas espalhadas por várias regiões do estado. É também um portal onde os internautas podem ter acesso a esses mesmo veículos, com versões digitais do jornal, matérias veiculadas nas TVs e links para ouvir as rádios ao vivo, em qualquer lugar do planeta onde haja sinal de internet.

Está, portanto, em seu DNA o espírito jornalístico que busca sempre a apuração aprofundada dos acontecimentos que impactam a vida da população mato-grossense, com credibilidade e respeito à verdade dos fatos, algo que tem conquistado cada vez mais importância em tempos de fake news. Sem esquecer das reportagens sobre cultura, comportamento e atualidades que hoje são responsáveis por grande parte dos acessos.

Para ser relevante durante estes 21 anos o *Gazeta Digital* teve também que se manter em constante evolução, tanto tecnológica como de visual e conteúdo. Para se ter uma ideia, no início, o máximo que se podia fazer em termos de plus para o visitante era disponibilizar o jornal impresso. Para "subir" as matérias era usada uma linha discada com conexão de 9.600 kbps (hoje é fácil contratar internetes que têm acima de 300 mbps) e tudo era feito somente em HTML (Hypertext Markup Language), bloco de construção mais básico da internet, próprio para materiais estáticos.

Ao longo dos anos foi preciso também acompanhar o crescimento das redes sociais e hoje o portal, além de estar integrado aos demais veículos do Grupo *Gazeta de Comunicação*, tem seu conteúdo expandido para as mais importantes, potencializando a conexão com o público e reforçando a presença nas principais mídias digitais. Para se ter uma ideia do tamanho dessa conexão, nada melhor do que números. Somente no Facebook são mais de 221 mil seguidores, e no Instagram já passa dos 88,7 mil seguidores orgânicos.

"Quando o *Gazeta Digital* nasceu, lá no começo dos anos 2000, a internet não tinha a velocidade que tem hoje, nem os canais de informação eram tão variados. Com o passar do tempo, fomos incentivados a acompanhar essa evolução não só na maneira de fazer o jornalismo, mas também para conquistar a atenção do leitor, que deixou de lado o computador 'caixotão' para ler as notícias na palma da mão", lembra a editora executiva *Jéssica Bachega*.

"Essa preocupação em assumir as novas imposições da maneira como consumir o conteúdo fez com que o site se atualizasse constantemente, não só em sua aparência, mas marcar presença forte das redes, que é onde o leitor está hoje. Com perfil no Instagram, Facebook, YouTube, *Giro de Notícias* na *TV Vila Real*, o grupo *Plantão Gazeta Digital* no WhatsApp e o canal direto com o leitor por meio do 'Mande Seus Whats', o *Gazeta Digital* se consolidou e se mantém numa posição de respeito e credibilidade entre os maiores portais de Mato Grosso", comemora.

DESAFIO DIÁRIO

"Nestes 21 anos de *Gazeta Digital* a maneira como as pessoas interagem e a forma como a notícia chegou até o leitor evoluíram rapidamente. As diversas plataformas de informação do que acontece no cotidiano também mudaram, mas algo que permanece desde as primeiras linhas do jornalismo é o compromisso com a checagem precisa e o texto apurado até que seja lido por quem acessa o portal", reforça a editora executiva.

Isso demanda uma atenção especial do portal. As redes sociais, que se multiplicam a cada dia, são um desafio para a equipe em busca de sempre levar o dado mais preciso e atualizado para os seguidores. "São muitos vídeos, textos, imagens compartilhadas tão rapidamente, sem contexto, que tornam nosso trabalho ainda mais minucioso em busca de dados e fontes para ter a certeza de que aquilo que se propaga é compatível com a verdade ou mais uma fake news a ser esclarecida antes que pessoas sejam prejudicadas por elas", explica *Jéssica*, citando como exemplo o que se vê agora no Rio Grande do Sul, quando desinformações sobre doativos e resgates dificulta a chegada da ajuda a quem precisa.

"São 21 anos noticiando fatos importantes da história de Mato Grosso, atendendo aos apelos da comunidade, denunciando irregularidades e também levando entretenimento ao leitor", salienta a editora executiva *Jéssica Bachega*.

SERVIÇO

Acesse o *Gazeta Digital* no endereço www.gazetadigital.com.br e acompanhe o site nas redes sociais com @gazetadigital.

APONTE O CELULAR E FALE COM O EDITOR

Fonte: Vieira (2014).

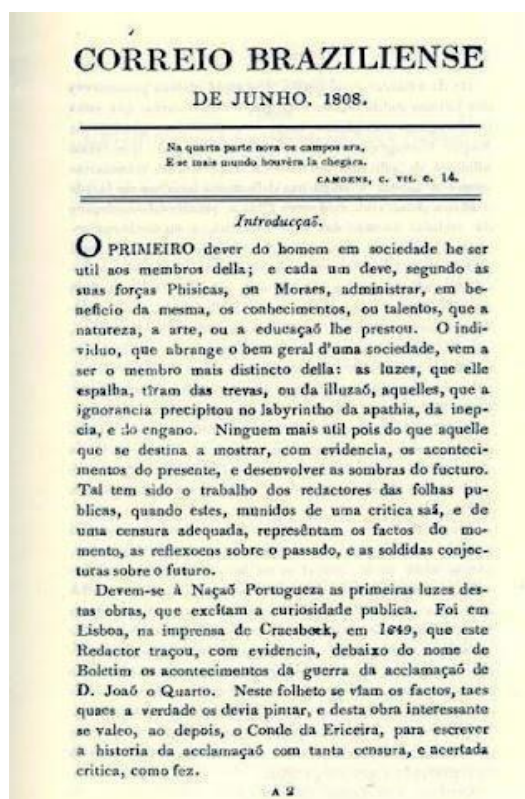
1.2 *Correio Braziliense*

Outro objeto empírico desta pesquisa, o *Correio Braziliense* tem um meio digital com 20,9 milhões de visitas em fevereiro de 2024, além de manter a versão impressa pertencente aos Diários Associados desde 1960, que abarca outros veículos de comunicação no Distrito Federal, estando sediado na capital, Brasília. Há dez anos, na data de primeiro de maio de 2014, o webjornal do *Correio Braziliense* anunciava reformulações no formato *online*, incluindo novos canais de acesso aos conteúdos publicados, que passaram a ser disponibilizados para assinantes por meio do aplicativo *CB Digital*, disponível para *smartphones* e *tablets*, além do acesso por computador. Conforme divulgado na época, ressaltava a adaptação do jornal impresso para formatos móveis.

Antes de lançar sua versão digital, o “novo” *Correio Braziliense* circulou apenas no formato impresso durante pouco mais de meio século. Seu fundador, Francisco de Assis Chateaubriand Bandeira de Melo ou, simplesmente, Assis Chateaubriand, decidiu lançar um jornal contemporâneo à criação da nova capital federal, Brasília (DF), fundada em 21 de abril de 1960, durante o governo de Juscelino Kubitschek.

Ao pensar no nome para o jornal, Chateaubriand se inspirou em Hipólito José da Costa, considerado o precursor da imprensa no Brasil por ter lançado, em 1808, o *Correio Braziliense* ou *Armazém Literário* (Cesar, 2006). Considerado o primeiro jornal brasileiro, apesar de ter sido publicado em Londres no período de 1808 a 1822, o periódico era crítico ao regime português e circulava clandestinamente no território brasileiro. Os exemplares chegavam entre 45 e 90 dias após a publicação, por causa do tempo de viagem marítima da época. Cada edição consistia em um volume com uma média entre 72 e 140 páginas, embora alguns números tenham ultrapassado 200 páginas. A estrutura do jornal era dividida em quatro seções gerais: Política; Comércio e Artes; Literatura e Ciência; e Miscelânea; com subcategorias de Correspondência e Reflexões.

Figura 6 – Primeira página da edição inaugural do *Correio Braziliense* ou *Armazém Literário*, em junho de 1808, fundado por Hipólito José da Costa



Fonte: Wikipedia (s.d.).

O nome *Correio Braziliense* foi resgatado por Chateaubriand como uma forma de homenagear Hipólito José da Costa e também para entrelaçar a história do Brasil com a modernidade representada pela nova capital, Brasília (DF), e o momento de “descobrimento” e colonização do país (Cesar, 2006). O “novo” *Correio Braziliense* tornou-se o jornal oficial da nova capital, sendo uma das primeiras publicações a registrar os acontecimentos históricos relacionados à fundação de Brasília. Desde então, o *Correio Braziliense* se consolidou como um dos jornais mais influentes e respeitados do Brasil, acompanhando de perto os eventos políticos e sociais que moldaram o país a partir de sua nova capital.

Chateaubriand, por sua vez, tornou-se um dos mais importantes empresários de comunicação do Brasil, como fundador do império de mídia conhecido como Diários Associados. Construiu uma vasta rede de veículos de comunicação que incluía 31 jornais, espalhados por diferentes regiões do Brasil, e quatro revistas, entre as quais se destacava *O Cruzeiro*, uma das mais influentes da época. Havia, ainda, uma agência de notícias, que ajudava a interligar seus veículos de comunicação e fornecer conteúdo para outros meios, além de 16 estações de rádio, pioneiras em sua época e que foram precursoras da difusão massiva de notícias e entretenimento via rádio no Brasil. O conglomerado midiático de Chateaubriand

incluía, ainda, cinco emissoras de televisão, sendo uma delas a TV Tupi, a primeira estação de televisão da América Latina, inaugurada em 1950. A cadeia noticiosa do empresário o tornou conhecido e garantiu a ele grande influência na política e na cultura do Brasil, elegendo-se senador, além de membro da Academia Brasileira de Letras (ABL) e tornando-se embaixador em Londres (Araújo; Souza, 2003).

2 EVOLUÇÃO DO JORNALISMO DIGITAL

Prática recente e que continua sendo remodelada, o jornalismo realizado com suporte da internet recebe diferentes nomenclaturas, como, por exemplo, “jornalismo digital”, “jornalismo *online*”, “jornalismo hipertextual”, “jornalismo eletrônico”, “webjornalismo” e “ciberjornalismo” (Mielniczuk, 2001; 2003; 2004). Para Canavilhas (2003; 2007), o webjornalismo é a transposição do jornalismo para o meio *online* e reúne em seus conteúdos elementos que o caracterizam, como texto e interatividade, texto e hipertexto, leitura não-linear com integração de elementos multimídia, utilização de som, de vídeo, elementos em *flash*, 3D e gráficos. Outros recursos que podem ser utilizados pelo webjornal são a personalização, distribuição e atualização constantes.

Os termos “jornalismo digital” ou “ciberjornalismo” são empregados por Nunes (2016) como sinônimos, abarcando com maior propriedade e amplitude – inclusive geográfica – o contexto atual, no qual conteúdos noticiosos expandem sua difusão por meio dos dispositivos móveis. Conforme explica:

Considera-se aqui que o termo jornalismo digital traz uma abordagem mais ampla, com utilização frequente para além da América Latina e englobando quaisquer apropriações jornalísticas que se situem no universo digital, conectado ou não. É importante pensar que em certos aplicativos mobile, não é mais preciso estar conectado para fazer a leitura dos conteúdos, sendo necessária conexão à internet somente para a sua atualização. Assim, deixa-se de certa forma o âmbito do ciberespaço: sua existência é dependente da internet (para atualização), mas sua leitura/visualização pode estar alheia à conectividade, mesmo que haja ainda perda de alguns conteúdos que necessitem de conexão, como hiperlinks e vídeos (Nunes, 2016, p. 25).

O jornalismo digital compreende uma época na qual publicações como as disponíveis para *tablets* encontram suporte na internet, mas não estão diretamente ligadas à *World Wide Web*. Difere-se, dessa forma, da nomenclatura “webjornalismo”, que trazia a ideia de produtos dependentes desse ambiente digital fechado em si mesmo, em que o “www” regia as linguagens, o espaço e inclusive certas características de formatação e *design* – um período em que a internet comercial era majoritariamente a “web”. Nas palavras de Nunes:

A web impulsionou a expansão da internet comercial como algo delimitado (com utilização padrão e linguagem “www”) ao mesmo tempo em que abrangente (possibilitando a inclusão de páginas de diversos assuntos em um universo sem fim). O termo webjornalismo se vincula a um período em que a informação jornalística estava disponibilizada digitalmente de forma majoritária em páginas de sites na web. Os aplicativos móveis e a chamada “internet das coisas”, que estão sendo popularizados pelos QR Codes e

etiquetas de RFID, ampliam esta noção de internet. Hoje a web é apenas uma fatia da internet, que figura como um universo amplo de possibilidades, com aplicativos móveis situados fora do espaço da web, e a própria internet das coisas, em um cenário em que a cada dia podem surgir novos usos e apropriações para a rede mundial que conecta computadores e dispositivos eletrônicos. A modificação do papel hegemônico da web na internet se mostra em dados, uma vez que já no ano de 2011, a utilização dos aplicativos móveis superou o uso da web tanto em dispositivos móveis como em computadores nos Estados Unidos (MDG Advertising Agency, 2012, compilados pela Flurry Analytics) (Nunes, 2016, p. 25).

A ampliação da internet de banda larga potencializou a capacidade de transmissão de dados, permitindo que conteúdos mais complexos e variados fossem disponibilizados *online*. Isso inclui vídeos de alta definição, transmissões ao vivo, gráficos interativos, infográficos detalhados e outros recursos multimídia que enriquecem as reportagens. A melhoria na infraestrutura de internet também facilitou o acesso contínuo e rápido a informações atualizadas, característica basilar do jornalismo digital.

Além disso, contribuiu para a popularização das redes sociais e plataformas de distribuição de conteúdo, o que modificou profundamente a forma como as notícias são consumidas e disseminadas. Os leitores não apenas acessam informações de maneira mais ágil e fácil, mas também participam da criação e disseminação de conteúdo, interagindo com jornalistas e veículos de comunicação.

Na era pós-industrial, o jornalismo digital incorpora uma adaptação contínua às novas tecnologias e às mudanças nas expectativas e comportamentos do público, impulsionadas pela disponibilidade e pelo avanço das conexões de alta velocidade.

No Brasil, a velocidade média da internet fixa de banda larga alcançou 188,34 Mbps entre janeiro e março de 2024. Um estudo atualizado no dia 19 de abril de 2024 com dados de testes de conexão residencial realizados no período por meio do *Ranking Minha Conexão* analisou 2,6 milhões de medições de velocidade no período. A conexão é 5,89% maior em comparação ao último trimestre de 2024, em que a média de velocidade do país foi de 177,86 Mbps. Mesmo com o aumento de desempenho, a média de conexão do país para banda larga ainda está 4,35% abaixo do registrado no mesmo período do ano anterior. No primeiro trimestre de 2023, a velocidade média no território brasileiro era de 196,90 Mbps. Também está muito aquém dos países com as conexões de internet fixa de banda larga mais rápidas do mundo, como Mônaco (261,82 Mbps), Singapura (255,83 Mbps), Hong Kong (254,70 Mbps), Romênia (232,17 Mbps) e Suíça (229,96 Mbps), segundo dados da Ookla.

O Centro-Oeste se posiciona como a região brasileira com maior velocidade média de internet fixa de banda larga. Três dos quatro entes federados da região central do país registram

internet banda larga fixa com velocidade superior aos demais estados brasileiros, segundo o *Ranking Minha Conexão* de 2024. No topo do *ranking* nacional está o Piauí, com 193,19 Mbps, seguido de Mato Grosso (163,85 Mbps), Goiás (156,6 Mbps) e Distrito Federal (127,07 Mbps). Mato Grosso do Sul está em décimo primeiro lugar no país, com velocidade média de internet de 109,18 Mbps.

Em relação à conexão de internet móvel no Brasil, após um ano da implementação da tecnologia de quinta geração (5G) no país, a velocidade média *mobile* era de apenas 30,40 Mbps. Em junho de 2023, a velocidade média nacional foi de 29,56 Mbps, equivalente à média verificada em abril de 2022, de 29,65 Mbps. A conexão é considerada boa para navegar em redes sociais e assistir a vídeos em alta resolução.

2.1 Jornalismo digital de primeira a quinta geração e o Jornalismo de Dados

Barbero (2008) observa que desde o final dos anos 1980 o cenário da comunicação na América Latina é protagonizado pelas novas tecnologias, representando um processo contínuo de aceleração da modernidade, desde a Revolução Industrial até a Revolução Eletrônica, do qual nenhum país pode estar ausente sob pena de morte econômica e cultural:

O surgimento de tais tecnologias na América Latina se inscreve, em todo caso, num velho processo de esquizofrenia entre modernização e possibilidades reais de apropriação social e cultural daquilo que nos moderniza. Informatização ou morte! - é o lema de um capital em crise, precisando com urgência vital expandir o consumo de informática (Barbero, 2008, p. 256).

Com a aparição das plataformas digitais na década de 90 do século XX, o jornalismo deixou de ser um produto exclusivo dos meios de comunicação tradicionais, incorporando outros canais de divulgação de conteúdos noticiosos. Segundo Amado:

O jornalismo, como atividade responsável pela produção de informações sociais a partir de uma variedade de práticas e modelos, está mudando na medida em que não é mais um produto exclusivo dos meios para se desenvolver em plataformas, portais de informação, redes sociais e outras mídias digitais. Sem a relação estreita entre os meios e as redações, o jornalismo começa a enfrentar outros atores e novas práticas que também fornecem informações. Nos modelos jornalísticos implícitos nas discussões públicas, permanecem em vigor as funções clássicas das formas de supervisão ou monitoramento do entorno, a correlação dos diferentes setores da sociedade em resposta ao contexto e a transmissão de valores sociais e entretenimento (Amado, 2019, p. 12).

Contudo, devido às limitações de conexão com a internet, o texto verbal escrito continuou sendo preponderante no jornalismo digital, uma vez que o *download* de versões simplificadas de páginas, com menos imagens e vídeos, acaba sendo relativamente mais rápido em condições de baixa velocidade.

Ainda no final da década de 80, com a disseminação de computadores e *softwares* de edição nas redações, os jornais conseguiram antecipar o fechamento das edições e passaram a realizar divulgação *online*, no formato digitalizado, das mesmas notícias da versão impressa. As notícias e reportagens publicadas na internet, simples reproduções dos jornais impressos e sem inovações de apresentação nas narrativas, eram atualizadas a cada 24 horas, após o fechamento das edições diárias pelos veículos tradicionais. Essa fase de transposição marca a primeira geração do jornalismo digital, caracterizada pela adaptação do impresso ao ambiente online.

A fase posterior do jornalismo digital, definida como segunda geração, é caracterizada como uma metáfora do jornalismo, por ser análoga e atrelada ao modelo dos jornais impressos. Nessa fase, que surge no final dos anos 90, os conteúdos noticiosos começam a explorar as características oferecidas pela internet, utilizando recursos de hipertexto, *links* com chamadas para notícias de fatos ocorridos entre as edições impressas, *e-mail* para comunicação entre jornalista e leitor ou entre leitores através de fóruns de debates.

No jornalismo, as comparações figurativas por meio de metáforas costumam ser empregadas para fazer referências às funções e aos propósitos dessa prática profissional. É recorrente a referência à ideia de que o jornalismo atua como vigilante ou “cão de guarda” (*watchdog*), implicando que os jornalistas estão constantemente observando e reportando eventos para manter a sociedade informada e atenta sobre temas de interesse público. Essa metáfora ressalta a responsabilidade dos jornalistas em serem “os olhos e ouvidos” do público, buscando a verdade, investigando questões relevantes e comunicando-as de forma clara e imparcial. Assim como vigias que protegem uma comunidade, os jornalistas desempenham a função de fiscalizar o poder, além de combater apagamentos sociais dando visibilidade e reconhecimento a indivíduos ou grupos excluídos ou marginalizados, contribuindo para uma sociedade democrática. Outro exemplo comumente utilizado para descrever seu papel é a metáfora do Quarto Poder, sugerindo que o jornalismo tem uma atribuição significativa na sociedade, assim como os três poderes tradicionais (Executivo, Legislativo e Judiciário). Os jornalistas têm a capacidade de influenciar a opinião pública, responsabilizar os líderes e instituições e moldar o debate público. Também atuariam como uma “mão invisível”, em referência à metáfora inspirada na ideia de Adam Smith sobre o livre mercado, que sugere que

o jornalismo orienta e regula a sociedade ao expor informações, revelar verdades e promover a transparência.

Essas expressões figurativas ilustram diferentes aspectos da função do jornalismo na sociedade, destacando seus papéis como mediador, fiscalizador, influenciador e educador. Sobre as metáforas, Amado aponta:

As metáforas são formas abstratas de pensamento que se relacionam com o pensamento simbólico e criativo (DAMASIO, 2018) e ajudam a entender uma coisa em termos de outra. Elas organizam a realidade, facilitam a percepção dos fatos e atribuem funções em uma ordem compartilhada. A palavra imprensa (prensa, press) é uma metáfora permanente da atividade jornalística. Designa o dispositivo técnico com o qual foram impressos panfletos e jornais, que antecederam as publicações periódicas que no século XIX consolidaram a democratização da esfera pública. Inicialmente, imprensa era a atividade das gráficas responsáveis pela divulgação de informações, notificações oficiais, boatos e outros tipos de conteúdo (WAISBORD, 2013c). Por metonímia, ela passou a designar a indústria e os jornalistas, cuja atividade é protegida pelo direito à liberdade de imprensa, independentemente dessa atividade ser realizada por meio de outro dispositivo técnico. E também sem considerar que a mídia impressa enfrenta a principal crise do século XXI, com tendências indicando que essa forma de distribuição é a que mais sofra retrocessos (Amado, 2019, p. 13).

Amado (2019) cita doze metáforas para o jornalismo, cinco delas usadas nos estudos da área pelos acadêmicos: jornalismo como profissão, vocação, instituição, prática ou texto. Outras sete são acrescentadas para se tratar do jornalismo como prática profissional: o jornalismo como espelho, instinto, conteúdo, história, criatura, serviço e compromisso. Para cada uma dessas metáforas há títulos de meios de comunicação que recorrem à analogia para explicar sua missão, alguns deles cristalizados desde o século XIX.

Às vezes, a mídia quer se identificar como mais um cidadão do seu povo, fazendo referência à nacionalidade. O meio ou a tecnologia como suporte para o conteúdo noticioso que reproduz também está em títulos, a exemplo do *Gazeta Digital* e do *Correio Braziliense*, que compõem o objeto desta pesquisa. O jornalista seria o executor da ferramenta, aquele que produz o que é apresentado ao público.

O compromisso do jornalismo costuma estar associado ao partidário, mas no ambiente digital se consagrou a ideia de engajamento para designar o envolvimento do público com a proposta informativa, medida pela soma dos cliques. Entende-se que é mais relevante comentar ou compartilhar uma notícia nas próprias redes com amigos e seguidores do que apenas visualizá-la. O lugar central que os leitores adquirem na circulação de informações nas redes sociais faz com que a participação seja um fator importante na mídia, assim como a

personalização e as narrativas colaborativas que acrescentam ao jornalismo o adjetivo cívico, cidadão, público, comunitário. Tudo isso está expresso nos títulos de jornais que incluem palavras como cidadão, liberal, popular etc. (Amado, 2019).

Muitas das diferentes metáforas empregadas para expressar os papéis do jornalismo mostram, contudo, cristalizações linguísticas que aparecem em diferentes idiomas sem captar as mudanças pelas quais o jornalismo está passando. Essa transformação tem afastado o jornalismo de figuras de linguagem nostálgicas que o descrevem de maneira idealizada. Agora, surgem novas metáforas que sugerem que o jornalismo pode enfrentar os desafios do mundo contemporâneo. Uma delas indica que o jornalismo pode transcender os problemas causados pela comercialização em massa, através do apoio de comunidades interessadas, seja por meio de assinaturas ou financiamento coletivo, também chamado de *crowdfunding*. Essa perspectiva coletiva também amplia o papel do jornalismo como uma ferramenta de vigilância e colaboração, abrangendo público diverso e propondo soluções para problemas sociais na forma de um jornalismo de soluções.

Observamos, ainda, que em um cenário no qual a imprensa já não detém o monopólio da visibilidade dos eventos sociais, o fato de ser o primeiro a relatar uma notícia – furo jornalístico, no jargão técnico – perde relevância, uma vez que qualquer pessoa pode compartilhar uma informação inédita quase instantaneamente, contribuindo para um jornalismo participativo.

Continua sendo indispensável, contudo, a verificação profissional dos fatos, garantindo a precisão da informação em meio à pluralidade de conteúdos gerados por fontes diversas e disseminados em diferentes canais digitais.

No esforço para organizar os significados em mutação para o jornalismo, Amado (2019) retoma classificações metafóricas clássicas e considera, para cada uma delas, novas metáforas, conforme apresentamos a seguir:

Quadro 1 – Metáforas jornalísticas clássicas e metáforas futuras

Metáforas clássicas do século XX	Metáforas básicas para o século XXI
Massivo	Coletivo
Primazia	Verificação
Produção	Curadoria
Observatório de ética	Diálogo Ético
Problema	Solução
Informação interessada	Informação interessante
Serviço público	Serviço ao público

Metáforas clássicas do século XX	Metáforas básicas para o século XXI
Receptor Visibilidade Liberdade de Imprensa	Gamer Credibilidade Liberdade de rede

Fonte: Amado (2019, p. 21).

As recentes mudanças na prática jornalística na era digital não são, contudo, meramente tecnológicas, porque o jornalismo não se reduz aos canais ou ferramentas que agora compartilham com a sociedade em geral, conforme aponta Amado (2019). O jornalismo se desenvolve atualmente nas mesmas plataformas em que circulam as informações institucionais e interpessoais. Às vezes, encontra-se até em desvantagem perante comunidades de interesses que conseguem conexão maior do que com a mídia, que ainda está ligada à massa e que entende os destinatários como receptores. Nesse contexto, o jornalismo continua sendo pensado em outros suportes, em novos cenários, com outras ferramentas:

Não se pode esquecer que todo processo de mediação vai além da tecnologia, pois envolve parâmetros de sociabilidade: os dispositivos técnicos dos fenômenos midiáticos se estabilizam de maneira institucional e geram seus próprios grupos (VERÓN, 2013). A mediação não é reduzida apenas à tecnologia, mas envolve também agentes, instituições, canais e práticas de comunicação pública como um todo. O jornalismo é um agente, uma instituição, um canal e um ofício e está experimentando transmutações em todos esses aspectos. Sua reinvenção deve passar por todas as esferas do sistema e se expressar em novas metáforas para designar o jornalismo e os espaços em que se desenvolve (Amado, 2019, p. 19).

2.2 Transposição e simbiose geracional

Barbosa (2005; 2007; 2008; 2013) reforça que o jornalismo continua a incorporar novidades tecnológicas. Na evolução digital, novos aparatos são introduzidos na prática jornalística ao ponto de se tornarem indissociáveis e intrínsecos, como por exemplo o computador. Com a internet e o surgimento de dispositivos conectados à rede – como os *smartphones* e *tablets* – a coleta, produção, publicação e disseminação de conteúdos ganharam nova dimensão, processo denominado jornalismo digital, jornalismo *online*, jornalismo hipertextual, jornalismo eletrônico, webjornalismo ou ciberjornalismo (Mielniczuk, 2001; 2003; 2004; Nunes, 2016; Canavilhas, 2003; 2007), entre outras nomenclaturas.

As mudanças evoluíram a tal ponto que, no início dos anos 2000, surgiram diversos *websites* jornalísticos, versões de congêneres impressos, *sites* de emissoras de TV e de rádios,

até culminar no surgimento de produtos eminentemente digitais, chamados de “nativos digitais”, criados por empresas de comunicação, informática e telecomunicações.

Gradualmente, os diversos tipos de web sites jornalísticos (edições digitais de jornais, revistas, emissoras de tevê, agências de notícias, portais, portais regionais, e, mais recentemente os blogs) vão experimentando narrativas diferenciadas, nas quais o uso de recursos hipertextuais - como o link -, associados ao áudio, ao vídeo, à fotografia, aos gráficos animados para a redação das informações auxilia na construção e na consolidação de uma linguagem para essa modalidade, que passa a ter nos usuários (ou no público) importantes colaboradores para a formulação dos conteúdos dentro de uma estrutura já descentralizada. Interatividade, hipertextualidade, multimidialidade, personalização, atualização contínua e em fluxo e a memória (arquivamento) constituem-se os elementos característicos do jornalismo digital, conforme apontam diversos investigadores (Palacios, 1999; 2002; 2004; Bardoel & Deuze, 2000; Zamorra, 2002; Salaverría, 2005) (Barbosa, 2007, p. 128).

Com a utilização de bases de dados (BDs) para a estruturação e organização das informações nasce a terceira geração ou terceira onda de evolução no jornalismo digital. Nesse estágio, as características predominantes são o uso de base tecnológica ampliada, favorecida pelas conexões de banda larga; a proliferação de plataformas móveis; redação descentralizada e adoção de sistemas que permitam a participação do usuário; produtos criados originalmente para veiculação no ciberespaço; conteúdos dinâmicos formatados em narrativas multimídia como sons e animações; e experimentação de novos elementos conceituais para organização da informação. Além disso, os produtos jornalísticos oferecem recursos de interatividade, como *chats*, enquetes e fóruns de discussões, sendo que os recursos de hipertexto começam a ser empregados na narrativa dos fatos (Mielniczuk, 2001; 2003; 2004). Segundo Barbosa:

A linha evolutiva sistematizada em gerações – a primeira caracterizada pela reprodução ou *shovelware*¹⁰ dos conteúdos e, a segunda, a da metáfora, com produtos ‘atrelados’ ao modelo do jornal impresso, na qual já se começa a explorar os recursos oferecidos pela rede, como o e-mail e o link – não é excludente, ainda que demarque períodos distintos desde 1995 até o momento atual. Tampouco pode-se precisar o tempo de duração para cada geração, pois, ainda hoje há vários web sites jornalísticos que operam segundo o modelo da transposição e da metáfora. De acordo com Pryor (2002), a terceira onda ou, para nós, terceira geração, inicia-se em 2001, e o seu desenvolvimento está ainda em processo (Barbosa, 2007, p. 130).

As bases de dados são intrínsecas ao jornalismo digital e serão cada vez mais utilizadas. A evolução para um novo *status* que envolve a incorporação de banco de dados ao jornalismo

¹⁰ *Shovelware* é um termo da computação que se tornou pejorativo por referir-se a pacotes de *software* mais notáveis pela sua quantidade do que por sua qualidade ou utilidade (Alves, R., 2006).

digital culmina em uma definição operacional para as funcionalidades existentes, conceituada por Barbosa (2005; 2007; 2008; 2013) como Paradigma do Jornalismo Digital em Base de Dados (JDBD), identificado na transição da terceira para a quarta geração, e que reflete uma complexificação dos processos para implementação de produtos jornalísticos no ciberespaço. Ocorre uma evolução significativa na forma como as notícias são produzidas, distribuídas e consumidas. Esta fase é marcada por uma base tecnológica ampliada, que oferece novas possibilidades e desafios para jornalistas e organizações de mídia.

Entre os principais aspectos dessa transformação estão o acesso expandido à informação e conexões de banda larga, permitindo que os usuários obtenham conteúdos abundantes em dados e com alta qualidade de áudio e vídeo sem interrupções. Isso ampliou o alcance das publicações jornalísticas e facilitou a disseminação de notícias em tempo real. Outro fator que caracteriza essa mudança é a proliferação de plataformas móveis, como *smartphones* e *tablets*, revolucionando a forma como as notícias são consumidas e exigindo que os conteúdos sejam otimizados para diferentes dispositivos, levando à criação de aplicativos dedicados e *sites* responsivos que melhoram a experiência do usuário/leitor.

Diante da complexidade dos novos processos tecnológicos são demandadas equipes especializadas e multidisciplinares, incluindo jornalistas, desenvolvedores, *designers* gráficos, analistas de dados e especialistas em SEO (*Search Engine Optimization*). Esse conjunto diversificado de habilidades moldam uma nova estrutura laboral para a produção de conteúdo multimídia de alta qualidade e gestão eficiente de plataformas digitais.

Com o uso expandido de bases de dados e algoritmos, o jornalismo baseado em dados ou *data-driven journalism* ganha destaque, utilizando grandes conjuntos de dados para criar reportagens mais aprofundadas e analíticas. Algoritmos são empregados para personalizar a experiência do usuário, sugerindo conteúdos alinhados com suas preferências e histórico de navegação.

São desenvolvidos Sistemas de Gestão de Conteúdo (*Content Management Systems - CMS*) mais complexos, que evoluíram para suportar a integração de diferentes formatos de mídia, facilitar a publicação em múltiplas plataformas e permitir a colaboração em tempo real entre membros de equipes, os quais atuam de forma remota e estão até distribuídos geograficamente. Esses sistemas automatizam tarefas administrativas, aumentando a eficiência operacional.

Ocorre, ainda, maior incorporação dos *blogs* e participação do usuário na produção de informações que se tornaram elementos determinantes para o jornalismo digital. Plataformas que permitem comentários, *blogs* de colaboradores e seções de opinião proporcionam uma

maior interatividade e engajamento do público.

Outro aspecto que se sobressai é a incorporação de gráficos, vídeos, áudios e demais elementos multimídia nas narrativas jornalísticas, enriquecendo a experiência informativa. A infografia interativa permite que os leitores explorem dados de maneira visualmente atraente e compreensível. Nota-se, ainda, o emprego do RSS (*Really Simple Syndication*) para facilitar a distribuição e o compartilhamento de conteúdo em diferentes formatos – em texto, áudio e vídeo –, permitindo que leitores recebam atualizações de seus *sites* e *blogs* favoritos em tempo real.

Há uma integração de arquivos e metadados, que aprimoram a organização e a acessibilidade da informação. *Data mining*¹¹ é utilizada para extrair conhecimento de grandes conjuntos de dados, auxiliando na identificação de tendências e padrões que podem gerar novas histórias jornalísticas. São incorporados novos métodos de organização e visualização de informação, como mapas interativos e *timelines* dinâmicas, oferecendo ao leitor formas inovadoras de compreender e interagir com os conteúdos. Essas técnicas melhoram a contextualização e a compreensão dos fatos noticiados.

A quarta geração do jornalismo digital representa um marco no uso da tecnologia para criar uma experiência informativa mais variada, interativa e personalizada. Tais mudanças, conformadas a partir da concepção de bases de dados como um formato, favorecem o surgimento de uma nova metáfora (Barbosa, 2005; 2007; 2008; 2013) que permite superar aquela do impresso, também conhecida como *broadsheet metaphor* (Eriksen; Ihlström, 2000)¹².

A replicação de modelos que não servem mais, impedindo experimentações e inovações, são, contudo, um dos entraves ao desenvolvimento do jornalismo digital, segundo Barbosa (2005; 2007; 2008; 2013). A autora pressupõe, portanto, uma ruptura com um modelo padrão no jornalismo e que poderá conduzir para outro paradigma e uma nova metáfora para o jornalismo digital, configurada numa etapa posterior de desenvolvimento.

Na quarta geração, as bases de dados são estruturantes da atividade e agentes singulares no processo de convergência jornalística. Estão integradas nos sistemas publicadores, assegurando mais agilidade e qualidade à construção das narrativas, por um lado, e, por outro, consolidando um processo pós-industrial de produção no jornalismo digital (Schwingel, 2005;

¹¹ *Datamining* ou mineração de dados consiste em um processo para se obter e explorar grandes quantidades de dados.

¹² A metáfora do *broadsheet* (jornal de grande circulação) encapsula os ideais de jornalismo sério, profundo e confiável. Em um mundo onde o conteúdo digital é muitas vezes efêmero e superficial, a invocação dessa metáfora serve como um lembrete da importância de se manter a qualidade e a integridade jornalística. Assim como as grandes páginas de um *broadsheet* acomodam análises detalhadas e múltiplas perspectivas, o jornalismo de alta qualidade no ciberespaço deve buscar oferecer uma cobertura ampla, profunda e inclusiva das notícias e dos eventos globais.

Santos, 2017). Nesse sentido, Barbosa explica:

Ao propormos o modelo JDBD como paradigma de uma fase de transição, ressaltamos que o quarto estágio se encontra ainda em gestação. Porém, seja qual for a configuração que irá consolidá-lo, as bases de dados serão um elemento estruturante, pela flexibilidade intrínseca a elas, assim como o potencial de reinvenção que elas possuem (Barbosa, 2007, p. 152).

Bases de dados implicam em reconstrução do jornalismo digital, provocando modificações no sistema de obtenção e produção de informações, assim como em sua difusão. Em diálogo com a perspectiva de emergência de uma nova fase do jornalismo, Barbosa (2005; 2007; 2008; 2013) afirma que o JDBD é balizador para inferir a existência de uma quinta geração de desenvolvimento para o webjornalismo ou ciberjornalismo. A autora resalta características no que se refere aos meios, utilizando o conceito de medialidade, horizontalidade e “integração de processos e produtos no continuum multimídia dinâmico”.

Barbosa (2005; 2007; 2008; 2013), em sua pesquisa, debruça-se também sobre os produtos e destaca as aplicações autóctones como aquelas com maior potencial inovador, as quais se relacionam com demais gêneros. Assim, identifica a emergência de um novo estágio do jornalismo digital, cujas características estão ligadas às iniciativas móveis. Os produtos dessa geração têm particularidades e apresentam possibilidades jornalísticas não pertencentes às quatro fases anteriores.

A partir dessa análise, consideramos o advento de uma quinta geração de jornalismo digital, cujas características estão ligadas a certos delineamentos como produtos e aplicativos dos jornais, mobilidade e independência do espaço *web*, com amplificação do potencial de leitura *offline*, mas dependente da internet para acessos a *links*, vídeos e dados ao vivo (Nunes, 2016). Há dependência, com menor intensidade, da construção visual, de organização de conteúdo, além de hierarquização noticiosa e *design* gráfico ainda inspirado nos veículos jornalísticos de papel. Essa fase dialoga com influências digitais da própria *web*, tornando-se um híbrido.

No universo de iniciativas jornalísticas *mobile*, vislumbram-se também segmentações. Há edições digitais que são réplicas do papel ou de *sites* na *web*, ou seja, apresentam características similares à da impressão em papel, mas dialogam com a infinidade de conteúdo que a internet disponibiliza, denominadas “webmóveis” (Nunes, 2016). Outro atributo é a finitude, com a possibilidade de se esgotar a leitura do conteúdo compilado ou seguir uma linearidade de leitura com início, meio e fim das edições digitais móveis. Como exemplo de uma primeira segmentação do jornalismo digital *mobile*, Nunes (2016) cita o *Sunday Examiner*, de São Francisco, nos Estados Unidos, por manter edição digitalizada como uma réplica da

edição impressa, sem interação ou conteúdo extra, veiculada digitalmente em formato “PDF”¹³. Nesse modelo produtivo das redações de jornalismo, a página diagramada digitalmente segue dois caminhos, sendo um para a gráfica, para ser impressa, e o outro para o universo digital, ficando disponível nos dispositivos móveis e por vezes também na *web*. Esse processo de mobilidade digital continua sendo, contudo, encabeçado pelas edições impressas dos jornais.

Tal hibridização também é identificada e perdura na imprensa regional, tendo como exemplos os conglomerados de mídia da região Centro-Oeste brasileira, como é o caso do Grupo Gazeta de Comunicação, em Mato Grosso, mantenedor dos noticiários *Gazeta Digital* e jornal *A Gazeta*, e do *Correio Braziliense*, do Grupo Diários Associados S/A, no Distrito Federal. Os dois grupos de comunicação mantêm versões impressas digitalizadas ou na forma de webjornais, que também podem ser acessados a partir de dispositivos móveis, conforme demonstramos a seguir:

Figura 7 – Edição digitalizada da publicação impressa do diário *Correio Braziliense* em 21 de junho de 2023, disponível no site e pelo aplicativo *CB Digital*



Fonte: Correio Braziliense (2023).

¹³ Sigla em inglês para *Portable Document Format* ou formato portátil de documento, uma extensão de arquivos destinados a leitura.

Figura 8 – Edição digitalizada da publicação impressa do diário *A Gazeta* em 21 de junho de 2023, disponível no site e pelo aplicativo *Gazeta Digital*



Fonte: Gazeta Digital (2023).

Os aplicativos móveis (apps) exclusivos dos veículos jornalísticos apresentam, quase sempre, características semelhantes aos *sites da web*, constituindo apenas uma adaptação ou mesmo transposição de edições impressas para o formato *mobile*, denominadas webmóveis.

Em diálogo com o cenário nacional, Nunes (2016) traz uma amostragem complementar dos dez aplicativos selecionados a partir dos 15 maiores jornais do Brasil¹⁴, percebendo que os aplicativos das publicações *O Globo*, *Super Notícia* (por meio do app do *O Tempo*), *Extra*, *Aqui* (MA), *O Tempo* e *Dez Minutos* disponibilizavam uma versão em “PDF” sem interatividade, sendo, portanto, similares ao modelo adotado pelos dois conglomerados de mídia da região Centro-Oeste citados nesta pesquisa, mantidos pelo Grupo Gazeta de Comunicação, em Mato Grosso, e ao *Correio Braziliense*, do Grupo Diários Associados S/A, no Distrito Federal, como descreveremos adiante.

O jornalismo de quinta geração que emerge a partir dos dispositivos móveis e de

¹⁴ Baseado em levantamento da Associação Nacional de Jornais (ANJ) (apud Nunes, 2016).

mudanças culturais de leitura delineia-se com fortes bases no jornal impresso, com hibridismo entre as características da *web* e dos meios móveis. Evidencia-se, desse modo, uma tendência de posicionamento dos produtos digitais como forma complementar de distribuição de um veículo material (*offline*), situação percebida na análise dos aplicativos brasileiros, inclusive da região Centro-Oeste, segundo apontamos neste estudo.

Frente às diferentes apropriações presentes no jornalismo digital de quinta geração, podemos considerar as publicações como produtos do contexto social digital móvel, por se embasarem na mobilidade, e com características específicas que perfazem não só o contexto, mas tornam possível a identificação e a divisão desses produtos digitais. As segmentações apresentadas evidenciam o desenvolvimento de iniciativas móveis jornalísticas atuais e representam experiências das indústrias midiáticas em produzir conteúdo para as plataformas móveis, as quais crescem em uso e podem representar um caminho em relação à “crise do papel” (Nunes, 2016).

Destacamos, ainda, que as mudanças nos produtos jornalísticos digitais estão relacionadas ao contexto cultural de leitura e de transformação cognitiva de apreensão de conhecimento. Conforme registra Santaella (2019), com o surgimento da internet emergiu o leitor imersivo, que navega e interage nas infovias dos ambientes da *web*. Assim, o advento das mídias móveis trouxe consigo o leitor ubíquo, aquele que tem acesso à informação em qualquer lugar onde possa estar.

As fases do jornalismo digital evoluem, portanto, em consonância com uma mudança nas formas de leitura que trazem referenciais cognitivos de outras gerações do jornalismo para apoiar e suportar culturalmente certas rupturas. Mais do que mudanças tecnológicas, as referências ao papel no jornalismo de quinta geração interligam-se com mudanças culturais de leitura e também de leitores.

Nessa nova fase, o jornalismo adapta-se às demandas de uma sociedade cada vez mais conectada e exigente, buscando oferecer conteúdo relevante, transparente e interativo, de forma segmentada e personalizada. Uma das características distintivas do jornalismo de quinta geração é a sua natureza multimídia e multiplataforma. Para isso, utilizam algoritmos e análise de dados.

Nesse sentido, os veículos de comunicação exploram uma variedade de formatos, incluindo texto, áudio, vídeo, infográficos e realidade virtual, para contar histórias de maneira mais envolvente e acessível a diferentes tipos de audiência. Transparência e colaboração também estão presentes nessa fase, quando os veículos buscam estabelecer uma relação de confiança com sua audiência, fornecendo informações sobre suas práticas editoriais, fontes de

financiamento e processos de produção de notícias, bem como incentivando a participação ativa do público, o engajamento, o *feedback* e até mesmo a colaboração direta na produção de conteúdo. Para demonstrar transparência, precisão e credibilidade jornalística, investem na verificação de fatos e checagem de dados, especialmente diante da proliferação de desinformação e informações falsas (Da Silva, Medeiros e Ceretta, 2020).

Quadro 2 – Características geracionais do jornalismo digital

1a Geração - Transposição	Transposição do jornal impresso para internet no início dos anos 80.
2a Geração - Metáfora	Fase da “metáfora” do jornalismo impresso. Uso de alguns recursos multimídia, como links e emails para comunicação entre jornalistas e leitores, fóruns de debates. Surge na década de 90.
3a Geração - Jornalismo Digital	Inicia a partir de 2001. Incorpora características da 2a fase do jornalismo digital com apresentação de novos recursos em multimídia, como sons e animações, interatividade com chats, enquetes, fóruns de discussões e hipertexto.
4a Geração – JDBD	Surgimento do JDBD na transição da 3a para 4a geração. Base tecnológica ampliada por conexões banda larga; Plataformização (redes sociais); Equipes especializadas; Uso de base de dados, algoritmos, linguagens de programação; Metadados; Mineração de dados; Sistemas de gestão de conteúdos; Blogs; Interação do usuário/leitor na produção de conteúdo; Automatização de produtos; Sites dinâmicos; Narrativas multimídias; Infografia interativa; Emprego do RSS (Really Simple Syndication ou Rich Site Summary) para recolher, difundir e compartilhar conteúdos em áudio e vídeo;
5a Geração - Mobilidade	A 5a geração de jornalismo digital tem características ligadas à mobilidade. Variedade de formatos de mídia (texto, áudio, vídeo, infográficos e realidade virtual); Interatividade, convergência e multimídia.

Fonte: elaborado pela autora (2024) com base em Barbosa (2005; 2007; 2008; 2013); Mielniczuk (2001; 2003; 2004); Canavilhas (2003; 2007); Nunes (2016).

2.3 Jornalismo digital no Centro-Oeste: características gerais

Recortando para o contexto regional, observamos que os dois webjornais escolhidos como objetos empíricos para este estudo, *Gazeta Digital* e *Correio Braziliense*, apresentam elementos que caracterizam uma simbiose do jornalismo digital de primeira a quinta geração (Barbosa, 2005; 2007; 2008; 2013).

Nesse sentido, notamos que o webjornal *Gazeta Digital* apresentava, inicialmente, configuração semelhante à versão impressa do periódico, com divisões das mesmas editorias e suplementos. Mantinha, ainda, botões (*links*) de acesso à programação de outros veículos do Grupo Gazeta de Comunicação, como rádios e TV. Para estimular a interação com o público, disponibilizava um espaço para enquete virtual, opção mantida até hoje. Ao longo dos seus 21 anos, incorporou características que evidenciam a evolução geracional do jornalismo digital.

O *Gazeta Digital* não possui, contudo, uma versão exclusiva *mobile*, que seria uma configuração típica do jornalismo digital de quinta geração (Barbosa, 2005, 2007, 2008, 2013). Por meio do aplicativo *Gazeta Digital* (GD) é possível ler em dispositivos móveis, como *smartphones* e *tablets*, apenas a versão impressa digitalizada do jornal *A Gazeta*, característica que exemplifica a primeira etapa evolutiva da quinta geração do jornalismo digital. Mantém, porém, alternativa responsiva, ou *seja*, o site se adapta e funciona em qualquer dispositivo – seja *smartphone*, *tablet* ou *desktop* –, ajustando-se à tela e mudando sua apresentação de acordo com o mecanismo utilizado pelo leitor, promovendo uma navegação fluida e acessível. Além disso, utiliza recursos multimídia e multiplataforma. O *Gazeta Digital* emprega variedade de formatos de mídia, como texto, áudio e vídeo. Incentiva a participação dos leitores, promovendo engajamento e *feedback* na produção de conteúdo, como por meio do canal “Tô na Bronca”. A interação com o público é estimulada com a disponibilização de um número de telefone com aplicativo de mensagem instantânea e espaço para comentários mediados ao final da matéria, além de enquete virtual. Possibilita também o compartilhamento dos conteúdos por meio de redes sociais, como *Facebook* e *X* (antiga *Twitter*).

Seu predecessor, o jornal diário *A Gazeta*, além de ter suas edições digitalizadas hospedadas no site *Gazeta Digital*, com acesso livre, e que podem ser compartilhadas pelos leitores por meio de serviços de comunicação multimídia, como redes sociais e *e-mail*, também é disponibilizado através de aplicativo móvel (app), conforme citado anteriormente, possibilitando ao público leitor que acesse os conteúdos das edições diárias impressas a partir da instalação de programas específicos desenvolvidos para dispositivos eletrônicos móveis, disponíveis em lojas *online*.

O *Correio Braziliense*, por sua vez, apresenta características mais conformadas à quinta geração do jornalismo digital, por disponibilizar aos leitores o acesso tanto aos conteúdos informativos publicados no diário impresso como no digital por meio do seu aplicativo, CB. Ajustado à mobilidade e plataformização, permite o compartilhamento de notícias por meio de redes sociais e aplicativos de mensagens, com *WhatsApp*, *Twitter/X*, *Instagram*, *Facebook* e *YouTube*. Também permite incluir comentários relacionados às notícias e reportagens, que são publicados tanto na versão digital como edição impressa, que reserva espaço de, no máximo, dez linhas, para manifestação dos leitores. Também insere *hiperlink* em todas as notícias e reportagens publicadas.

Observamos, portanto, alguns deslocamentos que buscam transformar esses produtos em uma experiência multimidiática, interativa e convergente para os usuários, com produção e distribuição pensada diretamente para o suporte *mobile*, típicos da quinta geração do jornalismo digital, oferecendo conteúdos informativos exclusivos do webjornal e também da edição impressa digitalizada, como é o caso do *Correio Braziliense*.

3 DADOS NO JORNALISMO: ORIGENS E PERSPECTIVAS

Alves e Gehlen (2023) lembram que o uso de dados, números, anotações, sempre esteve presente em periódicos de comércios portuários para fazer registro das idas e vindas das embarcações desde o século XVIII. Essa prática de anotações deu-se inclusive durante a diáspora africana, quando aproximadamente 11 milhões de indivíduos vindos da África foram traficados nos transatlânticos denominados tumbeiros ou navios negreiros, sendo retirados à força de países como Angola, Senegal, Nigéria, Moçambique, entre outros, em direção à América, a partir do século XV até o século XIX.

O ato de coletar e registrar dados, inclusive a respeito dos próprios seres humanos, não é novo e está evidente em outros capítulos da história da humanidade, como durante o movimento eugênico no final do século XIX e início do século XX, que buscou utilizar dados para quantificar a superioridade dos brancos sobre os demais indivíduos, acrescentam D'Ignazio e Klein (2020), na tentativa de criar uma “ciência” capaz de legitimar o racismo.

Anteriormente, ainda no século XVIII, o filósofo suíço Jean-Jacques Rousseau destacou o uso do método estatístico em oposição à racionalidade comunicativa. Em sua obra *Do contrato social - Ensaio sobre a origem das línguas* ([1762] 1997), Rousseau define a vontade geral como uma medida numérico-matemática encontrada além da ação comunicativa. Não é a comunicação, mas a operação aritmética – um algoritmo, portanto –, que averigua a vontade geral:

Há comumente muita diferença entre a vontade de todos e a vontade geral. Esta se prende somente ao interesse comum; a outra, ao interesse privado, e não passa de uma soma das vontades particulares. Quando se retiram, porém, dessas mesmas vontades, os mais e os a-menos que nela se destroem mutuamente, resta, como soma das diferenças, a vontade geral (Rousseau, [1762] 1997, p. 91).

Para Byung-Chul Han (2022), Rousseau inaugura o dataísmo ao defender que, quanto mais dados estiverem disponíveis, mais autêntica é a vontade geral averiguada. Sua racionalidade aritmética se aproxima da racionalidade digital por abdicar do discurso e da comunicação, ou seja, da racionalidade comunicativa que consiste na autonomia e na liberdade dos indivíduos.

Lévy (1998) lembra que “a informação é hoje torrencial ou oceânica” e que “o hiato entre o caráter diluviano dos fluxos de mensagens e os modos tradicionais de decisão e orientação faz-se cada vez mais evidente” (Lévy, 1998, p. 61). O autor entende que os sistemas de governo ainda utilizam, contudo, majoritariamente técnicas de comunicação molares, ou

seja, baseadas em tecnologias midiáticas. A mídia constitui uma tecnologia molar que só age sobre as mensagens a partir de fora, por alto e em massa, contentando-se em fixar, reproduzir e transportar uma mensagem somaticamente produzida. Sendo assim, o digital sempre pairou acima da mídia. No vocabulário do digital não se fala de montagem, mas de computação, de cálculo ou de tratamento da informação. Sendo assim, a informática é uma técnica molecular, pois não se contenta em reproduzir e difundir mensagens – algo que faz melhor que a mídia clássica –; ela permite sobretudo engendrâ-las, modificá-las à vontade, conferir-lhes capacidade de reação de grande sutileza, graças a um controle total de sua microestrutura. O digital autoriza a fabricação de mensagens, sua modificação e mesmo a interação com elas, *bit* por *bit*.

Na comunicação escrita tradicional, todos os recursos de montagem são empregados no momento da redação. Uma vez impresso, o texto material conserva certa estabilidade, mesmo podendo haver desmontagens e remontagens de sentido pelo leitor. O hipertexto digital, porém, automatiza, materializa operações de leitura e amplia seu alcance. Bases de dados, sistemas especialistas, tabuladores, hiperdocumentos, simulações interativas e outros mundos virtuais são potenciais de textos, de imagens, de sons ou mesmo de qualidades táteis que situações específicas atualizam de mil maneiras. O digital reencontra, assim, a sensibilidade juntamente ao contexto das tecnologias somáticas, ao mesmo tempo em que conserva a potência de gravação e de difusão da mídia. Isso só vale, no entanto, para a exploração efetiva das possibilidades da informática. Um *compact disk* (CD), um hiperdocumento, mesmo que digitalmente codificado ou que conserve algumas das características interativas próprias do digital, oferece menos plasticidade, dinamismo, sensibilidade à evolução do contexto que um hiperdocumento enriquecido e estruturado em tempo real por uma comunidade de autores e leitores em rede. Segundo Lévy:

O tratamento molecular da informação abre um ciberespaço que interconecta virtualmente todas as mensagens digitais, multiplica os captadores e semáforos, generaliza as interações e os cálculos em tempo real. O ciberespaço tende a reconstituir em escala mais ampla o plano, o continuum indivis, o caldo vivo e flutuante que unia os signos e os corpos, como os signos entre si, antes que a mídia isolasse e fixasse as imagens (Lévy, 1998, p. 54).

As técnicas moleculares opõem-se, portanto, às técnicas molares. O avanço das tecnologias moleculares no tratamento da matéria promete um aumento sem precedentes da produtividade do trabalho humano e uma aceleração das mutações econômicas. A abertura do ciberespaço permite conceber formas de organização econômica e social centradas na inteligência coletiva e na valorização do humano.

No campo da política, Lévy (1998) defende a existência de uma política molecular, em que os grupos seriam auto-organizados e não seriam considerados fontes de energia a ser utilizadas no trabalho, nem forças a explorar, mas, sim, inteligências coletivas que elaboram e reelaboram seus projetos e recursos, refinam suas competências, visando indefinidamente ao enriquecimento de suas qualidades. Dessa forma, seriam opostos aos grupos organizados, molares, que passam por uma transcendência ou exterioridade para se constituir e se manter. À medida que grupos molares se tornam demasiadamente numerosos para que os indivíduos se conheçam por seu nome e possam compreender em tempo real o que fazem juntos, diversos líderes, chefes, reis e representantes unificam e polarizam o espaço do coletivo. As instituições lhes conferem um tempo contínuo e a burocracia torna-se seu órgão separado de gestão e de tratamento da informação. Nas palavras de Lévy:

A transcendência e a separação são igualmente tecnologias molares, pois, nos grupos que se organizam as mudanças se efetuam de maneira custosa, brutal e, muitas vezes, catastróficas: golpes de Estado, revoluções, motins. Durante esses períodos de transição violenta, sob o impacto da revolta, do entusiasmo, da identificação com um chefe carismático, o grupo se funde. Torna-se uma fonte de energia explorada pelos profissionais da mediação (Lévy, 1998, p. 56).

Na esfera do humano, as tecnologias moleculares propõem aos grupos e às pessoas instrumentos que lhes permitam valorizar a si próprias, qualidade por qualidade. Promovem o reconhecimento mútuo e a sinergização das qualidades antrópicas. Assim como a nanotecnologia constrói suas moléculas átomo por átomo, a nanopolítica cultiva cada hipercórtex comunitário de maneira mais fina, mais precisa, mais individualizada possível, favorecendo a conexão das capacidades cognitivas. Assim como as mensagens no ciberespaço interagem e invocam-se de um extremo a outro de um plano desterritorializado, os membros dos coletivos moleculares se comunicam transversalmente, reciprocamente, fora de categorias, sem passar pela via hierárquica. Em relação ao excesso de dados, Lévy aponta:

Entretanto, para responder à aceleração da mudança, um uso maciço das técnicas digitais de simulação, de acesso à informação em tempo real e de comunicação interativa pode se revelar muito útil, entre as mãos de todos os cidadãos. Como tratar enormes massas de dados concernentes a problemas interligados em situação de mobilidade? Provavelmente adotando estruturas de organização que favoreçam uma verdadeira socialização das resoluções de problemas, em vez de seu tratamento por instâncias isoladas, com risco de rapidamente se tornarem concorrentes, serem engolfadas, superadas e expulsas da disputa. O tratamento cooperativo e paralelo das dificuldades requer a concepção de ferramentas de filtragem inteligente dos dados, navegação em meio à informação, simulação de sistemas complexos,

comunicação transversal e observação recíproca das pessoas e grupos em função de suas atividades e de seus saberes (Lévy, 1998, p. 62).

A democracia só progredirá, portanto, explorando da melhor forma as ferramentas de comunicação contemporâneas. Uma possível objeção a essa proposição tecnopolítica, ainda segundo Lévy (1998), poderia ser que as ferramentas de navegação seriam demasiado caras e difíceis de manejar. A ágora eletrônica seria um luxo elitista reservado a pessoas ricas e educadas. Contudo, os cidadãos poderiam participar de uma administração sociotécnica de um novo tipo, permitindo a grandes coletividades que se comuniquem entre si em tempo real.

O ciberespaço cooperativo deve ser concebido como um verdadeiro serviço público. Essa ágora virtual facilitaria a navegação e a orientação no conhecimento, promoveria trocas de saberes, acolheria a construção coletiva do sentido, proporcionaria visualização dinâmica das situações coletivas. Permitiria, enfim, a avaliação por múltiplos critérios, em tempo real, de uma enorme quantidade de proposições, informações e processos em andamento.

Nesse sentido, o ciberespaço poderia se tornar o lugar de uma nova forma de democracia direta em grande escala, visto que o atual modo de expressão permitido pelo voto é grosseiro e descontínuo, possibilitando pouca iniciativa por parte dos cidadãos, principalmente pelo fato de que eleições importantes só ocorrem a cada quatro ou cinco anos. Já um dispositivo de democracia direta em tempo real, no ciberespaço, permitiria a cada um contribuir de maneira contínua para a elaboração e o aperfeiçoamento dos problemas comuns, para a abertura de novas questões, para a formulação de argumentos, para enunciar e adotar posições independentes umas das outras sobre grande variedade de temas.

3.1 Fundamentos conceituais do mundo orientado por dados

Dataísmo, dataficação ou digitalização são conceitos recentes impostos pelo aumento vertiginoso de dados que circulam na internet, caracterizando o fenômeno emergente de *Big Data*. A conceituação deste termo é fundamentada em cinco aspectos, classificados como “cinco Vs” (Stochero, 2020): volume; velocidade; variedade; variabilidade; e vínculo. Diariamente são originados dados de variadas fontes que constituem um volume gigantesco de informações. Tais dados são gerados, atualizados, utilizados e compartilhados de forma rápida. Além disso, os dados têm formatos, cores e pesos variados. São produzidos de forma irregular ou sazonal, mas em velocidade e variedade que aumentam constantemente, influenciando em sua inconsistência.

Apesar de imbricados, os termos “*Big Data*”, “dataficação” ou “digitalização” e “dataísmo” não são sinônimos. O *Big Data* é definido como o agigantamento crescente dos dados gerados, armazenados e disponibilizados pelos meios digitais. A dataficação (Mayer-Schönberger; Cukier, 2013) ou digitalização (Lemos, 2021, 2023), por sua vez, designa a transformação em dados de todos os aspectos da vida em sociedade, de forma ampla e onipresente (Segata; Rifiotis, 2021), colocando os eventos em um formato quantificável, permitindo que sejam analisados e comparados; em outras palavras, trata-se de transformar as atividades cotidianas em informação, esta percebida como nova forma de valor (Mayer-Schönberger; Cukier, 2013). Segundo Schönberger e Cukier: “No futuro – bem mais cedo do que pensamos – muitos aspectos do nosso mundo serão aumentados ou substituídos por sistemas informáticos que hoje são da exclusiva competência do julgamento humano” (Mayer-Schönberger; Cukier, 2013, p. 12 - tradução nossa).

A dataficação afeta todas as áreas da vida social contemporânea, conforme exemplifica Lemos, apontando aquelas que chama de “megaplataformas”:

[...] saúde e educação (Mascheroni 2020), finanças e trabalho (Gidaris 2019), governo (Gozman, Currie, e Seddon 2015; Hong et al. 2019; Redden 2018), transportes (Schor et al. 2020), redes sociais (Helmond 2015; Plantin e Punathambekar 2019), jornalismo e produção cultural (Nieborg e Poell, 2018), para citar algumas (Lemos, 2021, p. 195).

O dataísmo, por sua vez, é um termo utilizado pela primeira vez por David Brooks, do *New York Times*, em 2013, que passou a se referir ao culto aos dados como fonte suprema de compreensão do mundo, visão originada entre tecnólogos do Vale do Silício. O dataísmo tem sido criticado como uma nova forma de religião, especialmente pelo uso ininterrupto das redes e as novas contradições do capitalismo que são alimentadas por essa prática, bem como pelo culto aos dados propriamente ditos:

À semelhança de qualquer religião, o dataísmo tem mandamentos práticos. Um seguidor dessa nova religião, o dataísta, deve querer “maximizar o fluxo de dados, interligando-se cada vez mais”, pois acredita piamente na liberdade de informação e confia mais no big data e nos algoritmos de computador do que no conhecimento e sabedoria humanos (Kaufman; Santaella, 2021, p. 4).

A transformação ocorre em tal medida que decisões relevantes estão sendo tomadas por meio do *Big Data* e da Inteligência Artificial (IA) no atual regime de informação ou infocracia (Han, 2022). A influência de informações geradas a partir de conjuntos de dados, coletadas e tratadas para fins específicos, escreve capítulos da história mundial, conforme pôde ser

observado nos resultados das eleições presidenciais dos Estados Unidos em 2016 e no Brasil em 2018, quando foram eleitos Donald Trump¹⁵ e Jair Bolsonaro¹⁶, respectivamente:

Os dados, indiscutivelmente o bem mais valioso do mundo, estão sendo armados para travar guerras culturais e políticas. O obscuro mundo da exploração de dados é descoberto através das jornadas pessoais imprevisíveis de jogadores em lados diferentes da história explosiva de dados da Cambridge Analytica / Facebook (Privacidade Hackeada, 2019).

Diante do volume crescente de informações, vem se tornando imperativa a integração entre dados e diferentes plataformas. Assim, os dados tornam-se um produto estratégico para empresas (Stochero, 2020) e para a sustentação do capitalismo da informação (Han, 2022) ou capitalismo de dados, capitalismo de vigilância (Zuboff, 2018) ou neocolonialismo digital (Kaufman; Santaella, 2021), entre outras nomenclaturas. Segundo Kauffman e Santaella:

Sessenta anos atrás, os computadores digitais tornaram a informação legível. Trinta anos atrás, a internet tornou-a acessível. Vinte anos atrás, o primeiro motor de busca transformou-a em um único banco de dados. Agora as empresas Google e afins estão tratando esse corpus massivo como um laboratório da condição humana. Essas empresas são as filhas geradas pelos petabytes (Kauffman; Santaella, 2021, p. 2).

Nesse contexto, o jornalismo não fica à margem e vem sendo afetado e reconfigurado. Na era do *Big Data*, as transformações, os hábitos e a cultura da sociedade continuam sendo refletidas pelo jornalismo. Como observam Roberto, Fidalgo e Buckingham (2015), a sociedade da informação incorpora e provoca mudanças no mundo, gerando novas dinâmicas econômicas, sociais e culturais. A utilização exponencial das novas tecnologias reflete-se, por exemplo, nas crescentes taxas de adesão à internet, que pela sua versatilidade afeta, direta ou indiretamente, a sociedade em geral. Em 2024, a população mundial totalizou oito bilhões de pessoas, sendo 0,9% ou 74 milhões acima de 2023. Do total de habitantes do planeta, 5,61 bilhões ou 69,4% possuem conexões por celular, sendo 2,5% ou 138 milhões a mais que no ano anterior. Além

¹⁵ Donald Trump foi eleito presidente dos Estados Unidos em 2016. Durante sua campanha presidencial como candidato da extrema direita, a empresa britânica Cambridge Analytica ofereceu seus serviços de consultoria após obter indevidamente acesso a dados de 87 milhões de usuários do *Facebook*, incluindo de 443.117 brasileiros, com o intuito de traçar perfis psicológicos dos usuários da rede social e fazer propaganda política. Esse fato foi denunciado pelos jornais *The New York Times* e *The Guardian*, levantando dúvidas sobre a transparência e o compromisso da Meta – empresa proprietária do *Facebook* – com a proteção de dados dos usuários.

¹⁶ Jair Bolsonaro foi eleito presidente do Brasil em 2018. Antes, em agosto de 2016, um dos filhos do então candidato brasileiro à presidência esteve nos Estados Unidos, onde se reuniu com Steve Bannon, principal estrategista de campanha de Donald Trump. A campanha do presidencial brasileiro da extrema direita, Jair Bolsonaro, foi caracterizada como uma guerrilha virtual com um esquema industrial de propagação de notícias falsas (*fake news*), financiadas por empresários bolsonaristas.

disso, 5,36 bilhões de pessoas ou 64,4% da população mundial utiliza a internet. O número equivale a um crescimento anual de 1,9% ou 98 milhões de indivíduos. Em relação ao uso de redes sociais, são 5,04 bilhões ou 52,3% da população mundial, contingente 5,6% superior a 2023 e que representa a adesão de 266 milhões de novos adeptos, segundo informa o *DataReportal* (Kemp, 2024). O relatório, que reúne estatísticas relativas ao uso de internet e redes sociais, aponta ainda que em janeiro de 2024 havia 187,9 milhões de conexões de internet no Brasil, quando a penetração era de 86,6%. No intervalo temporal de doze meses foram incorporados 6,1 milhões de usuários, contingente que representa um acréscimo de 3,35% em comparação com o início de 2023, quando havia 181,8 milhões de brasileiros conectados à *web* e a penetração da internet era de 84,3%.

Diante da possibilidade de conversão de fotos, textos, sons e imagens em sistemas binários, reconhecidos por dispositivos eletrônicos como computadores e celulares, o jornalismo vem se apropriando de tais dados para produção e difusão de notícias. Na propagação de informações, as mídias sociais têm sido uma alternativa cada vez mais buscada pelos veículos de comunicação.

Os números do *DataReportal* (Kemp, 2024) mostram que havia 144 milhões de identidades ativas para uso dos serviços de mídia social no Brasil em janeiro de 2024, número equivalente a 66,3% da população total. Constatamos, pelos dados, um aumento de dois milhões ou 1,4% de utilizadores de mídias sociais entre o início de 2023 e 2024. Faz-se importante registrar que os perfis nas mídias sociais podem não representar indivíduos únicos. De forma mais ampla, 76,6% da base total de conexões de internet no Brasil usaram pelo menos uma plataforma de mídia social em janeiro de 2024, sendo 55,6% mulheres e 44,4% homens.

Em relação às conexões móveis no Brasil, são 210,3 milhões de celulares ativos no início de 2024, segundo dados da *GSMA Intelligence*¹⁷. Ressaltamos, ainda, que muitos utilizam mais de uma ligação móvel, motivo pelo qual é comum que o número de ligações móveis exceda o número total de habitantes, com uma teledensidade superior ao contingente populacional. No Brasil, os números das conexões móveis equivaliam a 96,9% da população total em janeiro de 2024, com diminuição de dois milhões ou 1% entre o início de 2023 e 2024.

Cimadevilla (2009) observa que a taxa de adoção da telefonia móvel em todos os países alcançou, nos últimos anos, tendências nunca antes vistas para esse tipo de dispositivo de comunicação, superando claramente a telefonia tradicional e abrangendo todas as classes sociais. Quase o mesmo é possível afirmar sobre a internet e sua versão mais recente de banda

¹⁷ <https://www.gsmainelligence.com/>

larga ou acesso sem fio, mais restrito a certos segmentos sociais.

O ritmo de crescimento da banda larga condiciona o tipo de conteúdo oferecido pelo jornalismo que se faz na *web*, segundo argumentam Roberto, Fidalgo e Buckingham (2015). Desencadeiam, ainda, ajustes na sociedade da informação, com consequentes proposituras de políticas promotoras de acesso às Tecnologias de Informação e Comunicação (TIC) e tentativas de reformulação de programas de aprendizagem.

Persiste, contudo, o reverso da convergência das TIC, ilustrado pelas dificuldades em se garantir acesso universal aos meios tecnológicos que permita reduzir a infoexclusão. A desigualdade de acesso às TIC deve-se, entre outros fatores, às diferenças nos usos que as pessoas dão à informação, aos instrumentos que utilizam, ao nível de intensidade de acesso, sendo que fatores como condições socioeconômicas, idade, gênero e educação podem também ter um peso relevante no agravamento da população infoexcluída.

Apesar do crescimento das conexões de internet no Brasil, 29,1 milhões de pessoas ou 13,42% da população não usava a rede no começo de 2024. Um ano antes, 33,9 milhões ou 15,62% dos brasileiros permaneciam *offline*, considerando-se o contingente demográfico de 217 milhões de pessoas no país. Os dados reunidos pelo *DataReportal* (Kemp, 2024) mostram que o total de brasileiros aumentou 1,2 milhão (0,6%) entre o início de 2023 e de 2024. O relatório destaca, ainda, que 50,9% da população brasileira é feminina e 49,1% é masculina. A ocupação do território nacional está concentrada, majoritariamente, em centros urbanos (87,9%), enquanto apenas 12,1% vive em áreas rurais – outro fator a influenciar no acesso à internet.

Enquanto uma parcela da humanidade usufrui dos benefícios das TIC na sua plenitude, outros permanecem apartados dos avanços que caracterizam o modo de vida da sociedade atual. Idealmente, as TIC devem garantir oportunidades equitativas no seu uso e acesso, permitindo que o conhecimento e a informação sejam fontes universais de bem estar e progresso, argumentam Roberto, Fidalgo e Buckingham (2015), condição defendida também pelos países ibero-americanos signatários da *Declaração de Bávaro* (Cúpula Ibero-Americana, 2002):

Reconhecemos que uma das principais características da mundialização é o desenvolvimento das novas tecnologias da informação e das comunicações, que propiciam uma maior difusão dos conhecimentos, pelo que se impõe o desenvolvimento de planos de cooperação, a fim de que os nossos povos estejam em melhores condições de aproveitar as oportunidades que estas oferecem, e por sua vez enfrentar com êxito o desafio que representa a atual tendência de homogeneização cultural. Em defesa da nossa identidade cultural, enfatizamos a importância do acesso às novas tecnologias de comunicação e informação e o desenvolvimento de programas específicos que permitam diminuir a brecha tecnológica entre os países (Cúpula Ibero-

Americana, 2002).

Posteriormente, o tema foi abordado pelo governo italiano e corporações do ramo da tecnologia no documento *Call for an AI Ethics* (em tradução livre, “Apelo por uma Ética da Inteligência Artificial”), emitido em 2020, no qual se apela pelo combate à infoexclusão:

As transformações atualmente em curso não são apenas quantitativas. Acima de tudo, são qualitativas, pois afetam o modo como essas tarefas são realizadas e o modo como percebemos a realidade e a própria natureza humana, tanto que elas podem influenciar os nossos hábitos mentais e interpessoais. Novas tecnologias devem ser pesquisadas e produzidas de acordo com critérios que assegurem que elas realmente sirvam a toda a “família humana” (Preâmbulo, Declaração Universal dos Direitos Humanos), respeitando a dignidade inerente a cada um de seus membros e todos os ambientes naturais, e levando em conta as necessidades daqueles que são mais vulneráveis. O objetivo não é apenas garantir que ninguém seja excluído, mas também expandir aquelas áreas de liberdade que podem ser ameaçadas pelo condicionamento algorítmico (*Call for an AI Ethics*, 2020).

Em junho de 2024, durante participação inédita em reunião do G7¹⁸, o Papa Francisco observou que, no contexto da sociedade da informação, a IA é um dos instrumentos cada vez mais incorporados em diversas áreas do conhecimento, como comunicação, política saúde e educação, moldando o modo de vida da humanidade e a maneira como a própria identidade humana é concebida. As transformações vão além de um progresso científico, resultando em uma revolução cognitivo-industrial que cria um novo sistema social marcado por complexas transformações epocais. O pontífice também destacou a necessidade de controle humano sobre a escolha de programas de inteligência artificial para preservar a própria dignidade humana, sua liberdade e responsabilidade. Defendeu, além disso, uma moderação ética dos algoritmos, que devem ser ordenados para o bem de cada ser humano:

A inteligência artificial poderia permitir a democratização do acesso ao conhecimento, o progresso exponencial da pesquisa científica, a possibilidade de delegar trabalhos árduos às máquinas; mas, ao mesmo tempo, poderia trazer consigo uma injustiça maior entre nações avançadas e em desenvolvimento, entre classes sociais dominantes e classes sociais oprimidas, colocando em perigo a possibilidade de uma “cultura do encontro” e promovendo uma “cultura do descarte” (Papa Francisco, 2024).

Uma abordagem “algorética” – combinação de ética com tecnologia – é fundamentada em cinco princípios, sendo eles: transparência, inclusão, responsabilidade, imparcialidade,

¹⁸ Grupo dos Sete, organização informal de líderes de algumas das maiores economias mundiais: Canadá, França, Alemanha, Itália, Japão, Reino Unido e Estados Unidos

privacidade, confiabilidade e segurança (Benanti, 2021).

Os dados são um mapa da realidade, representam uma redução da realidade e, por isso, são úteis para tomar decisões. Além disso, as inteligências artificiais trabalham com bancos de dados e sensores. Mas os sensores também não leem toda a realidade: tomam apenas uma parte dela, transformando-a em dados.

Aqui chegamos ao ponto-chave da questão. Como as inteligências artificiais baseiam suas decisões nos dados, e como esses dados não são uma cópia perfeita da realidade, é impensável a priori que a máquina dotada de inteligência artificial possa fazer uma escolha livre de erros. A *Machina sapiens* será sempre e constitutivamente falível.

As inteligências artificiais precisam constitutivamente de uma ética. Como elas podem cometer erros, é preciso entender, então, como gerir esse equívoco. A questão ética é fundamental, importantíssima e urgente. É preciso encontrar um sistema ético compartilhado, para que a utilização desses sistemas não produza injustiças, não prejudique as pessoas e não crie fortes desequilíbrios globais.

A existência de *Machine sapiens* pede que se institua uma nova linguagem universal que saiba traduzir essas diretrizes éticas em diretrizes executáveis pela máquina. Mas como fazer isso? O mundo na era da *Digital Age* é regulado pelos algoritmos. Vários falam de uma algocracia. Para evitar que haja esse domínio do algoritmo também graças às inteligências artificiais, devemos começar a desenvolver essa linguagem comum da algorética (Benanti, 2021).

Combater a infoexclusão requer estratégias de literacia e inclusão digital que possibilitem gerir novas competências e fomentar a coesão e a participação social. Pensar na infoexclusão e na sua proximidade com a desigualdade social implica, também, considerar a forma como as pessoas que nascem na era da informação, ou seja, os nativos digitais, podem ser atingidos por essa problemática. Para Roberto, Fidalgo e Buckingham (2015), a vivência neste novo mundo de informação global vai requerer uma estreita relação com a educação de cariz digital, no sentido de permitir a aquisição e o desenvolvimento de competências que possibilitem acompanhar o progresso tecnológico. Em outras palavras, o acesso à internet deve ser percebido como um novo direito humano, bem como direito à educação, à liberdade de expressão e acesso à informação, uma vez que as TIC alteram a forma como a sociedade vive, aprende e trabalha.

Olhando para o futuro, Donoho (2017) afirma que o alcance e o impacto da Ciência de Dados se expandem à medida que a Ciência de Dados e os dados sobre a própria ciência tornam-se disponíveis de maneira onipresente. A motivação central para a expansão da Ciência de Dados é intelectual. A aprendizagem a partir de dados é um campo do conhecimento impulsionado por motivações científicas e não industriais, por estudar métodos envolvidos na análise e no processamento de dados e propor tecnologias para melhorar tais métodos de forma baseada em evidências.

Anteriormente, Lévy (1998) argumentara que a humanidade passava por uma nova era

de produção e circulação do conhecimento, impulsionada pelas tecnologias da informação. Nesse contexto, o autor propõe o conceito de inteligência coletiva, que é a capacidade das comunidades de compartilharem conhecimentos e cooperarem intelectualmente, potencializada pelo ciberespaço, que se configura como um novo ambiente antropológico no qual se constroem novas formas de sociabilidade, comunicação e pensamento, em que a inteligência coletiva pode se desenvolver.

Dessa forma, surge uma nova ecologia cognitiva, distribuída por toda parte, constantemente valorizada, coordenada em tempo real e que resulta em uma mobilização efetiva das competências. Não está centralizada em uma pessoa ou instituição, mas se manifesta na colaboração em rede. O saber deixa de ser algo acumulado por especialistas e passa a ser um processo aberto e em constante construção coletiva. Nesse sentido, poderia representar o fim do saber hierárquico, em que a estrutura vertical do saber seria substituída por redes horizontais de troca de conhecimento. Cada indivíduo deixaria de ser apenas consumidor e se tornaria também produtor de informação:

Com o projeto da inteligência coletiva, damos sequência à proposta de emancipação da filosofia das Luzes. E, no entanto, mais uma vez, não podemos mais manter a ficção de um progresso linear, automático, garantido. Neste final de século XX, existe um extraordinário desencanto em relação à “modernidade”. O arcaico, o bárbaro estão à espreita, prestes a surgir ainda mais arcaicos e bárbaros do que jamais foram. Tudo coexiste: a mundialização (que faz agora de toda guerra uma guerra civil) e os fanatismos nacionais; as máfias triunfantes e os refinamentos da bioética; o continente cultural transversal da juventude urbana, suas insígnias e músicas, e o trabalho infantil; a fome e as megamáquinas mundiais de produção de sonhos das indústrias de diversão interativa; as multinacionais de alta tecnologia e a escassez de água; o ciberespaço e o analfabetismo...O tempo não é linear, é múltiplo, em espiral, em turbilhões. Talvez não sejamos pós-modernos, talvez não vivamos depois, mas antes da história, enquanto todas as durações estão ainda mescladas, momento fabuloso, fonte de uma história futura que ainda não começou. Viveríamos no “tempo das origens”, na própria *arché*, no interior do tempo do mito, na grande época das metamorfoses e dos animais que falam. Ritmos, espaços, identidades possíveis são marcas sobre os dados de marfim agitados pelo tempo. Não Cronos, o horrível deus que come seus filhos, o castrador de seu pai, o deus da sucessão linear, mas *aion*, o tempo dos tempos, a eternidade, a inocência. O Tempo é uma criança que joga dados. Que humanidade resultará disso? Um mundo de guerra civil planetária ergue-se sob nossos olhos, dominado por redes do crime e pelas elites high-tech, condenando a maioria dos seres humanos a uma miséria sem esperança. Com o projeto da inteligência coletiva queremos trilhar outro caminho. Mais uma vez, é possível? Pode-se escapar à luta pelo poder, aos projetos de dominação, à guerra? (Lévy, 1998, p. 210).

Na economia do porvir delineada por Pierre Lévy (1998), o verdadeiro capital será o ser humano em sua totalidade, não apenas como força de trabalho, mas como inteligência sensível,

criativa e relacional. Em um contexto no qual a produção material torna-se cada vez mais automatizada e os ofícios ligados ao tratamento da informação são progressivamente assumidos por redes inteligentes, o que permanece como campo irreduzível à máquina é justamente o humano. A criação de mundos possíveis, a imaginação, o vínculo afetivo e a reinvenção de sentidos se afirmam como atividades essencialmente coletivas e não automatizáveis, consolidando-se como a nova fronteira do valor e da ação no ciberespaço.

3.2 Jornalismo baseado em dados: na busca por conceitos e nomenclaturas

A informatização e a emergência do *Big Data* exigem readaptações da sociedade contemporânea e, por consequência, do jornalismo. Como atividade profissional centrada na função democrática e constitucional de informar os cidadãos sobre fatos e temas de interesse público, o jornalismo revê continuamente suas práticas e se reconfigura para obter o máximo de informações e narrativas adequadas à demanda informativa. Concomitantemente com o avanço e a ampliação do acesso às tecnologias de coleta, análise e visualização de dados surgiu e vem se expandindo uma nova especialidade no jornalismo, denominada, a princípio, Jornalismo de Dados (Mancini; Vasconcellos, 2016; Bounegru e Gray, 2021), Jornalismo Guiado por Dados (Träsel, 2014; 2017) e, mais recentemente, Jornalismo Digital em Base de Dados (Barbosa, 2005; 2007; 2008; 2013).

Com origens no Jornalismo de Precisão, que foi impulsionado pela Reportagem Assistida por Computador (RAC), o Jornalismo de Dados surge em 1973 (Träsel, 2014; 2017), consolidando-se a partir de 2010 e evoluindo para os moldes atuais no Brasil por meio de projetos específicos e coberturas jornalísticas guiadas por dados, a exemplo da *Agência Tatu de Jornalismo de Dados*¹⁹, do *DeltaFolha da Folha de S. Paulo*²⁰, do *Nexo Jornal*²¹ e do *InfoAmazônia*²².

Diferentes nomenclaturas foram empregadas para se referir ao Jornalismo de Precisão, incluindo termos como “Jornal Computacional” (Anderson, 2012; Diakopoulos, 2021; Cohen; Hamilton; Turner, 2011; Hamilton; Turner, 2009; Lewis; Usher, 2013; Lima Jr., 2011), além de “Jornalismo Maquínico” (Van Dalen, 2012), “Jornalismo Robô” (Carlson, 2014),

¹⁹ Primeiro veículo especializado em Jornalismo de Dados do Nordeste: <https://www.agenciatatu.com.br/>.

²⁰ Iniciativa de jornalismo investigativo com dados: <https://www1.folha.uol.com.br/deltafolha/>.

²¹ Jornal digital com dados e estatísticas: <https://www.nexojornal.com.br/>.

²² Veículo independente que utiliza dados e mapas para produção de reportagens geolocalizadas na Amazônia: <https://infoamazonia.org/>.

“Jornalismo Algorítmico” (Dörr, 2015) e “Jornalismo Automatizado” (Graeffe, 2016), segundo afirmam Boselli, Ferreira e Varão (2022).

As principais fases evolutivas do Jornalismo de Dados incluem, portanto, o Jornalismo de Precisão e a Reportagem com Auxílio do Computador (RAC). Na década de 1970, o Jornalismo de Precisão emergiu, utilizando análise de dados para abordar temas complexos. Philip Meyer, um dos pioneiros na modalidade, usou essas técnicas para analisar distúrbios raciais em Detroit, trabalho que lhe rendeu um Prêmio Pulitzer.

A Reportagem com Auxílio do Computador (RAC) foi a nomenclatura utilizada para definir técnicas de coleta de dados públicos de forma sistemática, apurada e analítica sobre questões sociais, na década de 1950. Ainda em 1952, a rede CBS News usou um computador Remington Rand Univac (*Universal Access*) para prever a vitória de Dwight Eisenhower nas eleições presidenciais estadunidenses, iniciando um novo marco no jornalismo assistido por computadores. De acordo com Boselli, Ferreira e Varão:

Em 1968, Clarence Jones, do jornal The Miami Herald, contratou estudantes do curso de Direito para inserir registros jurídicos em um computador para análise de preconceito racial no condado de Miami-Dade. Essa análise resultou no trabalho *A Scientific Look at Dade Crime* que apontou discrepâncias entre os números e taxas de crimes relatados na cidade e as prisões realizadas em diferentes delegacias, tornando-se o primeiro trabalho jornalístico de análise de dados governamentais com apoio de computadores (Boselli; Ferreira; Varão, 2022, p. 2).

Outra expressão mais usual e contemporânea definidora dessa especialidade profissional que consiste no uso de dados para produção de notícias é “Jornalismo de Dados” (Bounegru; Gray, 2021; Parasie; Dagiral, 2013).

Abreviadamente, os conceitos de Jornalismo de Dados (JD), Jornalismo Guiado por Dados (JGD) e Jornalismo Digital em Base de Dados (JDBD) sintetizam práticas profissionais cujo ponto em comum é o uso de bases de dados como principal fonte de informação para a produção de conteúdo noticioso.

Na revisão bibliográfica realizada para esta pesquisa, notamos que não foi estabelecida diferenciação conceitual entre Jornalismo de Dados (JD) ou *Data Journalism* e Jornalismo Guiado por Dados (JGD) ou *Data-Driven Journalism*. Segundo Träsel (2014; 2017), jornalistas e alguns pesquisadores usam os termos como sinônimos. São atribuídas diferenças, contudo, a JD, JGD e JDBD, aspectos que apontaremos a seguir.

Optamos, neste estudo, pela adoção do conceito de Jornalismo Digital em Base de Dados (JDBD), por abarcar de forma mais específica o formato digital, que compõe o objeto

empírico desta pesquisa. Apesar de ainda haver poucos estudos relacionados ao assunto²³, a institucionalização e a socialização das modalidades JD, JGD e, mais recentemente, do JDBD, vêm acontecendo por meio de várias iniciativas, como conferências, encontros e comunidades *online*, que proporcionam trocas de experiências e capacitação. Bounegru e Gray (2021) identificam esse movimento de integração e formação no campo do jornalismo, consolidando a técnica de produção de notícias e reportagens assentada em base de dados como uma nova competência.

Nesse sentido, uma diferenciação trazida por Mancini e Vasconcellos (2016) é a subdivisão em Jornalismo **de** Dados (JD) ou Jornalismo Guiado por Dados (JGD) e jornalismo **com** dados. Os autores consideram que o Jornalismo de Dados (JD) ou Jornalismo Guiado por Dados (JGD) é exercido quando informações quantificáveis guiam a construção do discurso. Diferencia-se do jornalismo com dados, limitado à simples utilização de dados nas publicações noticiosas, sem que esses sejam o foco da narrativa jornalística. O JD ou JGD exige a coleta, organização e exploração dos dados para identificar relações que podem ser significativas. Essa modalidade tem nuances de uso de dados quantitativos nos conteúdos jornalísticos e contempla, ainda, as diferenças de abordagens entre uma notícia ou reportagem construída e orientada por base de dados daquelas que apenas incluem algumas informações quantificáveis em meio à narrativa.

Träsel (2014; 2017) adota especificamente o JGD e o define como um conjunto de práticas profissionais que comungam de uma semelhança: o uso de base de dados para produção de notícias. Envolve técnicas de Reportagem com Auxílio do Computador (RAC), visualização de dados com infografia, criação e manutenção de bases de dados e a política de acesso à informação e transparência pública de governos. O JGD vai além da apuração tradicional de notícias ao buscar não só registros numéricos como fontes de informação, mas também por tornar os dados o tema central da história e sua apresentação. Ele reside na intersecção de três áreas: primeiro, jornalismo visual ou infografia; segundo, multimídia e narrativas interativas; terceiro, jornalismo investigativo.

O termo jornalismo guiado por dados se refere à aplicação da tecnociência às rotinas produtivas do jornalismo informativo, num contexto de comunicação em rede, com o fim de estabelecer reclamações mais sólidas de autoridade profissional. Essas reclamações de autoridade profissional se baseiam na observação dos procedimentos de construção da objetividade, não como mero ritual (TUCHMAN, 1993), mas com vistas, em seus melhores momentos, à

²³ Conforme revisão bibliográfica em bases de dados dos repositórios *Capes*, *Google Acadêmico*, *Scopus* e *Mendeley* com os termos “jornalismo de dados”, “jornalismo guiado por dados”, “jornalismo digital em base de dados”, “*data journalism*”, “*data-driven journalism*” e “periodismo de datos” em julho de 2024.

objetividade disciplinar (PORTER, 1995), embora em geral apenas uma objetividade mecânica seja suficiente (Träsel, 2017, p. 14).

Para Matzat (2011), o JGD coaduna técnicas que ressaltam o caráter dinâmico da apuração e apresentação de informações através da navegação e interação com bancos de dados *online*. Conforme apontamos, essa prática se baseia em registros numéricos não como fontes de pesquisa, além de tornar os dados o tema central da história e sua apresentação. Essa modalidade possibilita ao jornalista encontrar informação com valor noticioso em bases de dados com milhares ou milhões de registros, dificilmente manejáveis sem a ajuda de computadores. Facilita, ainda, o trabalho de comparar diferentes bases de dados, para a produção de novo conhecimento sobre a sociedade, criando *mashups*²⁴ ou relatando os resultados em texto, imagem e conteúdo audiovisual.

Portanto, bases de dados são entendidas, no âmbito do JGD, como fontes de informação para a produção de narrativas que podem ser apresentadas em formatos digitais ou tradicionais, como papel ou ondas eletromagnéticas, uma vez que apurações guiadas por dados também resultam em matérias para jornais e revistas ou emissoras de rádio e televisão.

O conceito de Jornalismo Digital em Base de Dados (JDBD), por sua vez, apreende que todos os produtos são expressões de código binário armazenado em computadores. Os bancos de dados não são apenas fontes de informação, mas compõem a base material sobre a qual todos os produtos se assentam. Barbosa e Torres (2013) consideram o JGD como um fenômeno pertencente ao campo do JDBD, caracterizando-o como uma extensão ou ampliação das práticas deste.

O JDBD envolve algoritmos, linguagens de programação, metadados e *data mining*, tendo uma base tecnológica ampliada, estando amparado na plataformização, com equipes especializadas, sistemas de gestão de conteúdo; *blogs*; interação do usuário/leitor na produção de conteúdo; automatização; sites dinâmicos e não mais estáticos; narrativas multimídias; infografia interativa; e emprego do RSS (*Really Simple Syndication* ou *Rich Site Summary*) para distribuição de informações em tempo real pela internet.

No JDBD são compilados e compartilhados conteúdos em áudio e vídeo, utilizando uma variedade de formatos de mídia, como texto, áudio, vídeo, infográfico e realidade virtual,

²⁴ Um *mashup* é compreendido, no contexto da informática, como uma aplicação para a *web* produzida pelo amálgama de duas ou mais fontes diferentes, gerando um terceiro produto. Em geral, as fontes colocam os dados voluntariamente à disposição de usuários, através de APIs (*Application Programming Interface*), isto é, há uma abertura que permite a programadores usarem algumas funcionalidades de um determinado *software*. Uma lista com exemplos de *mashups* pode ser encontrada na página: <http://www.programmableweb.com/apis/directory/1?sort=mashup> (Träsel, 2017).

ligados às iniciativas móveis, em um “*continuum* multimídia”.

O modelo que tem as bases de dados como definidoras da estrutura e da organização, bem como da composição e da apresentação dos conteúdos de natureza jornalística, de acordo com funcionalidades e categorias específicas, que também vão permitir a criação, a manutenção, a atualização, a disponibilização, a publicação e a circulação de cibermeios dinâmicos em multiplataformas (Barbosa; Torres, 2013, p. 155).

Na revisão e atualização do conceito de JD, Barbosa (2005; 2007; 2008; 2013) propõe o Paradigma Jornalismo Digital em Base de Dados a partir da terceira para a quarta fase de transição geracional do ciberjornalismo.

Incorporando novas tecnologias, evoluindo a partir de um método experimental até as formas mais recentes de automação e análise de dados, o JDBD vem ampliando sua capacidade investigativa e inovando nas narrativas:

Atualmente, muitos estudos – circunscritos na área das novas mídias, da cibercultura, da comunicação, do jornalismo, das artes, das ciências da computação, da inteligência artificial, da economia – tem delineado perspectivas e apontado funções para as bases de dados que descortinam as novas linhas de forças, os modos de operação, de representação, as potencialidades que as atravessam e as colocam num patamar diferenciado (Barbosa, 2007, p. 20).

O JDBD objetiva a produção, o tratamento e o cruzamento de grandes quantidades de dados, de modo a permitir maior eficiência na recuperação de informações, na apuração de reportagens a partir de conjuntos de dados, na distribuição em diferentes plataformas por meio de computadores pessoais, *smartphones*, *tablets* e na geração de visualizações e infografias interativas.

Brenol (2019) reforça que o conceito de JDBD proposto por Barbosa (2005; 2007; 2008; 2013) é amplo e abrangente para compreender a relação funcional das bases de dados para o jornalismo em todas as dimensões, diferente do JGD, mais instrumental, metódico e técnico.

O JGD assume uma postura exploratória, característica do leitor imersivo experto descrito por Santaella (2019), que busca a informação na internet a partir do raciocínio dedutivo, deixando-se guiar tanto pela experiência quanto por pistas disponíveis na rede para encontrar seu objetivo. A imersão proporcionada pela experiência de interação com hipertextos mobiliza a cognição do leitor e o incentiva a adotar uma atitude dedutiva. O contato com aplicativos ou gráficos é uma forma de transferir o controle da narrativa para a audiência, conduzindo o leitor à problematização e reflexão sobre seu conteúdo.

Mancini e Vasconcellos (2016) consideram que a internet incentiva o desenvolvimento de publicações que permitem processos interativos entre os receptores e o conteúdo. A visualização interativa dos dados, ainda que desacompanhada de um texto narrativo, permite a participação ativa na interpretação de tais informações pelos leitores.

Além do conceito, outra nomenclatura contemporânea é a de “Jornalismo Automatizado”, proposta por Dalben (2019) e resultante da *Natural Language Generation* (NLG), campo da Ciência da Computação dedicado a realizar experiências de tradução e criação de pequenas frases com estruturas gramaticais simples:

Diferente dos robôs humanoides que vemos nos filmes e nos romances de ficção científica, estes jornais adotaram softwares de Natural Language Generation (NLG), um subcampo da Inteligência Artificial e da linguística computacional, para publicar de maneira automatizada notícias curtas com estruturas repetitivas sobre esportes, finanças, eleições, crimes, terremotos, trânsito, para citar alguns exemplos (D’Andréa & DalBen, 2017). Todas seguem uma narrativa pré-determinada escrita por um jornalista e são formatadas e, posteriormente, publicadas após o processamento de informações fornecidas por grandes bancos de dados (Dalben, 2019, p. 3).

A partir dos anos 2000, o uso de IA e aprendizado de máquina permitiu aprimorar a coleta e análise de dados. *Softwares* como *Narrative Science* e *Automated Insights* começaram a produzir notícias automaticamente. Além disso, no Brasil, “as primeiras discussões sobre a proposta de automação da redação de leads de notícias tiveram início com Nilson Lage (1997, 2003), evidenciando a inviabilidade de automação para reportagens em razão da extensão e profundidade peculiares à forma narrativa” (Boselli; Ferreira; Varão, 2022, p. 3).

Apesar disso, agências de checagem como Aos Fatos têm experimentado *bots* para automatizar a verificação de fatos. As etapas evolutivas tecnológicas, portanto, transformam o jornalismo em várias dimensões, incluindo como os profissionais da área trabalham, o conteúdo das notícias, a organização das redações e as relações com o público.

3.3 Rotinas produtivas do jornalismo baseado em dados

Arrese (2022) considera que o JGD move o jornalismo desde o princípio, particularmente o jornalismo econômico, que reúne cinco aspectos centrais dessa modalidade, sendo elas: trabalho com dados e bases de dados; desenvolvimento de uma infraestrutura conceitual para análise de dados e narrativa; uso regular de ferramentas de visualização; aplicação de novas tecnologias às peculiaridades dos dados econômicos; e integração de diferentes perfis profissionais nas redações.

Na conceituação de JD, o jornalista Paul Bradshaw (2011) propõe um diagrama com cinco etapas que caracterizam a rotina produtiva do que ele chama de “Pirâmide Invertida do Jornalismo de Dados”. Pensando no processo de produção de uma notícia ou reportagem alicerçada em dados, é preciso ter certeza da qualidade dos dados coletados, sendo necessária a limpeza e padronização das informações quantificáveis, eliminando possíveis falhas humanas. Em seguida, deve ser feita a contextualização dos dados. E, ainda, faz-se necessário recorrer a múltiplas fontes, combinando mais de uma base de dados, que poderá incluir a visualização por meio de infográficos, gráficos ou mapas.

A rotina produtiva do que Träsel (2014; 2017) considera como JGD obedece ao esquema de pirâmide invertida de Bradshaw (2011), cumprindo as etapas de compilação (*compile*), limpeza (*clean*), contextualização (*contexto*) e combinação (*combine*), além da comunicação (*communicate*), sendo esta última desdobrada em seis características.

A compilação consiste na coleta de dados. Segundo Bradshaw (2011), as bases de dados podem ser construídas pelos jornalistas através de pesquisa de campo por meio de técnicas qualitativas e quantitativas, como questionários, análise de documentos e etnografia. Podem, ainda, ser obtidas junto às fontes de informação, como instituições e indivíduos, que podem disponibilizar planilhas contendo os dados buscados pelo jornalista. Planilhas de dados podem também ser geradas através de *web scraping*²⁵. Essa etapa envolve o processamento dos dados para posterior análise que possibilite encontrar informações com valor-notícia.

No processo de busca pelo valor-notícia, o jornalista de dados busca responder a perguntas que exigem dados ou encontra-se em posse de dados que exigem uma investigação.

Na sequência, o jornalista deve realizar a limpeza da base de dados, processo similar à checagem durante a produção de uma reportagem tradicional. Essa ação tem o objetivo de eliminar erros humanos e transformar os dados para que apresentem um formato coerente e eventualmente passível de combinação e cruzamentos com outras bases de dados. Esse processo evita erros que poderiam induzir a equívocos matemáticos, factuais e dificuldades para visualização dos dados.

Na contextualização da base é feita uma análise de possíveis vieses dos dados, que

²⁵ *Web scraping* ou “raspagem de dados” é a prática de copiar ou transformar páginas da *web* em bases de dados, usando aplicativos específicos. Com isso, é possível automatizar o processo de coleta de dados em *websites* que não os colocam à disposição em formatos diretamente manejáveis, como “CSV”, “XLS” ou “TXT”. Os dados a serem raspados podem estar integrados a um documento “HTML”, ou em formatos ilegíveis por aplicativos de tratamento de dados, como “DOC” ou “PDF”, ou ainda distribuídos em diversas páginas. A alternativa à “raspagem de dados” é copiar os dados manualmente, página por página, linha por linha, usando o recurso “copia e cola” para inseri-los diretamente numa planilha do *Excel* ou programa similar.

deverão ser contrastados e contextualizados com outras informações para produzir uma narrativa relevante.

Por último, é realizada a combinação de bases de dados com o objetivo de produzir um novo conhecimento. Para Bradshaw (2011), combinar bases de dados é similar ao procedimento de entrevistar múltiplas fontes durante uma apuração e produção jornalística tradicional, com a diferença, porém, de que possibilitará alicerçar um ângulo jornalístico sobre determinado fato.

Após as seis etapas da rotina produtiva no JGD é realizada a publicação ou comunicação dos resultados, que consiste em seis possibilidades: a) visualização (*visualize*) por meio de infografia ou cartografia; b) narração (*narrate*) em formatos tradicionais, como texto ou audiovisual; c) comunicação social (*social communication*), ou seja, compartilhamento dos resultados com a audiência; d) humanização (*humanise*), a partir de entrevistas com indivíduos que ilustrem as informações obtidas no processamento dos dados; e) personalização (*personalise*) com abertura dos produtos à interação com o leitor; e f) utilização (*utilise*), que envolve a criação de ferramentas ou serviços de informação a partir dos dados coletados.

O processo produtivo no JGD também é abordado pelo estatístico e jornalista Nate Silver, fundador do site *FiveThirty Eight*, em 2014, conforme registram Bounegru e Gray (2021). Silver estabelece quatro etapas, similares às fases de produção de notícias e reportagens no jornalismo tradicional (*apud* Träsel, 2014; 2017). São elas: a) coleta de dados ou evidências por meio de aplicação de questionários, experimentos sociais ou raspagem de dados, sendo um processo equivalente à realização de entrevistas, observação ou produção de documentários do jornalismo tradicional; b) organização, por meio do tratamento e encadeamento dos dados, consistindo na descrição estatística, no armazenamento em bases de dados relacionais e na criação de infográficos, comparada à narração em pirâmide invertida do jornalismo tradicional; c) explicação, ou seja, a interpretação dos resultados da organização dos dados, comparada ao jornalismo interpretativo; d) generalização, estabelecendo previsões a partir da interpretação dos dados, verificando hipóteses por meio de experimentos estatísticos, sendo semelhante ao jornalismo opinativo.

Ainda em relação ao JGD, Gehrke (2021) estabelece seis elementos de transparência, congregando diferentes etapas: coleta, análise, produto, replicabilidade, uso de código aberto e reprodutibilidade. Esses parâmetros servem como guia para um “Protocolo de Transparência no Jornalismo Guiado 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 (Gehrke, 2021, p. 4).

Para Mancini e Vasconcellos (2016), a compreensão do conceito e da prática de JGD é facilitada a partir de uma matriz metodológica, fundamentada em três dimensões, que o jornalismo pode adotar. São elas: dimensão investigativa (atuação proativa na busca de dados e revelações); interpretativa (capacidade ou interesse em expor relações de causas ou consequências entre os dados); e comunicativa (centralidade da visualização dos dados, importante para o leitor/receptor entender a relação entre os dados por meio das imagens). A matriz proposta pelos autores (Mancini e Vasconcellos, 2016) é inspirada no modelo de pirâmide invertida de Bradshaw (2011).

Bertocchi (2005; 2016) acrescenta que o JGD diz respeito ao processo jornalístico que vai da captura de dados e sua curadoria até a visualização em um formato a ser acessado pelos usuários finais nas interfaces digitais. Ele pode emergir de dados estruturados, semiestruturados ou não estruturados. Os dados estruturados compõem uma base, estruturados em blocos semânticos, com atributos definidos, organizados em uma mesma estrutura de representação, como em uma base de dados em “SQL”. Dados semiestruturados, por sua vez, são coleções de dados organizados de forma heterogênea, com estrutura irregular, como o que se vê na *web*, em formatos variados, como “XML”, “RDF”, “OWL”. Por fim, os dados não estruturados são dados sem estrutura definida em documentos variados, como um texto no *Word* ou um arquivo em formato “HTML” simples.

Bertocchi (2005; 2016) contribui com a propositura de um modelo teórico que consiste em um protótipo do sistema narrativo do jornalismo digital, por meio da articulação de dados e metadados em um formato renderizado. Tal modelo é apresentado em três atos, sendo o primeiro a antenarração dos dados, ou seja, o levantamento e a seleção das informações de interesse público guardados nas bases de dados, o fortalecimento do jornalismo de dados e as questões relativas aos *softwares de mídia*. O segundo ato consiste da antenarração dos metadados com a semantificação destes para compreensão de *softwares* e algoritmos, para a apresentação de narrativas nas interfaces digitais e também para a organização interna de dados em sistemas publicadores. Por último, há a narração por meio da organização corporificada na interface digital para acesso dos usuários finais (renderização da narrativa) e demais possibilidades relacionadas, como experiências *cross-channel*²⁶, computação ubíqua com a

²⁶ Personalização de interações em multicanais.

onipresença das tecnologias no cotidiano das pessoas e arquiteturas de sentido. Nas palavras de Bertocchi:

O modelo coloca o sistema narrativo nos seguintes termos: antenarrar é criar e manipular dados e metadados; formatar é interfacear, renderizar a narrativa, conferindo-lhe forma. O formato da narrativa revela o comportamento de todo o sistema. Observamos ainda que um gênero de texto jornalístico (notícia, reportagem, infografia, etc.) pode se desdobrar em vários formatos. O formato ganha corporeidade na interface gráfica exposta na tela, e o que se visualiza na interface é a pele da narrativa (Bertocchi, 2016, p. 208).

Conforme se amplia a compreensão dos jornalistas acerca do agenciamento das camadas do sistema narrativo e de como aplicar estratégias narrativas em mais camadas/estratos da composição da narrativa digital, formatos mais interessantes vão sendo estimulados nas redações em que o jornalista é protagonista da experiência narrativa.

Barbosa (2005; 2007; 2008; 2013) observa que as bases de dados (BDs), mesmo não sendo um meio de comunicação, um espaço visual, social ou urbano, são melhoradas pelo fato de que a internet vai garantir novas técnicas e linguagens para a sua construção e aplicação, de um lado, e sofrerão remediações e possibilitarão rupturas, ganhando novas funcionalidades de acordo com os usos e apropriações no jornalismo digital, de outro.

Quadro 3 – Diferenciações do jornalismo baseado em dados

Características	Jornalismo com Dados (Mancini/Vasconcelos)	Jornalismo de Dados ou Jornalismo Guiado por Dados (Trasel, Matzat, Mancini e Vasconcellos)	Jornalismo Digital em Base de Dados (JDBD) (Barbosa)	Jornalismo de Dados (Gerke/Bertocchi)
Competência Investigativa	Extração, organização, tratamento de dados	Registros Numéricos como fontes de informação; jornalismo investigativo (pesquisa exploratória); coleta, limpeza, análise dos dados;	Bancos de dados são principal fonte de informação;	Coleta
Competência Interpretativa	Narrativa textual, argumentação, relações entre dados	Dados no centro da narrativa jornalística	Dados são base para todos os produtos jornalísticos, que são expressões de código binário.	Análise
Competência Comunicativa	Visualização gráfica dos dados	Visualização dos dados (podem ou não ser apresentados em formatos digitais) Multimídia ou Narrativas interativas.	Bases de dados vão permitir a criação, manutenção, atualização, disponibilização, publicação e circulação de cibermeios dinâmicos em multiplataformas (continuum multimídia)	Produto
Transparência				Uso de código aberto, reprodutibilidade

Fonte: adaptado de Träsel (2014, p. 114).

3.4 Jornalismo de dados e a Lei de Acesso à Informação (LAI)

O direito de acesso é um direito-meio que permite que se conheça os demais, que se exijam condições para exercê-los, e que nos mobiliza como grupos ativos, comprometidos e interessados (Christofolletti, 2019, p. 191).

Na contemporaneidade, o avanço da era digital e a crescente disponibilidade de dados binários e estatísticos induziram transformações nas práticas jornalísticas, conforme apontado por Barbosa (2005; 2007; 2008; 2013). A incorporação de instrumentos normativos, como a Lei Geral de Proteção de Dados (LGPD) (Brasil, 2018) e a Lei de Acesso à Informação (LAI) (Brasil, 2011), aliada ao uso de recursos tecnológicos como visualizações de séries históricas, tabelas dinâmicas e infográficos – inclusive sonoros –, reconfigurou as condições de busca,

tratamento e exposição das informações. Ainda assim, persistem entraves estruturais relacionados à cobertura informacional, à conectividade e à inclusão digital, os quais limitam o pleno exercício do direito à informação.

No que tange ao uso de dados no jornalismo, a transparência (Brenol, 2019; Gehrke, 2021) apresenta-se como requisito indispensável para a manutenção da credibilidade das informações veiculadas. A clareza quanto às fontes utilizadas, aos métodos empregados na análise e às conclusões alcançadas constitui elemento primordial para o fortalecimento do vínculo de confiança entre a imprensa e o público.

Nesse contexto, é necessário considerarmos a trajetória histórica brasileira marcada por um regime autoritário, cujo ápice se deu durante o período militar (1964–1985), caracterizado por censura institucionalizada e restrições severas à liberdade de expressão. Com a redemocratização e a promulgação da Constituição Federal de 1988, o acesso à informação foi consagrado como um direito fundamental, essencial à consolidação de uma esfera pública democrática (Brasil, 1988). Para regulamentar esse direito e operacionalizá-lo no âmbito da administração pública, foi sancionada a Lei nº 12.527, de 18 de novembro de 2011, conhecida como Lei de Acesso à Informação (LAI), cuja aplicação se estende aos três Poderes da União, aos estados, ao Distrito Federal e aos municípios (Brasil, 2011). No âmbito do Poder Executivo Federal, a LAI é regulamentada pelo Decreto nº 7.724, de 16 de maio de 2012 (Brasil, 2012).

Conforme destaca Seibt (2022), a LAI constitui um marco jurídico necessário para o exercício do jornalismo de verificação (*fact-checking*) ao permitir que jornalistas e cidadãos acessem dados públicos e, a partir deles, estruturem conteúdos verificáveis e confiáveis. A curadoria de dados públicos, nesse sentido, é basilar para o enfrentamento da desinformação e para adaptação do jornalismo às exigências de um ecossistema informacional em constante mutação. Para que esse modelo seja plenamente efetivado, contudo, são indispensáveis condições estruturais que garantam não apenas a transparência e o acesso irrestrito às informações públicas, mas também a qualificação técnica e crítica dos profissionais envolvidos na mediação e interpretação desses dados.

Para Brenol e Baccin (2022), o acesso à informação pública em regimes democráticos configura-se como um dos principais instrumentos de participação cidadã, ao permitir que os indivíduos não apenas tomem conhecimento das ações governamentais, mas também possam demandar inovações e melhorias na prestação de serviços à população com base em princípios cívicos. No cenário brasileiro contemporâneo, a tecnologia digital atua tanto como meio de acesso à informação quanto como vetor de transformação na forma como os dados públicos são consumidos e ressignificados, contribuindo para a formação de uma cidadania mais crítica e

bem-informada, dialogando com o conceito de inteligência coletiva de Lévy (1998). A disponibilização de dados públicos e seu consumo ativo pelos cidadãos representam uma forma de cognição compartilhada, na qual o saber não está mais concentrado em instituições estatais ou midiáticas, mas é co-construído por meio de interações em rede. Lévy (1998) destaca ainda que a inteligência coletiva se concretiza quando há valorização mútua das competências individuais, em benefício de um projeto comum – no caso, a formação de uma cidadania crítica e bem-informada. Assim, a ressignificação de dados públicos por meio de ferramentas digitais, como plataformas de dados abertos, visualizações interativas e jornalismo de dados contribui para a transparência, o controle social e a democratização do conhecimento.

4 PROCEDIMENTOS METODOLÓGICOS

França (2016) afirma que o método não é dado por dedução, mas construído de acordo com o problema de pesquisa. Uma boa formulação do problema é primordial e surge de incitações da própria realidade. Transformar, porém, problemas do mundo em questões a serem investigadas e respondidas requer clareza sobre nossa área de estudo e escolha de uma abordagem.

Na era da informação ou do *Big Data*, a realidade é cada vez mais organizada com base nas tecnologias digitais, centralizadas nos algoritmos, que, por sua vez, seguem uma lógica sistemática, organizando uma sequência de informações mediadas por raciocínios, instruções ou operações bem definidas, com um objetivo e visando a um resultado. Sobre os algoritmos, Pinto e Barbosa explicam:

Entenda-se algoritmo como a sequência conceitual de etapas, que deve ser expressa em qualquer linguagem de computador, ou em linguagem humana, ou lógica (Gillespie, 2014). Eles performam processos de mediação, dataficação (Lemos, 2021) e a consequente abundância de informações (Boczkowski, 2021) que nos leva à era do Big Data (Kitchin, 2014; Siegel, 2013), onde bits, sobre os mais diversos aspectos da vida cotidiana, são processados, armazenados, tratados, analisados etc. Para Lemos (2021), a dataficação possibilita a conversão de toda e qualquer ação em dados digitais rastreáveis, produzindo diagnósticos e inferências nos mais diversos domínios. Essa dataficação só é possível devido às tecnologias digitais, que têm em seu âmago os algoritmos (Pinto; Barbosa, 2024, p. 58).

A sociedade atual vive, portanto, em um contexto de ubiquidade algorítmica, que traz a reboque outras características como mobilidade e velocidade. Atualmente tudo se encaixa em algoritmos, desde relacionamentos (redes sociais) até negócios (*e-commerce*) e comunicação, condição potencializada com a pandemia e o rastreamento de contato (*contact tracing*).

O sistema de “inteligência algorítmica” (Lemos, 2021; 2023) é um conjunto de métodos de coleta, processamento e tratamento de dados para realizar previsões. Tais sistemas usam dados criados por eles próprios, resultando continuamente em novos dados e retroalimentando o *Big Data*. Esta performatividade dos dados desempenha papel ativo na sociedade, gerando simulacros participativos, influenciando o comportamento e as decisões de indivíduos e instituições. Somente algoritmos e Inteligência Artificial (IA) são capazes de dar conta da capacidade performática dos dados. A Inteligência Artificial (IA) também tem o poder, portanto, de transformar a realidade, tal como o jornalismo.

Assim, o JDBD se insere numa concepção de comunicação como um processo de

globalidade, com relações cruzadas dos sujeitos entre si, com a linguagem, com os dispositivos, com o contexto, numa relação de mútua afetação e um fenômeno em movimento (França, 2016), condições observadas com as mutações geracionais e hibridização no jornalismo digital (Barbosa, 2005; 2007; 2008; 2013; Mielniczuk, 2001; 2003; 2004; Canavilhas, 2003; 2007; Nunes, 2016).

Sob essa perspectiva, esta pesquisa busca evidências de uma prática regionalizada de Jornalismo Digital em Base de Dados e como ela se configura em dois veículos de comunicação digitais do Centro Oeste, sendo *Gazeta Digital* e *Correio Braziliense*. Nesse sentido, é preciso que consideremos a estrutura dos conteúdos noticiosos a partir de um esforço proativo de investigação de dados, estabelecendo a coleta, a extração, a comparação e a análise estatística aplicada ao jornalismo, visualização interativa de painel de dados – por meio de gráficos, mapas, redes, linhas do tempo – e adequações para produtos diversos, como os aplicativos móveis (apps), incorporando características de mobilidade, multimidialidade e convergência. A produção de conteúdos noticiosos ancoradas em dados deve incluir sistemas *open source* (código aberto), assegurando transparência.

Ponderando, ainda, que os dados não oferecem representações neutras e diretas do mundo por estarem enredados em meio à política e relações de poder, é necessário questionarmos quais são os dados, as fontes dos dados e como são obtidos (Bradshaw, 2011). Além disso, precisamos considerar quais temas estão atrelados ao uso de dados e para quais fins, levando em conta quem é capaz de se beneficiar deles, bem como quais públicos são agrupados através dos dados, quais competências comportam, quais políticas podem sancionar e quais tipos de participação engendram. Nosso estudo buscou identificar alguns desses aspectos, expostos a partir de análise de produções jornalísticas, amparado na criação e aplicação de livro de códigos e atinentes às características do JDBD, tal como propõe Barbosa (2005; 2007; 2008; 2013).

A classificação que compõe o livro de códigos, que desenvolvemos nesta pesquisa, foi fundamentada na análise proposta por Bardin (2011) e nos procedimentos sistemáticos descritos por Sampaio e Lycarião (2021), com o objetivo de oferecer um instrumento metodológico claro e replicável para a categorização de dados em investigações quantitativas e qualitativas. A partir da definição de categorias e subcategorias, organizadas em uma estrutura hierárquica e operacionalizada por meio de descritores e exemplos, o código busca garantir consistência analítica e rigor interpretativo. Essa abordagem possibilita a codificação de conteúdos com base em referenciais teóricos previamente definidos, assegurando a validade e a confiabilidade da análise. Ao alinhar os fundamentos clássicos da análise de conteúdo com técnicas

contemporâneas de codificação, neste caso voltadas para o campo do JDBD, o livro de códigos se propõe a ser uma ferramenta para uma compreensão aprofundada das narrativas midiáticas que compõem o *corpus* desta pesquisa.

Sendo assim, foram classificadas 18 unidades de registro da seguinte forma:

1. data da publicação;
2. total de conteúdos noticiosos na página coletada;
3. total de notícias/reportagens selecionadas para análise;
4. veículo em que foi publicada a notícia/reportagem;
5. *link* na internet de acesso à publicação;
6. assinatura do responsável pela notícia/reportagem;
7. local de apuração das informações publicadas;
8. título da notícia;
9. fontes das informações;
10. gênero jornalístico da publicação;
11. tema/caderno/localização da publicação;
12. recurso visuais - contém imagem (foto ou ilustração);
13. recursos multimídia - contém vídeo;
14. recursos multimídia - contém áudio ou *podcast*;
15. replicação e/ou distribuição multiplataforma/*crossmedia* - compartilhamento via redes sociais;
16. replicação e/ou distribuição multiplataforma/*crossmedia* - configuração para dispositivos móveis;
17. interação (comentários);
18. *hiperlink*.

As 11 unidades de análise, por sua vez, foram classificadas buscando identificar características do conceito de JDBD atualizado para a quinta geração do jornalismo digital, estabelecendo categorias de busca e inclusão de dados, visualização de dados, uso de recursos multimídia, como:

1. características do título - contém dados;
2. características das notícias/reportagens - contém de dados;
3. busca proativa por banco/base de dados;
4. somente reprodução de dados;
5. requer aprofundamento de dados;
6. desinformação com citação sem dados;

7. fonte dos dados nas publicações;
8. solicitação dos dados feita pelo veículo;
9. menciona LAI;
10. menciona LGPD;
11. recursos visuais - contém infográfico (estático ou interativo).

As 29 categorizações estão demonstradas no livro de códigos, conforme exposto a seguir.

Quadro 4 – Livro de códigos categorial – Unidades de registro

Código	Categoria	Descrição	Critério de Inclusão	Exemplo
UR01	Data da publicação	Data em que o conteúdo foi publicado no site	Registro visível na página da notícia	03/05/2025
UR02	Total de conteúdos na página	Número total de conteúdos jornalísticos publicados na página coletada	Contagem manual dos itens disponíveis no <i>site</i> no momento da coleta	15 notícias listadas
UR03	Total de notícias selecionadas para análise	Quantidade de notícias analisadas a partir da amostra total	Contagem dos conteúdos que cumprem os critérios da pesquisa	3/15
UR04	Veículo da publicação	Nome do portal, <i>site</i> ou veículo onde foi publicada a matéria	Identificável pelo logotipo	<i>Gazeta Digital, Correio Braziliense</i>
UR05	<i>Link</i> de acesso à publicação	URL completa da matéria	Deve estar ativa e acessível	https://gazetadigital.com.br
UR06	Assinatura do responsável	Profissional	Função	Repórter
UR07	Local de apuração	Onde a reportagem foi apurada	Indicação explícita no texto	Da redação

Código	Categoria	Descrição	Critério de Inclusão	Exemplo
UR08	Título da notícia	Título principal da matéria	Copiado diretamente da publicação	“Brasil atinge marca histórica de vacinação”
UR09	Fontes das informações	Fontes citadas no corpo do texto	Pode ser órgão público, ONG, empresa etc.	IBGE, Datafolha
UR10	Gênero jornalístico	Classificação do tipo de conteúdo	Notícia, reportagem	Reportagem
UR11	Tema/caderno	Área temática ou seção em que o conteúdo foi publicado	Política, economia, saúde etc.	Economia
UR12	Recurso visual – Imagem	Presença de fotos ou ilustrações	Ao menos uma imagem presente	✓
UR13	Recurso multimídia – Vídeo	Presença de vídeo incorporado	Vídeo visível na página	✓
UR14	Recurso multimídia – Áudio	Presença de áudio ou <i>podcast</i>	Reprodutor de áudio disponível	✓
UR15	<i>Crossmedia</i> – Redes sociais	Possui botões ou menções a redes sociais para compartilhamento	Ícones de <i>Facebook</i> , <i>X</i> , <i>WhatsApp</i> etc.	✓
UR16	<i>Crossmedia</i> – Dispositivos móveis	Estrutura adaptada para <i>mobile</i> (responsividade ou app)	Testado no navegador ou presença de app	✓
UR17	Interação – Comentários	Possibilidade de interação via comentários	Área de comentários aberta ao público	✓
UR18	<i>Hiperlink</i>	Presença de <i>links</i> no corpo da matéria	<i>Links</i> internos ou externos	✓

Fonte: elaborado pela autora (2025).

Quadro 5 – Livro de códigos categorial – Unidades de análise

Código	Categoria	Descrição	Exemplo	Inclusão
UA01	Título com dados	Título menciona números ou estatísticas	“Brasil tem 12 milhões de desempregados”	✓
UA02	Conteúdo com dados	Corpo da matéria contém dados estruturados	Tabelas, números oficiais, pesquisas	✓
UA03	Busca ativa por dados	Jornalista buscou dados em bancos públicos ou APIs	Citação de consulta ao Portal da Transparência, IBGE etc.	✓
UA04	Reprodução de dados	Dados apenas copiados de outra fonte sem análise	“Segundo o IBGE...” sem interpretação	✓
UA05	Requer interpretação	Matéria analisa, cruza ou contextualiza dados	Gráficos, relações causais, explicações	✓
UA06	Desinformação com dados	Uso de dados distorcidos, imprecisos ou manipulados	Números sem fonte, ou interpretados erroneamente	✓
UA07	Fonte dos dados citada	A origem dos dados é claramente indicada	“Fonte: IBGE, 2023”	✓
UA08	Solicitação ativa de dados	O veículo requisitou dados via pedidos formais	Citação à resposta de órgão público	✓
UA09	Menciona LAI	A matéria menciona a Lei de Acesso à Informação	“Dados obtidos via LAI”	✓
UA10	Menciona LGPD	A matéria aborda a Lei Geral de Proteção de Dados	Pode tratar da proteção de dados pessoais	✓
UA11	Infográfico	Presença de infográficos, estáticos ou interativos	Gráficos, linhas do tempo, mapas	✓

Fonte: elaborado pela autora (2025).

Tais classificações, adaptadas do método bardiniano de análise de conteúdo e de atualizações trazidas por Sampaio e Lycarião (2021) no intuito de identificar características de JDBD conforme proposto por Barbosa (2005; 2007; 2008; 2013), foram aplicadas para analisar a prática jornalística da imprensa regional em dois webjornais mantidos por conglomerados de mídias pioneiras em dois entes federados do Centro-Oeste do Brasil, que compõem nosso objeto empírico. Conforme mencionamos anteriormente, requisitos como nível de audiência, longevidade e acessibilidade foram critérios para a seleção dos sites noticiosos *Gazeta Digital*, do Grupo Gazeta de Comunicação, em Mato Grosso e *Correio Braziliense*, do Grupo Diários Associados S/A, no Distrito Federal.

Assim, realizamos a busca dos conteúdos informativos destacados nas capas dos dois portais noticiosos a partir de consultas ao banco de dados digital *Wayback Machine*²⁷.

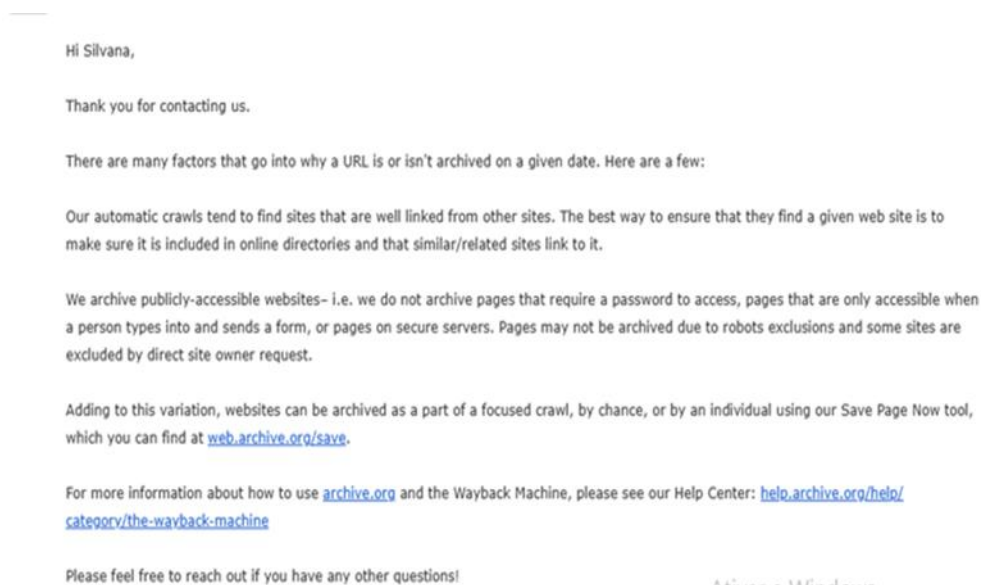
Inicialmente, pretendíamos analisar a cobertura jornalística com uso de base de dados de quatro webjornais do Centro-Oeste a partir dos critérios de tempo ao longo do qual são publicados e vínculo com conglomerados de mídia. O recorte inicial considerou os webjornais *Gazeta Digital*, de Mato Grosso, *Correio Braziliense*, do Distrito Federal, além do *Correio do Estado*, de Mato Grosso do Sul, e *O Popular*, de Goiás. A intenção, à época, era coletarmos e analisarmos, com base em um livro de códigos em pesquisa, as notícias e reportagens publicadas nas capas desses quatro webjornais. Nesse sentido, procedemos à busca por referenciais bibliográficos para contextualização da produção jornalística para formatos digitais e com apropriação de bases de dados em um cenário de crescente digitalização e dataficação. Na sequência, iniciamos o levantamento de conteúdos publicados nas capas dos quatro webjornais dentro do recorte temporal estabelecido na pesquisa. Nesse movimento de coleta do material informativo publicado pelos quatro portais do Centro-Oeste esbarramos, porém, em entraves: instabilidades operacionais na utilização da ferramenta *Wayback Machine* dificultaram estratificar as notícias e reportagens publicadas nos webjornais *Correio do Estado*, de Mato Grosso do Sul, e *O Popular*, de Goiás, conforme previamente definido para este estudo.

De acordo com esclarecimento da *Internet Archive*, a ausência de determinadas páginas na *Wayback Machine* está relacionada a diversos fatores técnicos e operacionais. O serviço realiza rastreamentos automáticos que priorizam sites amplamente referenciados por outros domínios, o que favorece a visibilidade e o arquivamento. No entanto, não são arquivadas páginas que exigem autenticação, estão protegidas por formulários, utilizam servidores seguros ou possuem restrições via “robots.txt”. Além disso, alguns sites podem ser excluídos a pedido

²⁷ Disponível em: <https://archive.org/web/>. Acesso em: 19 de set. 2025.

dos próprios administradores. O arquivamento também pode ocorrer de forma aleatória, por meio de projetos específicos ou por solicitações individuais por meio da ferramenta “Salvar Página Agora”. Essas variáveis explicam a razão pela qual nem todas as URLs estão disponíveis em datas específicas no repositório, conforme justificado em resposta por *e-mail*:

Figura 9 – Cópia de mensagem eletrônica da *Internet Archive*



Fonte: captura de tela pela autora (2025).

Diante desse obstáculo inicial, buscamos, então, mudar o recorte temporal de coleta dos conteúdos noticiosos publicados nos webjornais, aproximando-o para um período mais recente, em 2024, na intenção de avançar na busca pelas informações que comporiam o *corpus*. Desse modo, evitamos lacunas no levantamento das publicações noticiosas a serem consideradas. Contudo, as oscilações no funcionamento da plataforma *Wayback Machine* persistiram. Após incontáveis tentativas e sem conseguir prosseguir de forma abrangente e satisfatória no recolhimento dos dados dos produtos comunicacionais do *Correio do Estado*, de Mato Grosso do Sul, e *O Popular*, de Goiás, optamos por restringir a estratificação, categorização e exame a dois webjornais, sendo *Gazeta Digital*, de Mato Grosso, e *Correio Braziliense*, do Distrito Federal, ambos relevantes por sua longa permanência no mercado e pela representatividade como produtos de conglomerados de mídias. A decisão por esse ajuste no objeto de pesquisa também ocorreu após amadurecimento de aportes teóricos durante as seis disciplinas cursadas, de forma presencial ou remota por videoconferência, durante o primeiro e segundo semestres de 2023 do Programa de Pós-Graduação em Comunicação e Poder (PPGCOM) da Universidade Federal de Mato Grosso (UFMT): Comunicação Ambiental, Sustentabilidade e Poder;

Metodologias de Pesquisa em Comunicação; Comunicação e Poder: Perspectivas Teóricas e Epistemológicas; Tópicos Especiais em Comunicação III; Seminário de Pesquisa; e a disciplina transversal Discurso Desinformativo e Plataformas no Brasil, ofertada pelo Instituto Nacional de Ciência e Tecnologia em Disputas e Soberanias Informacionais (INCT/DSI).

Conforme apresentado introdutoriamente neste estudo, para realizar as coletas durante 15 dias, optamos pela técnica do mês construído, obedecendo à ordem cronológica crescente, a partir do dia 1º de março até 21 de junho de 2023, abarcando pautas diferenciadas. Dessa forma, contemplando sequencialmente a primeira, segunda, terceira e quarta semanas de março – 01, 09, 17 e 25/03 –; a primeira, segunda, terceira e quarta semanas de abril – 02, 10, 18 e 26/04 –; a primeira, segunda, terceira e quarta semanas de maio – 04, 12, 20 e 28/05 –; e, ainda, a primeira, segunda e terceira semanas de junho – 05, 13 e 21/06 – do ano de 2023.

As coletas foram realizadas nas datas conforme quadro que se segue.

Quadro 6 - Datas e dias de coletas

01/03/2023	Quarta-feira
09/03/2023	Quinta-feira
17/03/2023	Sexta-feira
25/03/2023	Sábado
02/04/2023	Domingo
10/04/2023	Segunda-feira
18/04/2023	Terça-feira
26/04/2023	Quarta-feira
04/05/2023	Quinta-feira
12/05/2023	Sexta-feira
20/05/2023	Sábado
28/05/2023	Domingo
05/06/2023	Segunda-feira
13/06/2023	Terça-feira
21/06/2023	Quarta-feira

Fonte: elaborado pela autora (2025).

Justificamos a escolha desse recorte temporal diante da expectativa de identificarmos uma cobertura jornalística ancorada em dados, uma vez que em 2023 inicia o período pós-pandêmico²⁸, quando foi oficializado o fim da emergência global decorrente da covid-19. Durante a pandemia, iniciada em 2020, houve uma aceleração da digitalização, culminando em 2023 num alargamento do uso da Inteligência Artificial (IA) generativa e dos debates sobre a necessidade de regulação das plataformas digitais. Além disso, o ano de 2023 marca a transição de um governo de extrema direita para o retorno da esquerda ao poder político no Brasil, após eleições presidenciais em 2022.

Dentro do espaço temporal estabelecido, foram publicados na capa dos dois webjornais que compõem o *corpus* 1.959 conteúdos informacionais. Desse total, consideramos 54 notícias²⁹ e reportagens³⁰ (Melo; Assis, 2016) para análise, por apresentarem aparente utilização de dados que foram extraídos de repositórios digitais de dados abertos ou vazamentos de dados, com inclusão e apresentação de dados quantitativos, numéricos ou não, buscando, com isso, identificar características do JDBD. Foram dispensados da coleta outros textos e gêneros jornalísticos, como *releases*, artigos e outras produções que não eram o foco da nossa pesquisa, após leitura e observação pormenorizada de cada conteúdo informacional, o que permitiu reconhecer aqueles que não se enquadravam nos objetivos do estudo.

Para análise do *corpus*, aplicada a 54 notícias e reportagens do total de 1.959 conteúdos informacionais coletados, após visualização e leitura, das capas dos webjornais *Gazeta Digital* e *Correio Braziliense* e no período demarcado, criamos 29 categorias, divididas em 18 unidades de registro e 11 unidades de análise, com auxílio de um editor de **planilhas eletrônicas**³¹. A escolha pelas capas foi motivada por funcionarem como a principal porta de entrada para o leitor e, por isso, anteciparem a forma como ele será sensibilizado ou direcionado em relação ao conteúdo apresentado (Portari, 2009).

Do *Gazeta Digital* foram analisadas 18 notícias e reportagens de um total de 595 coletadas, enquanto do *Correio Braziliense* foram analisadas 36 publicações noticiosas de um total de 1.364.

²⁸ Em 5 de maio de 2023, após 3 anos e 3 meses desde a adoção da emergência global em janeiro de 2020, a Organização Mundial da Saúde (OMS) comunicou o “fim da pandemia”.

²⁹ Gênero jornalístico informativo de caráter factual.

³⁰ Gênero jornalístico informativo aprofundado em temas de interesse público.

³¹ Disponibilizado em: <https://tinyurl.com/5f4nsvxh>.

5 ANÁLISE DO USO DE DADOS EM NOTÍCIAS E REPORTAGENS DE *GAZETA DIGITAL E CORREIO BRAZILIENSE*

Mobilidade e velocidade caracterizam um novo ambiente para o exercício do jornalismo, decorrente da chegada dos computadores às redações e do advento da internet. A conexão em rede das máquinas possibilitou acesso e processamento de um grande volume de informações sem a necessidade de deslocamento físico, acelerando a produção de notícias e o alcance do público. Barbosa (2005; 2007; 2008; 2013) complementa que as bases de dados propiciaram o surgimento, a construção e a gestão de produtos jornalísticos digitais, desempenhando função estruturante, auxiliando e condicionando processos e práticas nas diferentes gerações do jornalismo digital, desde a primeira até a etapa seguinte de desenvolvimento (Barbosa, 2005; 2007; 2008; 2013; Barbosa; Torres, 2013; Pinto; Barbosa, 2024).

A incorporação de computadores influenciou as rotinas produtivas em todas as áreas. Na pesquisa, a investigação e a análise de conteúdo foram afetadas a partir dos anos 1960 e início de 1970, período que marca a segunda geração de computadores e sucede a chegada do primeiro “cérebro eletrônico”, em 1944, conforme contextualiza Bardin (2011, p. 28). Nas décadas seguintes, a evolução informática se acentua cada vez mais, permitindo assimilar quantidade de dados impossíveis de se manipular manualmente e testes estatísticos antes impraticáveis. Segundo Bardin:

Após meados dos anos 1970, a proliferação dos computadores pessoais e as experiências em inteligência artificial aumentam as esperanças nas possibilidades informáticas. Por influência de uma linguística mais aberta, e graças também às investigações levadas a cabo por certo tipo de análise do discurso, a experimentação informática, depois de ter ultrapassado os obstáculos anteriores de programação e de ter relativamente dominado as relações lexicais, concentra-se nas apresentações das estruturas sintáticas dos textos (Bardin, 2011, p. 31).

Segundo os preceitos bardinianos, a Análise de Conteúdo (AC) tem duas funções, que podem ou não se dissociar uma da outra. Uma é a função heurística, exploratória, que aumenta a propensão para descoberta. A outra é a função de administração da prova, quando a análise sistemática é aplicada para verificar hipóteses ou afirmações provisórias. A análise de conteúdo é, portanto, uma tarefa de interpretação que oscila entre o rigor da objetividade e a fecundidade da subjetividade. O analista, nesse processo, se vê atraído pelo latente, pelo não dito e pelos sentidos subentendidos nos textos. O método, nesse sentido, valida essa curiosidade

investigativa ao mesmo tempo em que impõe uma postura sistemática e científica, tornando essa leitura profunda legitimada pela metodologia. Essa técnica se diferencia de outras práticas como a análise documental, que se ocupa com documentos, enquanto a análise de conteúdo trabalha com mensagens. O objetivo da análise de conteúdo é, portanto, ir além da mera descrição e evidenciar indicadores que permitam inferir realidades não imediatamente observáveis na superfície textual.

Para cumprir as funções elencadas, a AC bardiniana se divide em três fases: a pré-análise, a exploração do material e o tratamento dos resultados com inferência e interpretação.

Na pré-análise, o pesquisador organiza o *corpus* documental, define os objetivos e hipóteses da investigação e constrói os indicadores que fundamentarão a leitura analítica. Na fase de exploração do material, realiza-se a codificação, que é o processo de recorte, enumeração e classificação do texto em unidades significativas, como, por exemplo, frases, palavras ou expressões, que serão agrupadas por categorias de análise. Por fim, no tratamento dos resultados, os dados brutos transformados em unidades codificadas são interpretados à luz dos objetivos do estudo, com auxílio de procedimentos estatísticos, quantitativos ou qualitativos, que fortalecem a validade e a produtividade das inferências.

Neste estudo, as unidades significativas foram definidas e aplicadas às notícias e reportagens que compõem o corpus conforme característica conceitual JDBD, notadamente dados numéricos, bem como busca proativa e visualização de dados. A categorização é um processo tanto lógico quanto criativo, podendo ser pré-definida com base na teoria, numa abordagem dedutiva, ou emergir progressivamente da leitura do material, de forma indutiva, revelando estruturas internas não evidentes à primeira vista.

A complementaridade entre o quantitativo e o qualitativo permite uma leitura mais robusta do material, especialmente quando trabalhamos com mensagens jornalísticas, que combinam padrões de linguagem, estruturas narrativas e estratégias discursivas. O uso de técnicas de classificação, agrupamento e cruzamento de variáveis, como origem, implicação, sentimentos e descrição, permite gerar representações estruturadas do conteúdo, as quais servem como base para análises interpretativas mais amplas.

Reforçamos que o processo de codificação é basilar na metodologia de Bardin (2011), possibilitando transformar os dados brutos em unidades organizadas, capazes de revelar regularidades e também rupturas, tensões e silêncios. A construção de categorias e subcategorias pode ser orientada por um índice ou dicionário de conceitos-chave, ou seja, um livro de códigos, que atua como ponte entre a teoria do pesquisador e os dados textuais. Essa sistematização tem caráter estruturalista. Classificar é, antes de tudo, procurar o que há de

comum entre elementos distintos. Na pesquisa, a categorização deve ser precisa, teórica e validada. Os conceitos-chave operam como variáveis construídas, que permitem o estabelecimento de inferências sobre a realidade representada pelo texto.

Os resultados da análise não constituem uma prova irrefutável, mas podem corroborar hipóteses, revelar sentidos ocultos e abrir espaço para descobertas inesperadas. A AC, quando bem estruturada, é, portanto, uma ferramenta para compreender como os discursos se organizam, o que dizem e, sobretudo, o que deixam de dizer.

Sampaio e Lycarião (2021), por sua vez, contribuem com a criação de um manual ou guia para análise de conteúdo categorial quantitativa, buscando explicar a aplicação de tal técnica de pesquisa a partir da definição e diferenciação entre códigos, codificação e categorias, elementos usualmente tomados como sinônimos na literatura:

As análises estatísticas permitem estabelecer a análise quantitativa de eventos e as características de uma população. Isso pode se dar de maneira mais elementar por intermédio da estatística descritiva, o que compreende técnicas e maneiras mais usuais, como frequências e gráficos (de pizza, barras, por exemplo.), medidas de tendência central (média, mediana, por exemplo) e dispersão (desvio padrão, por exemplo). Porém, técnicas estatísticas mais complexas, geralmente associadas à estatística inferencial, também são recursos úteis para algumas pesquisas (Sampaio; Lycarião, 2021, p. 113).

A AC se configura como uma técnica de pesquisa científica voltada para a investigação sistemática, objetiva e metodologicamente rigorosa das comunicações. Seu objetivo central é gerar inferências válidas acerca das mensagens veiculadas em diferentes suportes, sejam verbais, visuais ou escritos, considerando seus significados, intenções, efeitos e contextos de produção e recepção.

Embora existam variadas definições da técnica, os autores (Sampaio e Lycarião, 2021) se apoiam em uma tradição que remonta a Bardin (2011), cuja formulação clássica descreve a AC como um conjunto de técnicas sistemáticas e objetivas de detalhamento do conteúdo das mensagens, com vistas à produção de indicadores que permitam inferir aspectos das condições de produção, circulação ou recepção desses conteúdos. Sua natureza é híbrida, situada entre abordagens quantitativas e qualitativas, mas invariavelmente orientada por princípios da ciência.

Nesse sentido, Sampaio e Lycarião (2021) ressaltam que qualquer aplicação da AC deve ser norteadas por três princípios epistemológicos fundamentais, sendo eles a validade, a confiabilidade e a replicabilidade. Esses critérios asseguram o rigor da investigação e distinguem a AC como técnica científica e não como mero exercício interpretativo

descomprometido com regras analíticas explícitas. A validade refere-se à correspondência entre os dados analisados e os fenômenos que se pretende compreender. A confiabilidade diz respeito à consistência dos procedimentos de codificação e interpretação. A replicabilidade, por sua vez, implica na possibilidade de outro pesquisador, com acesso aos mesmos dados e regras analíticas, chegar a conclusões semelhantes.

Mais do que uma ferramenta de codificação ou quantificação de mensagens, a AC é um método de leitura controlada, fundamentada e teoricamente orientada, que permite ao pesquisador passar do conteúdo manifesto, aquilo que é explicitamente dito ou mostrado, ao conteúdo latente, ou seja, aos sentidos subjacentes, valores implícitos, estratégias discursivas ou intencionalidades comunicacionais. Nesse percurso, a AC busca reduzir a complexidade inicial do material empírico por meio da categorização sistemática, que organiza os dados em unidades de análise coerentes e analiticamente significativas. A definição adotada por Sampaio e Lycarião (2021) é, assim, uma síntese crítica e atualizada das contribuições clássicas e contemporâneas, formulada nos seguintes termos: Análise de Conteúdo é uma técnica de pesquisa científica baseada em procedimentos sistemáticos, intersubjetivamente validados e públicos para criar inferências válidas sobre determinados conteúdos verbais, visuais ou escritos, buscando descrever, quantificar ou interpretar certo fenômeno em termos de seus significados, intenções, consequências ou contextos.

Justificamos a adoção desses referenciais teóricos nesta pesquisa por sua abrangência metodológica e por sua capacidade de sustentar diferentes níveis de análise, do descritivo ao interpretativo, assegurando ao mesmo tempo transparência analítica e robustez epistemológica. Ao empregarmos a AC como eixo metodológico para investigar reportagens digitais com uso de dados, pretendemos compreender como determinados elementos simbólicos e informacionais são selecionados, organizados e apresentados nas narrativas jornalísticas.

Reforçamos, ainda, que esta pesquisa tem enfoque quali-quantitativo, fazendo uso também de elementos da abordagem categorial quantitativa, conforme orientações de Sampaio e Lycarião (2021) e Bardin (2011), que ressaltam que muitos dos passos metodológicos descritos para a AC quantitativa são igualmente úteis a análises qualitativas rigorosas. Isso inclui, por exemplo, a definição clara de categorias, a codificação sistemática das unidades de registro e, quando pertinente, a quantificação da frequência de ocorrência de determinados elementos no *corpus* analisado.

Neste capítulo, portanto, apresentamos os resultados da aplicação dos fundamentos metodológicos da técnica de AC para investigar 54 reportagens digitais com aparente uso de dados, publicadas em ambientes de webjornalismo. Nesse sentido, o uso da AC nesta

dissertação tem como finalidade identificar, de forma estruturada e sistemática, como os dados são incorporados nas reportagens digitais, em termos de sua presença narrativa, estrutura informacional e função jornalística. Seguindo os princípios defendidos por Sampaio e Lycarião (2021) e Bardin (2011), o processo de análise foi planejado para garantir transparência, coerência analítica e possibilidade de replicação, o que implica na exposição clara das etapas de categorização e nos critérios utilizados para o recorte e interpretação do corpus.

A escolha pela AC fundamentada em Sampaio e Lycarião (2021) também se sustenta em sua contribuição recente e relevante para o campo metodológico brasileiro, especialmente no contexto da análise de mensagens midiáticas. Os autores destacam que, embora a técnica tenha suas origens em estudos clássicos como os de Harold Lasswell (1927) sobre propaganda política durante as guerras do século XX, sua aplicação tem sido progressivamente ampliada para uma gama diversa de áreas, incluindo comunicação, ciências sociais, ciência da informação e saúde pública. Essa capilaridade atesta sua robustez epistemológica e sua adaptabilidade a diferentes objetos de estudo, incluindo os produtos jornalísticos veiculados em meios digitais.

Além disso, Sampaio e Lycarião (2021) alertam para a tendência contemporânea de substituição ou eclipsamento das análises manuais pelas técnicas automatizadas de análise de *Big Data*. No entanto, defendem a continuidade e relevância da análise de conteúdo humana como base de validação, complementação e refinamento das abordagens computacionais:

Finalmente, no atual estágio da pesquisa e do emprego de Big Data, existem diversos questionamentos sobre o uso de Big Data e sobre a falsa impressão de que eles geram, necessariamente, resultados melhores que as pesquisas de baixo *n*. Alguns autores têm justamente debatido como o Big Data e os algoritmos de plataformas podem (e, em alguns casos, tendem a) distorcer o debate público (e.g. O'NEIL, 2016). Ao menos no atual momento da tecnologia e da ciência, é inócuo acreditar que toda a AC manual poderá ser substituída por formas computadorizadas de análise. Em verdade, acreditamos que, não importando a evolução de tais técnicas, sempre haverá espaço para análises mais minuciosas e humanas do conteúdo que geramos (Sampaio; Lycarião, 2021, p. 13).

Nesta pesquisa, assumimos o uso da técnica manual como escolha metodológica intencional, por se tratar de um *corpus* reduzido e de natureza complexa, que demanda atenção interpretativa e sensibilidade analítica que os métodos automatizados ainda não são capazes de substituir integralmente. Por fim, ao considerarmos o contexto do webjornalismo, no qual as narrativas se estruturam cada vez mais a partir de dados, interfaces interativas e dinâmicas de plataforma, a aplicação da AC tal como proposta por Sampaio e Lycarião (2021) e Bardin

(2011) oferece um caminho metodológico viável para entendermos criticamente os modos de produção, circulação e significação da informação jornalística digital.

Com base nos pressupostos de Bardin (2011) e Sampaio e Lycarião (2021) em articulação com os aportes teóricos de Barbosa (2005; 2007; 2008; 2013), Mielniczuk (2001; 2003; 2004), Bertocchi (2005; 2016), Mancini e Vasconcelos (2016), Ferrari (2010), Lévy (1998), cujos contributos convergem para compreensão da linguagem jornalística mediada por dados e da dinâmica informacional no ambiente digital, este capítulo retoma o percurso metodológico adotado e apresenta a análise de conteúdo de 54 notícias e reportagens publicadas em ambiente webjornalístico, selecionadas após visualização e leitura de 1.959 conteúdos informacionais, divididos em 595 publicações do *Gazeta Digital* e 1.364 do *Correio Braziliense*, por apresentarem o uso aparente de dados em sua construção narrativa e informativa.

Detalhamos, a seguir, os resultados da AC de 54 notícias e reportagens que compõem o *corpus*, a partir da aplicação dos critérios de seleção, procedimentos de categorização e fases operacionais da análise. Com isso, identificamos padrões, recorrências, ausências e estratégias narrativas relacionadas ao uso de dados no webjornalismo, contribuindo para refletir sobre os modos de produção e circulação da informação jornalística em meio digital. A partir desse referencial analítico, buscamos evidenciar como os dados são mobilizados nas notícias e reportagens.

Seguindo o modelo bardiniano, cumprimos três fases para análise de conteúdo nesta pesquisa: a pré-análise, a exploração do material e o tratamento dos resultados com inferência e interpretação. Na pré-análise, definimos os indicadores para leitura analítica orientados pelos objetivos e problema de pesquisa. Na fase de exploração do material, realizamos a codificação, com recorte, enumeração e classificação do texto em unidades significativas, sendo estas relacionadas, principalmente, ao conceito teórico de Barbosa (2005; 2007; 2008; 2013) sobre JDBD e presença de dados numéricos nos conteúdos jornalísticos, que foram agrupados por categorias de análise. Na sequência, observamos que as unidades codificadas com auxílio de procedimentos estatísticos quantitativos e qualitativos possibilitam interpretações conforme os objetivos do estudo, produzindo inferências e validando os resultados.

5.1 Identificação do uso de dados em notícias e reportagens do *Gazeta Digital* (MT)

Especificamente do *Gazeta Digital* reunimos 595 conteúdos informacionais, entre artigos de opinião, entrevistas, *newsletters*, *podcasts*, vídeos, enquetes, notícias e reportagens.

Deste total, escolhemos para análise, conforme critérios estabelecidos nesta pesquisa, 18 publicações, com características de notícias e reportagens produzidas pelos jornalistas do próprio webjornal.

Devido às falhas da plataforma *Internet Archive* por meio da ferramenta *Wayback Machine*, conforme justificamos anteriormente, não houve captura dos conteúdos publicados pelo webjornal *Gazeta Digital* nas datas de 10 e 18 de abril de 2023, bem como em 04, 12 e 20 de maio de 2023. Nas datas de 10 de abril e 12 de maio, embora a busca via *Wayback Machine* retornasse resultados de 57 e 45 conteúdos jornalísticos, respectivamente, não permitiu o carregamento individual de cada notícia, impossibilitando, portanto, análises dos conteúdos nas referidas datas, demonstradas a seguir:

Quadro 7 – Períodos sem capturas pelo *Wayback Machine*

10/04/2023	Foram identificadas 57 notícias na página inicial, no entanto, os conteúdos individuais não carregaram corretamente, impossibilitando sua captura e análise.
18/04/2023	A página não foi exibida corretamente, o que impossibilitou a captura dos conteúdos jornalísticos disponíveis.
04/05/2023	A página não foi exibida corretamente, o que impossibilitou a captura dos conteúdos jornalísticos disponíveis.
12/05/2023	Foram identificadas 45 notícias na página inicial, no entanto, os conteúdos individuais não carregaram corretamente, impossibilitando sua captura e análise.
20/05/2023	A página não foi exibida corretamente, o que impossibilitou a captura dos conteúdos jornalísticos disponíveis.

Fonte: elaborado pela autora (2025).

Analisando cada uma das 18 publicações do webjornal *Gazeta Digital* que atendem aos preceitos teóricos estabelecidos nesta pesquisa e que compõem o *corpus* de 54 notícias e reportagens coletadas em dois webjornais, incluindo o *Correio Braziliense*, identificamos, conforme categorizações estabelecidas em livro de códigos orientado por Bardin (2011), Sampaio e Lycarião (2021) e conceitos teóricos de JDBD (Barbosa, 2005; 2007; 2008; 2013), obtivemos os resultados que descreveremos a seguir.

A notícia intitulada “**Estado anuncia realização de concurso na Saúde este ano**” (<https://tinyurl.com/2psudet4>), publicada no dia 1º de março de 2023 no webjornal *Gazeta Digital*, não menciona números ou estatísticas logo no título (UA01), porém, deixa pressuposta essa informação na matéria, conforme explicitada no texto, que menciona 400 vagas, de forma numérica clara e estruturada, além de valores e prazos, como R\$ 800 milhões e 50 meses de repasses (UA02). Em relação à busca de dados, constatamos que ela não foi realizada, sendo a

matéria baseada em declarações oficiais, não evidenciando investigação ativa em bases públicas (UA03). Os dados foram apenas citados de outras fontes, tendo sido reproduzidos a partir de falas do governador e do secretário de saúde (UA04), sem cruzamento, crítica ou interpretação das declarações (UA05). Não há, contudo, indício de manipulação ou distorção dos dados apresentados (UA06). A origem dos dados é atribuída aos gestores públicos (UA07). Não há menção de que o veículo requisitou dados ou obteve resposta por meios oficiais (UA08). Nenhuma referência à LAI foi feita (UA09), bem como a matéria não aborda temas relacionados à proteção de dados pessoais, previstos na LGPD (UA10). A notícia não apresenta qualquer representação visual de dados (UA11).

Na sequência, em notícia publicada no dia 09 de março de 2023, com o título **“Operação Dissidência II - Força-tarefa prende 12 pessoas em operação contra facções criminosas em Cuiabá e interior de MT”** (<https://tinyurl.com/mr327hcu>), verificamos que o título menciona explicitamente o número de 12 pessoas presas, o que configura um dado quantitativo (UA01). O corpo do texto menciona mandados, número de prisões, datas e locais, além de citar números anteriores, como 36 réus e 70 mandados em 2022 (UA02). A matéria não evidencia investigação própria do jornalista, apenas reproduzindo dados fornecidos por autoridades, provenientes de informes da Polícia Federal e órgãos oficiais, sendo replicados no texto (UA03 e UA04). O conteúdo se limita ao factual, de forma que, para compreensão do contexto – facções, impacto local, processo judicial, entre outros –, seria necessária maior investigação jornalística (UA05). Não há indícios de dados manipulados ou informações falsas (UA06). Os dados são atribuídos a fontes oficiais, como a Polícia Federal (UA07). Não há evidência de que o jornalista tenha solicitado dados via canais oficiais (UA08). A matéria não menciona nem faz referência ao uso da LAI (UA09) e não trata de aspectos relacionados à proteção de dados pessoais prevista na LGPD (UA10). Não há visualizações, tabelas ou gráficos para facilitar a interpretação dos dados mencionados (UA11).

Na análise da notícia **“No Acre – Acusado de matar 3 em Poconé é preso 8 anos após crimes”** (<https://tinyurl.com/muue6n6w>), publicada no dia 09 de março de 2023, verificamos que o título informa que o indivíduo estava sendo acusado de “matar 3” pessoas e que a prisão ocorreu “8 anos após” o crime, portanto, houve uso de números (UA01). O texto apresenta datas (08/03/2015 e 08/03/2023), nomes das vítimas, localidade, idade do acusado (36 anos), entre outros dados numéricos e factuais (UA02). Não há evidência de que o jornalista tenha feito busca em bancos de dados (UA03). As informações estatísticas foram fornecidas por autoridades e não sofreram análise ou tratamento adicional pela reportagem (UA04) e o relato dos fatos foi feito conforme a versão da polícia, sem cruzamento, contextualização criminal ou

sociológica (UA05). Os dados apresentados são coerentes e atribuídos a fonte oficial, sem indícios de manipulação (UA06). Ademais, a Polícia Civil de Mato Grosso e do Acre são claramente indicadas como fontes primárias (UA07). A matéria não menciona nem dá sinais de uso de mecanismos formais de acesso à informação, como o Sistema Eletrônico do Serviço de Informações ao Cidadão (e-SIC) (UA08). Não há referência à LAI (UA09) e a LGPD também não é citada, embora o texto envolva dados pessoais (UA10). Não há gráficos, mapas, linha do tempo ou qualquer visualização de dados (UA11).

No dia 17 de março de 2023, selecionamos para análise a notícia **“MPF acompanha repactuação do contrato da BR-163/364 e cobra construção de viaduto”** (<https://tinyurl.com/mrxkbbyz>). Logo no título, observamos menção a trechos rodoviários específicos (BR-163/364), além de ter sido citado o objeto da ação, como a repactuação e construção de viaduto (UA01). O corpo da matéria traz informações numéricas, como localização (264 km ao Norte), procedimentos do Ministério Público Federal (MPF), termos técnicos e citações de documentos (UA02). A matéria se baseia em uma portaria oficial publicada no *Diário Eletrônico* do MPF, o que indica uma busca deliberada por documentos públicos (UA03). Apesar da busca documental, os dados são reproduzidos diretamente dos autos do MPF, sem análise crítica ou cruzamento com outras fontes (UA04). A matéria aborda questões técnicas e jurídicas, como repactuação, desequilíbrio econômico-financeiro, caducidade do contrato, exigindo interpretação ou conhecimento prévio para compreensão plena (UA05). Não há sinais de distorção, omissão ou uso incorreto de dados (UA06). O texto menciona claramente como fonte o *Diário Eletrônico* do MPF e a procuradora responsável (UA07). Não há evidência de uso de mecanismos formais para acesso às informações, uma vez que a matéria se apoia em documento já público (UA08). A LAI não é mencionada (UA09), nem a LGPD (UA10). A notícia é puramente textual, sem qualquer elemento visual ou gráfico (UA11).

Em outra notícia publicada no dia 17 de março de 2023, intitulada **“Vereador acusa petista de ‘nepotismo disfarçado’ por tentar empregar marido na Câmara”** (<https://tinyurl.com/3crxder6>), percebemos que, embora haja menção ao termo “salário de R\$ 32 mil” em destaque, o dado não aparece no título principal, o qual é integralmente textual (UA01). O corpo da matéria traz valores salariais (R\$ 32 mil), período da cessão (2 anos) e impacto estimado no erário (R\$ 840 mil) (UA02). A apuração é baseada exclusivamente em declarações públicas do vereador, sem investigação ou consulta a bases públicas de folha de pagamento ou contratos de cessão (UA03). Os dados e declarações foram reproduzidos da fala do vereador Dilemário Alencar na tribuna, sem verificação ou análise adicional (UA04). A

notícia aborda temas como nepotismo, impacto orçamentário e cessão de servidores públicos, mas sem aprofundamento por meio de análise legal, orçamentária e ética (UA05). Não há distorção visível dos dados apresentados, mas também não há checagem das informações, limitando a confiabilidade jornalística (UA06). A fonte é explicitamente identificada, sendo o vereador Dilemário Alencar (Podemos), com fala transcrita diretamente (UA07). O veículo não requisitou dados formais, como folha salarial, contratos de cessão ou posicionamento oficial da Câmara ou da vereadora petista (UA08). Não há qualquer menção ou referência ao uso da LAI (UA09). Embora trate de remuneração de servidor, não há debate sobre exposição de dados pessoais ou menção à LGPD (UA10). A matéria não apresenta elementos visuais como infográficos, tabelas ou gráficos que ajudem a visualizar o impacto financeiro apontado (UA11).

A notícia **“Sortudo: Morador de Sinop garante R\$ 41 mil em sorteio da Mega”** (<https://tinyurl.com/84p86j89>) traz título com valor numérico (R\$ 41 mil) (UA01). O texto, publicado no dia 17 de março de 2023, apresenta múltiplos dados numéricos, como o valor do prêmio de R\$ 41.754,98, concurso de número 2574, os números sorteados, 12, 17, 43, 44, 48, 60, a quantidade de 68 ganhadores, o valor estimado do próximo sorteio, de R\$ 45 milhões, bem como a distância geográfica da cidade em relação à capital, de 500 km. Esses elementos indicam presença efetiva de dados, embora não sejam estruturados ou explorados além da informação descritiva (UA 02). A reportagem reproduz informações que são disponibilizadas publicamente pela Caixa Econômica Federal, não havendo indícios de investigações ou coleta ativa de dados além do boletim oficial da Mega-Sena (UA 03). Os dados foram reproduzidos diretamente do sorteio da Mega-Sena, sem tratamento adicional, cruzamento com outras informações ou contextualização estatística (UA04). A matéria é factual e informativa, mas não propõe análise ou investigação sobre a loteria, nem explora padrões, frequência de prêmios por região ou perfis de ganhadores (UA05). Não há indícios de manipulação de dados ou descontextualização (UA06). A Mega-Sena, vinculada à Caixa Econômica Federal, é citada como origem dos dados (UA07), de forma que o conteúdo jornalístico não envolve requisição de dados por canais formais, como a LAI, baseando-se exclusivamente em fontes já disponíveis ao público (UA08). Não há qualquer menção ou uso da LAI no processo de apuração (UA09). Embora mencione o local do ganhador, nenhum dado pessoal é exposto e a LGPD não é citada (UA10). A matéria não apresenta visualização de dados, como tabelas, mapas ou gráficos (UA11).

No dia 25 de março de 2023, o webjornal *Gazeta Digital* publicou notícia intitulada **“Guia para Homens – Maioria acha ‘bobagem’ treinamento de coaches de**

relacionamentos” (<https://tinyurl.com/37ktdbxa>). Conforme as unidades de análise estabelecidas em livro de códigos, notamos que, embora trate de resultados de enquete, o título não explicita números ou percentuais, apenas o termo genérico “maioria” (UA01). O corpo da matéria apresenta números absolutos e percentuais dos votos de uma enquete, como os 365 votos e percentuais para cada resposta (UA02). Os dados vêm de uma enquete própria do portal, sem uso de bases externas ou dados públicos (UA03), sendo reproduzidos sem contextualização, metodologia ou crítica quanto à validade da amostragem (UA04). O tema envolve debates contemporâneos sobre masculinidade, comportamento afetivo e influência de “coaches”, exigindo maior análise sociocultural (UA05). Não há uso manipulativo ou distorcido dos dados, mas a fragilidade metodológica da enquete pode gerar impressões equivocadas sem que o texto sinalize isso (UA06). A matéria informa que os dados foram obtidos por meio de uma enquete publicada no próprio portal de notícias, embora sem *link* direto ou metodologia (UA07). A apuração se restringe à enquete interna e não envolve pedido de informações públicas ou base de dados externas (UA08). A LAI não é mencionada, nem implicitamente relacionada ao conteúdo (UA09); do mesmo modo, a LGPD também não é mencionada, embora a enquete envolva coleta de opinião de usuários (UA10). Os resultados da enquete são apresentados apenas em texto corrido, sem visualização gráfica (UA11).

Em análise de conteúdo da notícia “**Empresas suspeitas de cartel na Saúde mantêm contrato de R\$ 100 mi**” (<https://tinyurl.com/4s5vw95n>), publicada no dia 02 de abril de 2023, com base em livro de códigos contendo classificação das unidades de análise fundamentada, observamos que o título destaca diretamente um dado quantitativo relevante (“contrato de R\$ 100 milhões”), com valor-notícia centrado no volume financeiro envolvido (UA01). A matéria traz valores contratuais, nomes de empresas, órgãos públicos envolvidos, datas de operações policiais e valores de bloqueios judiciais, informações que estruturam o texto com base em evidências numéricas e dados oficiais (UA02). A apuração aponta que a reportagem teve acesso a informações além de comunicados oficiais, como o nome de administrador (O.G.C.), citações a informações obtidas em *Diário Oficial do Estado* (DOE) e da defesa de um empresário de São Paulo que estaria negociando colaboração com o Ministério Público de Mato Grosso (MPMT), configurando busca ativa de dados (UA03). A matéria reproduz informações de fontes institucionais e oficiais (MPMT, Delegacia Especializada de Combate à Corrupção - Deccor, juiz da 7ª Vara). Esses dados são inseridos sem reinterpretação crítica, caracterizando reprodução direta, mesmo que devidamente contextualizada (UA04). Apesar de conter dados importantes, o texto não aprofunda análises nem apresenta cruzamentos de dados, séries históricas ou investigações paralelas sobre as empresas citadas. Há indícios de investigação

jornalística, mas sem extrapolar os dados brutos ou interpretá-los criticamente (UA 05). Não há sinais de dados falsos, distorcidos ou fora de contexto. As informações são atribuídas a fontes identificadas e alinhadas à cobertura factual (UA 06), bem como as fontes institucionais são citadas com clareza (juiz João Bosco Soares da Silva; Delegacia Especializada de Combate à Corrupção - Deccor, Ministério Público de Mato Grosso – MPMT) (UA07). Não há evidência de que o veículo tenha feito pedidos formais de acesso às informações (UA08). A LAI não é citada (UA09) e, apesar de citar nomes de empresas e pessoas físicas, não há discussão sobre o uso de dados pessoais sensíveis, nem menção à LGPD (UA10). A matéria não traz recursos visuais como gráficos, tabelas, mapas ou *timelines* para a exploração visual dos dados (UA11).

Figura 10 – Captura de tela de notícia publicada pelo *Gazeta Digital*



Empresas suspeitas de cartel na Saúde mantêm contrato de R\$ 100 mi

O juiz da 7ª Vara Criminal da Comarca de Cuiabá, João Bosco Soares da Silva, determinou que o governo Mauro Mendes (União) e os 141 prefeitos do Estado fiquem proibidos de realizar contratação com 9 empresas de saúde, que estão sendo investigada na Operação Espelho, da Delegacia Especializada de Combate à Corrupção (Deccor). Tais empresas possuem contratos de mais de R\$ 100 milhões com o Estado.

A determinação ocorreu durante a segunda fase da operação no fim de março, quando foi bloqueado R\$ 35 milhões em bens dos investigados.

No despacho encaminhado aos prefeitos, Bosco alega que as investigações e auditorias estão em andamento, e que por isso novos contratos e a suspensão do pagamento para estas empresas são necessárias.

As empresas são suspeitas de participação de um cartel desde o início da pandemia da covid-19 e atuaram com conseguir contratos com o Estado e municípios.

Algumas das empresas possuem os mesmos sócios e administradores, o que aumenta as suspeitas de cartel.

Além da empresa LB Serviços Médicos Ltda, também estão na mira dos investigadores: Intensive Care Serviços Médicos Ltda, Surgery MT, Serviços de Anestesiologia Anestec, Bone Medicina Especializada, Curat Serviços Médicos Especializados Ltda, Gonçalves Preza Serviços, Samir Yoshio Matsumoto Bissi Eireli-ME e Medtrauma Centro Especializado em Ortopedia e Traumatologia LTDA.

Marcos Vergueiro/Secom-MT



Fonte: Rodrigo (2023).

Embora a notícia intitulada **“Empresas suspeitas de cartel na Saúde mantêm contrato de R\$ 100 mi”** incorpore informações numéricas e dados oficiais, tais como valores contratuais, datas de operações, órgãos públicos envolvidos e quantias bloqueadas judicialmente, sua estrutura narrativa permanece ancorada no modelo tradicional do jornalismo factual. O destaque atribuído ao valor financeiro no título e a constante mobilização de cifras ao longo da matéria caracterizam o uso instrumental dos dados, sem que estes efetivamente constituam o eixo estruturante da narrativa. Tal configuração aproxima o material noticioso do que Mancini e Vasconcellos (2016) definem como jornalismo com dados, na qual informações quantitativas servem predominantemente como complemento descritivo e não como fundamento analítico orientador da pauta.

Apesar da matéria demonstrar algum grau de busca ativa de informações, como a consulta ao Diário Oficial do Estado e a inclusão do nome de administradores de empresas investigadas, essa iniciativa não se desenvolve sequer nos termos estabelecidos por Träsel (2014; 2017) para o Jornalismo Guiado por Dados (JGD). Segundo o autor, o JGD pressupõe práticas que extrapolam a mera obtenção de documentos ou declarações oficiais, envolvendo análise sistemática de bases de dados, uso de técnicas de Reportagem Assistida por Computador (RAC), cruzamento de informações e produção de visualizações. A notícia, ao contrário, limita-se a reunir dados disponíveis e reproduzi-los de modo direto, sem submetê-los a processos interpretativos que permitam revelar padrões, relações ou inconsistências subjacentes.

Nesse sentido, também se observa distanciamento em relação ao Jornalismo Digital em Base de Dados (JDBD), proposto por Barbosa (2005; 2007; 2008; 2013). A autora argumenta que tal modelo exige a organização, tratamento e reconfiguração dos dados em um “continuum multimídia” que sinaliza a emergência de uma quinta geração do jornalismo digital, no qual diferentes formatos como texto, infografia, mapas, gráficos, áudio e vídeo funcionam de maneira integrada para ampliar a compreensão do fenômeno investigado.

A ausência total de recursos visuais ou elementos multimídia na notícia analisada evidencia a não aderência ao paradigma JDBD. A falta de infográficos, gráficos ou mapas também contraria o entendimento de Casadei, Rodrigues e Biernath (2016), que reconhecem nesses dispositivos alguns dos principais meios expressivos do jornalismo de dados.

Outro aspecto que reforça a distância entre a prática observada e os preceitos do jornalismo de dados refere-se à ausência de menções à Lei de Acesso à Informação (LAI) e à inexistência de indícios de que a redação tenha acionado mecanismos formais para obtenção de dados governamentais. Considerando que a transparência (Brenol, 2019; Gehrke, 2021) apresenta-se como requisito indispensável para a manutenção da credibilidade das informações

veiculadas e constitui elemento estruturante do jornalismo orientado por dados, a falta dessa dimensão evidencia a limitação da apuração. Do mesmo modo, a inexistência de referências à Lei Geral de Proteção de Dados (LGPD), mesmo diante da citação de dados pessoais e empresariais, aponta para a ausência de reflexão sobre o uso ético e regulado de informações sensíveis, componente central na produção jornalística contemporânea.

Assim, embora a reportagem mobilize dados de forma ostensiva e apresente indicadores numéricos relevantes, não os transforma em objeto de análise nem os utiliza como referência basilar da investigação jornalística, permanecendo alinhada ao jornalismo declaratório e informativo. A narrativa resultante carece de aprofundamento interpretativo, de cruzamentos de dados, de construção de bases próprias e de recursos de visualização, elementos indispensáveis ao jornalismo de dados. A matéria se enquadra, portanto, em um jornalismo factual tradicional, que utiliza dados como elementos acessórios, mas não atinge o nível de complexidade metodológica e analítica esperado nas práticas contemporâneas de Jornalismo Digital em Base de Dados (JDBD)

Nos dias 10 de abril de 2023 e 18 de abril de 2023 não houve capturas de notícias para serem analisadas.

Análise de conteúdo da notícia **“Sema ignora lei que proíbe PCHs e analisa 6 projetos”**, publicada no dia 26 de abril de 2023 (<https://tinyurl.com/avzjdjder>), revela que a notícia se insere em um contexto de conflito entre desenvolvimento energético, legislação ambiental e governança pública, o que qualifica sua complexidade e necessidade de interpretação crítica. O título menciona “6 projetos”, número que evidencia o volume de empreendimentos em análise. Isso confere um caráter quantitativo ao título (UA 01). A reportagem apresenta diversos dados relevantes, como nome dos empreendimentos (Guapira II, Tratambé I e II, Angatu I e II, Perudá), quantidade de projetos (6), nome da empresa interessada (Maturati Participações S.A.), número de servidores da comissão (8), tempo de análise (1 ano), resultado da votação legislativa (20 votos a favor e 3 contrários), bem como referências legais (Lei Complementar nº 38/1995, artigos 23 e 24). Esses elementos estruturam o texto com base documental e institucional, permitindo análise técnica e jurídica (UA 02). Há busca ativa por informações contendo dados quando o jornalista menciona conteúdo de portarias publicadas no *Diário Oficial*, trechos específicos da legislação estadual, detalhes do Termo de Ajustamento de Conduta (TAC), além de nome e função da secretária adjunta e do coordenador da comissão. Essas informações indicam pesquisa ativa em documentos oficiais, o que caracteriza um trabalho de jornalismo investigativo com uso de dados públicos (UA 03). A matéria, contudo, reproduz trechos diretos de portarias e leis, sem reorganizar as informações em formatos mais

interpretativos (UA04). A reprodução é qualificada, mas ainda não transita para transformação dos dados em conhecimento novo, sendo importante, nesse sentido, aprofundamento, uma vez que o texto discute divergência entre leis estaduais e decisões judiciais, articulação política, como demonstrado com veto do governador, e derrubada pela ALMT, contexto de julgamento no Supremo Tribunal Federal (STF), além de posicionamentos institucionais diversos (governo, Ministério Público Estadual - MPE, Tribunal de Justiça de Mato Grosso - TJMT, Agência Nacional de Águas - ANA). Essa sobreposição de camadas exige leitura crítica, contextualização jurídica e institucional, indicando complexidade interpretativa. A reportagem vai além da mera informação factual (UA05). Não apresenta desinformação com dados, uma vez que estes são verificados, institucionalmente atribuídos e contextualizados com posicionamentos múltiplos. Não há sinal de distorção, manipulação ou omissão que configure desinformação (UA 06). As fontes dos dados são citadas, sendo predominantemente oficiais, destacando o *Diário Oficial do Estado*, além de outras instituições, como ALMT, MPMT, Secretaria Estadual de Meio Ambiente (Sema), STF, TJMT, Fundação Eliseu Alves e ANA (UA07). Não há indício de pedidos formais de informações que provenham de documentos ou notas institucionais ou fontes públicas (UA 08). A reportagem não cita a LAI (UA 09). Apesar de tratar de nomes e cargos de servidores, nenhum dado pessoal sensível é explorado ou exposto de forma inadequada e não há, portanto, menção à LGPD (UA 10). Verificamos ausência de visualização gráfica, como mapa da área afetada, linha do tempo do processo ou fluxograma de tramitação legal (UA11).

Nos dias 04, 12 e 20 de maio não houve captura dos conteúdos noticiosos, por falha operacional do sistema *Internet Archive* e *Wayback Machine*.

Análise de conteúdo da notícia “**Mais de 170 mil eleitores estão com título cancelado em MT**” (<https://tinyurl.com/3f4bvf7f>), publicada no dia 28 de maio de 2023, revela que o título apresenta dado numérico (UA01) e conteúdo geral da notícia com uso intensivo de dados numéricos, como total de eleitores, títulos cancelados, prazos e localizações (UA 02). Não há, contudo, indício de apuração própria de dados, apenas reprodução de informações oficiais (UA03). A notícia reproduz dados da Justiça Eleitoral sem tratamento crítico (UA04), sendo que o texto é informativo, embora raso, por não aprofundar causas para o cancelamento, implicações sociais ou perfil dos atingidos (UA05). Não há desinformação com dados, uma vez que estão coerentes e sem distorções (UA06) e as fontes dos dados foram devidamente citadas (UA07). Não há indício de que o jornalista tenha solicitado dados extras ou inéditos (UA08). A notícia não cita a LAI (UA09), nem discute proteção de dados no âmbito da LGPD, mesmo falando sobre cadastros e biometria (UA10). Não há gráfico, mapa ou outro recurso visual de

apoio aos dados (UA11). Barbosa (2005; 2007; 2008; 2013) destaca que o jornalismo de dados deve ir além da reprodução de números e buscar contextualização, interpretação e visualização informativa.

Em outra notícia, publicada no dia 28 de maio de 2023 sobre o *ranking* das cidades que mais “enriqueceram” e “empobreceram” em Mato Grosso, intitulada “**Mapa da Riqueza - Ranking FGV: as cidades que mais ‘enriqueceram’ e ‘empobreceram’ em MT**” (<https://tinyurl.com/2hm7vrt2>), verificamos ausência de dados quantitativos no título, havendo apenas termos subjetivos (UA01). O corpo da matéria, por sua vez, traz valores numéricos detalhados, como renda média, variações percentuais e *ranking* (UA02). A matéria utiliza uma pesquisa inédita da Fundação Getúlio Vargas (FGV), demonstrando apropriação de dados disponíveis (UA03), uma vez que a matéria reproduz diretamente dados do estudo “Mapa da Riqueza” (UA04). O texto apresenta dados numéricos, mas não explora causas estruturais, impactos sociais ou políticas públicas relacionadas à questão (UA05). Os dados são apresentados com clareza e sem distorção aparente (UA 06). A FGV é citada como responsável pelo estudo, incluindo a origem da base de dados, que é a Receita Federal, com informações sobre o Imposto de Renda Pessoa Física (IRPF) e a Pesquisa Nacional por Amostra de Domicílio Contínua (PnadC) (UA 07). O jornalista não solicitou dados específicos, ou seja, o conteúdo baseia-se em divulgação pública (UA 08). Não há menção à LAI (UA09), nem à LGPD (UA10), apesar de usar dados sensíveis como renda. Os dados estão apenas em formato de texto, em lista e comparações numéricas, sem infografia (UA11).

Também em 28 de maio de 2023, na notícia intitulada “**Justiça registra 22 crianças adotadas até meio deste ano**” (<https://tinyurl.com/54xrfr3p>) verificamos, com base em livro de códigos adotado para análise de conteúdo neste estudo, que o título traz um dado quantitativo claro – 22 crianças adotadas até maio (UA01). A matéria apresenta diversos números específicos, como, por exemplo, 80 crianças em casas lares, 21 disponíveis para adoção (UA02). Os dados são reproduzidos de fontes oficiais e entrevistas, sem investigação própria (UA03), obtidos diretamente do Sistema Nacional de Adoção e da fala da juíza (UA04). Embora contenha informações importantes, faltam análises sobre políticas públicas, desafios legais ou estatísticas comparativas (UA05). Os dados são usados com clareza e fidelidade, não havendo distorções (UA06). As fontes são citadas, sendo elas o Sistema Nacional de Adoção e Acolhimento do Conselho Nacional de Justiça (SNA/CNJ) e a 1ª Vara da Infância e Juventude de Cuiabá (UA07). Não há menção de requerimento direto a bases de dados fechadas (UA08). Não há menção à LAI (UA09), nem à LGPD, mesmo tratando de informações sensíveis como HIV e deficiência (UA10). Não há uso de infográficos ou elementos visuais, sendo os dados

apresentados somente em texto corrido (UA11).

Avaliando o grau de presença e uso de dados, a notícia **“Energisa é condenada a pagar R\$ 16 mil a clientes após cobrar R\$ 10 mil de consumo recuperado”** (<https://tinyurl.com/5aw48kzt>), publicada no dia 28 de maio de 2023, revela presença de dados numéricos no título (UA01). O corpo da notícia traz o uso de dados ao informar valores extras das cobranças (R\$ 8.071,10 e R\$ 2.802,67) e das indenizações (R\$ 8 mil para cada cliente), além de mencionar os anos das cobranças (2017 e 2022) (UA02). Não há, porém, evidência de apuração própria ou investigação em busca de dados por parte do jornalista, que apresenta apenas aqueles derivados de decisão judicial, configurando simples reprodução de dados contidos nos autos judiciais, como decisões, trechos da sentença e valores determinados (UA04). Embora traga informações objetivas, a notícia poderia ser mais aprofundada ao contextualizar o que é “consumo recuperado”, como é feito o processo de medição e quais os procedimentos legais que a concessionária deve seguir nesses casos. Faltam entrevistas com especialistas, consumidores ou com a empresa, o que limitaria a pluralidade de fontes de dados (UA05). Não há desinformação na matéria. As informações são corretas e baseadas em fontes oficiais, como decisões da Justiça. A ausência de contextualização técnica, no entanto, pode gerar dúvidas no leitor, embora não configure desinformação (UA06). A fonte principal das informações é a 3ª Vara Cível de Cuiabá, e o nome da juíza responsável, Edna Ederli Coutinho, é citado. Há, portanto, identificação clara da fonte das informações e dados citados na matéria (UA07). Não há indícios de que o jornalista tenha solicitado dados, documentos, notas de esclarecimento da Energisa ou posicionamentos de outras partes e a cobertura se limita à decisão judicial já publicada (UA08). A LAI não é mencionada, nem há indício de que tenha sido utilizada como instrumento de apuração (UA09). A LGDP também não é mencionada, ainda que o tema envolva dados de consumo e clientes, o que poderia abrir espaço para discussões sobre uso e proteção dessas informações (UA10). Não há uso de recursos gráficos ou visuais, como quadros comparativos ou infográficos com os valores das cobranças e indenizações (UA11).

No dia 05 de junho de 2023, foi publicada pelo *Gazeta Digital* notícia intitulada **“Juíza condena Wilson e empresa a ressarcirem R\$ 154 mil aos cofres públicos por contrato sem licitação”** (<https://tinyurl.com/2ckjub22>). A análise do conteúdo revela que o título inclui dado quantitativo (R\$ 154 mil) (UA01). O corpo da matéria também apresenta dados numéricos, como o valor do contrato (R\$ 154.083,00), os períodos mencionados (1996 a 1999, 2006 a 2007, 2012 e 2013), bem como as partes envolvidas, sendo elas o deputado Wilson Santos, o Instituto de Tecnologia, a Prefeitura de Cuiabá e a decisão da magistrada (UA03). Não há

indícios de consulta a banco de dados ou documentos além da decisão judicial. O conteúdo é baseado exclusivamente na sentença, sem acréscimos investigativos ou entrevistas, caracterizando dependência de fonte oficial (UA04). Apesar da apresentação clara dos dados da decisão, o texto não explica tecnicamente como funciona a compensação de créditos de Pasep, não havendo apuração sobre o histórico de contratações similares, além da ausência de entrevista com as partes envolvidas ou especialistas. A notícia carece de aprofundamento explicativo e pluralidade de fontes dos dados (UA05). Não há sinais de desinformação. A matéria apresenta os dados de forma coerente com os autos judiciais (UA06). A Vara Especializada em Ações Coletivas é explicitamente citada como fonte da decisão, bem como o nome da juíza responsável, Célia Regina Vidotti (UA07). A reportagem não demonstra ter feito pedidos de esclarecimento à prefeitura, ao ex-prefeito ou à empresa envolvida (UA08). Não há menção à LAI nem indicativo de que ela tenha sido usada como mecanismo para obtenção de dados, documentos ou verificação de fatos (UA09). Apesar de tratar de recursos públicos e contratos administrativos, a matéria não menciona a LGPD, nem discute como os dados envolvidos na gestão pública são tratados ou divulgados (UA10). A notícia não conta com recursos visuais, como quadros, cronogramas ou esquemas explicativos, que poderiam ajudar a sintetizar os dados históricos e jurídicos complexos (UA11).

Uma análise de conteúdo da notícia **“INSS autoriza nomeação de candidatos aprovados em concurso; 22 de MT”** (<https://tinyurl.com/5n9bndcc>), publicada no dia 13 de junho de 2023, conforme retorno à busca no sistema *Wayback Machine* da *Internet Archive* na data de 12 de junho de 2023, revela que o título traz como dado numérico o salário inicial de R\$ 5,9 mil (UA01). A matéria apresenta informações quantitativas detalhadas, como número total de nomeações (1.000), quantidade de vagas por estado, valor da remuneração (R\$ 5.905,79), jornada de trabalho (40h semanais) e citação de norma oficial (Portaria 2.719/2023) (UA02). O conteúdo é baseado exclusivamente em fontes oficiais, sendo elas o Instituto Nacional do Seguro Social (INSS), Ministério da Previdência e a portaria mencionada (UA03). A notícia reproduz informações já disponibilizadas publicamente, especialmente da portaria e das redes sociais do ministro (UA04). A matéria não traz contextualização sobre déficit de servidores do INSS, impacto da nomeação no atendimento ao público e situação da fila de espera da Previdência. Faltam, portanto, elementos analíticos e explicativos que caracterizam aprofundamento informativo (UA05). A notícia não apresenta distorções ou erros nos dados. Os números, as fontes e as informações estão coerentes com os fatos oficiais, afastando qualquer tipo de desinformação (UA06). A fonte principal é o INSS, além da ministra da Gestão e do ministro da Previdência, todos identificados (UA07). Não há sinal de que o veículo tenha

solicitado informações adicionais por meio de canais oficiais ou assessorias. Isso demonstra passividade na obtenção de dados, reforçando o caráter reprodutivo do conteúdo (UA08). A LAI não é mencionada (UA09) e, apesar de o tema envolver nomes de aprovados e informações sensíveis sobre concursos públicos, não há menção à LGPD (UA10). A matéria não traz gráficos, quadros ou mapas com a distribuição das vagas por estado, o que poderia facilitar a leitura e ampliar o engajamento do público (UA11).

Na notícia **“Juíza nega bloqueio de R\$ 3,4 mi das contas de ex-secretário por esquema na Sinfra”** (<https://tinyurl.com/4s2p8vd7>), publicada no dia 12 de junho de 2023, observa-se inserção de dado quantitativo (R\$ 3,4 milhões) logo no título (UA01). A matéria também inclui informações numéricas, como prejuízo estimado de R\$ 3.445.175,36, datas, bem como nomes dos envolvidos. Há identificação de rodovias e trechos afetados, referência a contratos, delações premiadas e decisões judiciais (UA02). A reportagem acessa documentos oficiais, como decisões judiciais e termos de delações premiadas, além de contextualizar a questão com informações do Ministério Público e da Controladoria-Geral, demonstrando busca ativa por informações (UA03). Apesar da apuração, a matéria se baseia fortemente na reprodução de documentos oficiais e de conteúdo disponibilizado por órgãos públicos (UA04). A complexidade da matéria ao abordar corrupção, improbidade, contratos públicos, prescrição e recomendações técnicas exige compreensão de processos legais, administrativos e institucionais. Isso demanda aprofundamento do conteúdo (UA05). Não há erro, manipulação ou omissão de informações. O conteúdo é coerente com os dados oficiais, evitando distorções ou descontextualizações (UA06). As fontes são claramente indicadas, sendo elas a juíza Célia Regina Vidotti, da Vara Especializada em Ações Coletivas, o Ministério Público do Estado de Mato Grosso (MPMT), delações premiadas de ex-gestores e a decisão publicada no *Diário de Justiça* (UA07). A matéria não menciona solicitação de dados por mecanismos formais. A apuração se deu por acesso a documentos públicos e processuais já disponíveis (UA08). A LAI não é mencionada, embora o conteúdo seja baseado em dados públicos (UA09). A LGPD também não é citada (UA10). Não há apoio visual como infográficos, mapas ou linhas do tempo (UA11).

Por sua vez, na notícia **“PC recupera R\$ 30 mil subtraídos de comerciante que teve celular roubado”** (<https://tinyurl.com/ycxtwesp>), do dia 12 de junho de 2023, identificamos que o título apresenta dado numérico de R\$ 30 mil, quantia furtada e recuperada (UA01). O corpo da notícia reforça o valor total envolvido (R\$ 30 mil), o número de transferências (cinco), o número de contas bancárias distintas (quatro), a data do crime e das ações da polícia (UA02). Não há evidência de que o veículo tenha realizado apuração própria além da reprodução de

informações fornecidas pela Polícia Civil. Isso caracteriza ausência de busca ativa por dados (UA03). As informações são baseadas em declarações e materiais da Polícia Civil, limitada à reprodução de dados e declarações oficiais (UA04). Embora a estrutura da notícia seja simples, o caso envolve elementos que demandariam maior aprofundamento, como a vulnerabilidade de dados bancários após furto de celular, o funcionamento de esquemas de fraude via PIX, a eficiência da articulação entre delegacias e o setor antifraude. A ausência desse aprofundamento limita o potencial informativo da matéria (UA05). A notícia não apresenta elementos de desinformação e os dados são compatíveis com fontes oficiais, sem distorção, omissão ou extrapolação indevida (UA06). A fonte principal é claramente identificada como a Polícia Civil, com menção nominal ao delegado Marcos Paulo Batista de Oliveira (UA07). Não há menção à solicitação de informações via meios legais, caracterizando apuração passiva (UA08). A LAI não é mencionada (UA09) e, apesar do tema envolver uso indevido de dados pessoais e bancários, a LGPD também não é citada (UA10). Não há recursos visuais para ilustrar o caminho do dinheiro, etapas da investigação ou o vínculo entre vítimas, criminosos e contas bancárias (UA11).

Em análise da notícia **“Cuiabá aplica cerca de 1,4 mil multas todos os dias”** (<https://tinyurl.com/mr2ratvp0>), publicada no dia 21 de junho de 2023, identificamos título com o dado numérico referente a “1,4 mil multas por dia” (UA01). A matéria apresenta diversos dados, como o total de 514.444 multas em 2023, os locais monitorados, o número de pontos de fiscalização eletrônica (49), e menções a acidentes fatais evitados (UA02). A reportagem indica que os dados foram obtidos diretamente da Secretaria Municipal de Mobilidade Urbana (Semob), com entrevista do diretor de trânsito, Michel Diniz. Há menções a estudos técnicos e análise de laudos, evidenciando algum grau busca ativa de dados (UA03), embora a maior parte dos dados e argumentos sejam reproduções de discursos institucionais da Semob e de justificativas da gestão municipal (UA04). Apesar da cobertura ser informativa, o tema exige aprofundamento em áreas como efetividade real da fiscalização eletrônica, arrecadação com multas e sua destinação, comparação com índices de acidentes em outras capitais, bem como a percepção da população sobre as medidas (UA05). A notícia não contém informações falsas ou distorcidas. Os dados estão contextualizados com as falas das fontes institucionais (UA06). A Semob é citada explicitamente como fonte dos dados (UA07) e a informação foi fornecida por meio de declaração institucional, sem evidência de requisição pública documentada (UA08). A LAI não é citada, nem usada como ferramenta de apuração. (UA09). A LGPD também não é mencionada (UA10). Não há recursos visuais como mapas de pontos fiscalizados, evolução das multas por ano ou gráficos comparativos (UA11).

Os resultados da análise de conteúdo do *Gazeta Digital* com base nas categorias estabelecidas conforme livro de códigos (Bardin, 2011; Sampaio e Lycarião, 2021) estão detalhados no Quadro 8.

Quadro 8 – Análise de conteúdo do *Gazeta Digital*

Notícias Analisadas		Sim	Não
UA01	Título com dados	14	04
UA02	Conteúdo com dados	18	0
UA03	Busca por banco de dados	06	12
UA04	Reprodução de dados	18	0
UA05	Requer aprofundamento	14	04
UA06	Desinformação com dados	0	18
UA07	Fonte dos dados citada	18	0
UA08	Solicitação ativa de dados	0	18
UA09	Menciona LAI	0	18
UA10	Menciona LGPD	0	18
UA11	Infográfico	0	18

Unidade de Análise	Total
UA07 Fonte dos Dados	18
Oficial	17
Enquete	01

Fonte: elaborado pela autora (2025).

Figura 11 – Análise de conteúdo do Gazeta Digital



Fonte: elaborado pela autora (2025).

Verificamos que, ao todo, 14 ou 77,7% das 18 notícias e reportagens coletadas do webjornal *Gazeta Digital* que compõem o *corpus* incluíram dados numéricos no título (UA01). Isso se alinha ao critério de noticiabilidade reforçando o uso de dados como elemento de impacto jornalístico.

Outras 4 notícias, ou 22,2% das publicações, não incorporaram dados numéricos como informação nos títulos, apenas termos subjetivos (UA01).

Todas as 18 publicações submetidas à análise por meio do livro de códigos apresentaram conteúdos numéricos, porém, em um nível básico, aproximando-se de um Jornalismo **com Dados** (Mancini; Vasconcelos, 2016). As matérias ainda se enquadram, portanto, em um modelo mais tradicional, com uso superficial ou instrumental de dados, quase sempre apenas reproduzindo informações repassadas por fontes oficiais e não a partir da lógica de Jornalismo de Dados ou Jornalismo Guiado por Dados (Matzat, 2011) e mais distante ainda de um Jornalismo Digital em Base de Dados (JDBD), caracterizado por integrar algoritmos, linguagens de programação, metadados e técnicas de mineração de dados, operando sobre uma infraestrutura tecnológica ampliada e sendo sustentado pela lógica da plataformação, com uso de sistemas de gestão de conteúdo, equipes especializadas e ferramentas que possibilitam a automatização, a criação de *sites* dinâmicos e a incorporação de narrativas multimídia (Barbosa, 2005; 2007; 2008; 2013).

Das publicações, 6, ou 33,3%, demonstraram busca ativa em banco de dados, enquanto a maioria delas, sendo 12 ou 66,6% de 18, não realizou coleta em bases de dados, utilizando dados de maneira básica e passiva, principalmente com a reprodução de informações oficiais. Além disso, todas as 18 publicações se limitaram a reprodução de dados, sem construir novos sentidos, limitando seu potencial interpretativo.

Entendemos que 14, ou 77,7%, das 18 publicações requerem aprofundamento de dados, sendo predominantemente factuais, com espaço para ampliação de fontes informacionais, formatos e apuração. Em outras 4, ou 22,2%, não identificamos necessidade de aprofundamento de dados. Nenhuma das 18 notícias e reportagens induzem à desinformação, uma vez que as fontes dos dados são devidamente citadas. Contudo, ainda que não haja desinformação, o conteúdo requer aprofundamento analítico para que os dados ganhem maior valor interpretativo e jornalístico, evitando que o material fique restrito ao uso reprodutivo e limitado de dados, comum em coberturas que dependem de *releases* institucionais ou informações prontas.

Nenhum dos 18 conteúdos informacionais analisados utilizaram a LAI, tampouco citam a LGPD. A ausência de menções a esses dispositivos legais aponta um uso superficial das legislações que regem o acesso e uso de dados públicos no Brasil, quando poderia abrir espaço para discussões sobre acesso, uso e proteção dessas informações.

Além disso, nenhuma das 18 notícias e reportagens analisadas incluíram infográficos. A inovação no jornalismo digital requer visualização de dados, narrativas interativas e contextualização. A ausência de gráficos ou mapas no texto limita seu potencial informativo e

de engajamento, diminuindo a acessibilidade visual dos dados

Quanto à pluralidade de fontes, notamos predominância de fontes oficiais, estando presentes em 17 ou 94,5% do total de 18 publicações. Há, portanto, forte dependência de fontes oficiais, sem esforço jornalístico adicional de busca e coleta de dados para informações adicionais. O tratamento dos dados é reprodutivo e pouco elaborado, limitado à função de alerta público, não atendendo aos critérios de profundidade esperados no jornalismo. Há reprodução correta, mas falta cruzamento com outras fontes, interpretação crítica ou uso de tecnologias computacionais para análise mais profunda.

Apenas uma matéria, ou 5,5% de 18, foi produzida com base em indicadores (dados) coletados de enquete do próprio webjornal *Gazeta Digital*.

Segundo Barbosa (2005; 2007; 2008; 2013), o JDBD se baseia em pilares como busca ativa por bases de dados públicas (transparência, cruzamento e validação), interatividade e visualização (infográficos, mapas, *dashboards*) e produção autoral sobre dados estruturados. Para avançar em termos de jornalismo de dados, seria necessário maior exploração das bases públicas, pedidos via LAI, contextualização histórica e comparativa dos dados, além de uso de visualizações e articulação com a LGPD quando pertinente.

Constatamos, portanto, que as notícias publicadas pelo *Gazeta Digital* atendem predominantemente aos requisitos do Jornalismo **com** Dados (Mancini; Vasconcelos, 2016).

5.2 Análise do uso de dados em notícias e reportagens do *Correio Braziliense* (DF)

Do veículo *Correio Braziliense* reunimos 1.364 conteúdos informacionais, entre artigos de opinião, entrevistas, *newsletters*, *podcasts*, vídeos, enquetes, notícias e reportagens. Desse total, escolhemos para análise, conforme critérios estabelecidos nesta pesquisa, 36 publicações, com características de notícias e reportagens produzidas pelos jornalistas do próprio webjornal, conforme descreveremos a seguir.

Em notícia publicada pelo no dia 1º de março de 2023, intitulada “**Entenda mudanças que resultaram na alta no preço dos combustíveis**” (<https://tinyurl.com/2e2698yk>), percebemos que o título não apresenta dados numéricos, percentuais ou informações quantitativas, sendo de caráter mais geral, descritivo ou opinativo, sem trazer dados estatísticos ou referências diretas a bancos de dados (UA01). O corpo do texto contém informações quantificáveis, como números absolutos, percentuais ou valores monetários. Isso indica o uso de dados como elemento estruturante do conteúdo (UA02). A matéria analisada não evidencia que o jornalista tenha feito uma busca ativa em bases de dados públicas, privadas ou

especializadas (UA03). Os dados apresentados no texto são reproduzidos diretamente de uma fonte, sem elaboração própria ou interpretação aprofundada. Isso implica uma prática mais passiva de uso dos dados, com baixa incidência de análise crítica ou cruzamentos com outras informações (UA04). Apesar de conter dados, o conteúdo exige maior aprofundamento para oferecer uma compreensão mais completa do tema (UA05). Não há indícios de que as informações numéricas tenham sido usadas de maneira distorcida, enganosa ou manipulada (UA06). A fonte dos dados é identificada e tem caráter oficial, neste caso, a Petrobras (UA07). Não há evidência de que os dados tenham sido obtidos por meio de mecanismos de transparência ativa, como entrevistas exclusivas ou levantamentos próprios (UA08). O conteúdo não faz menção à LAI, o que reforça a percepção de que não houve solicitação ativa de dados públicos (UA09). A matéria também não trata da LGPD, o que indica que o conteúdo não aborda questões relacionadas ao uso, à proteção ou ao tratamento de dados pessoais (UA10). Não há apoio visual de infografia, como gráficos, tabelas ou mapas, apenas ilustração (UA11).

Em outra notícia publicada no dia 1º de março de 2023 pelo *Correio Braziliense* e intitulada “**Com restrições, Moraes libera 137 bolsonaristas presos por atos de 8/1**” (<https://tinyurl.com/4mbanb77>), verificamos que o título o apresenta informação numérica clara (“137 bolsonaristas”) (UA01). O texto também traz números como o total de presos, de liberados e as medidas aplicadas, como “137 bolsonaristas liberados”, “803 seguidores presos” e “603 foram liberados” (UA02). Não há informações de investigação em bases de dados públicas ou especializadas. A matéria é orientada por informações fornecidas pelo Supremo Tribunal Federal (STF) e outras autoridades (UA03). O conteúdo reproduz dados de fontes oficiais, como o STF e a Procuradoria Geral da República (PGR), sem análise própria (UA04). O conteúdo está centrado na informação factual, sem análise sobre o impacto jurídico, político ou social da decisão (UA05). Os dados são coerentes com os fatos notórios e públicos sobre os acontecimentos de 8 de janeiro de 2023 e a matéria não deturpa nem distorce os números, nem apresenta informações enganosas ou manipuladas (UA06). A matéria identifica como fonte o STF, porém sem mencionar nenhum outro contraponto, como, por exemplo, a defesa dos investigados, enfraquecendo o pluralismo de informações (UA07). Não há evidência de que o veículo tenha solicitado dados, ficando restrito ao conteúdo divulgado pelas autoridades, o que denota uma dependência das fontes oficiais. (UA08). A LAI não é mencionada (UA09), tampouco a LGPD (UA10). Não há apoio visual para facilitar a interpretação dos dados, como linha do tempo dos processos, mapa da localização dos detidos ou proporção entre presos e liberados (UA11).

Uma terceira notícia publicada no dia 1º de março de 2023 apresenta o título **“Até 174 mil brasileiros terão autorização de residência automática em Portugal”** (<https://tinyurl.com/yace3cju>), com informação quantitativa clara (“Até 174 mil brasileiros”) (UA01). A matéria apresenta vários dados numéricos, como “174mil brasileiros”, “233,1 mil brasileiros com registros válidos”, “15% do PIB de Portugal vem do turismo”, “R\$ 84 pela emissão do cartão” (UA02). Não há indicativos de que o jornalista tenha realizado levantamento próprio em bases públicas ou bases de dados abertas, como estatísticas oficiais acessadas via ferramentas interativas (UA03). As informações são reproduzidas diretamente das autoridades (UA04), não havendo transformação interpretativa dos dados (UA05). Os dados são coerentes, referenciados, sem indícios de manipulação ou distorção (UA06). A principal fonte de reportagem é oficial, como o Serviço de Estrangeiros e Fronteiras (SEF) de Portugal. Há também menções a fontes institucionais, como o ministro da Administração Interna e a Embaixada do Brasil (UA07). Não há evidências de que os dados tenham sido obtidos por meio de pedido formal do repórter (UA08). A LAI não é mencionada (UA09) e também não é citada a LGPD (UA10). A reportagem não apresenta gráficos, mapas ou visualizações de dados (UA11).

Também no dia 1º de março de 2023, uma quarta notícia, intitulada **“Governo anuncia Imposto sobre Exportação de petróleo cru para fortalecer receitas”** (<https://tinyurl.com/yc88uyb6>), apresenta título contendo uma informação factual de natureza econômica, relativa à criação de imposto sobre exportações de petróleo. Apesar de não citar um número exato, comunica uma medida concreta de impacto fiscal, o que se enquadra como uso informativo de dado econômico/político (UA01). A matéria é rica em números, como a alíquota do novo imposto de 9,2%, previsão de arrecadação de R\$ 6,7 bilhões em quatro meses, impacto no lucro da Petrobras de 1%, variação de ações PETR4: -3,48%, PETR3: -4,39%, PRIO3: -9,04%, RRRP3: -6,97%. Esses números mostram impacto fiscal e financeiro tanto para o governo quanto para o mercado, caracterizando uma abordagem jornalística voltada para dados econômicos e de mercado (UA02). Não há evidências de consulta ativa a bases públicas como dados brutos da Receita Federal, Agência Nacional de Petróleo (ANP), Comissão de Valores Imobiliários (CVM) ou Banco Central (BC). A informação é baseada em declarações institucionais e dados fornecidos por agentes públicos e pelo mercado financeiro (UA03). Os dados apresentados são reproduzidos diretamente de fontes oficiais ou do mercado, sem análise técnica (UA04). Apesar da presença de dados relevantes, o conteúdo não contextualiza a arrecadação em comparação com receitas anteriores, não delinea projetos econômicos em caso de prorrogação ou suspensão do imposto, bem como não detalha impactos para consumidores,

refinarias ou balança comercial. Há potencial para análise crítica, mas isso não é explorado (UA05). Os dados são consistentes com fontes oficiais e de mercado e não há distorções, manipulações ou omissões (UA06). As principais fontes são oficiais e institucionais, como Ministério da Fazenda, Ministério de Minas e Energia (MME), Ibovespa e dados de mercado (UA07). Não há pedido de informações ou investigação baseada em dados exclusivos (UA08). A LAI não é mencionada e não percebemos esforço para revelar dados que estão fora do circuito público tradicional (UA09). Não há referência à LGPD (UA10). Não há infográficos apesar do volume de números e porcentagens, que possibilitaria uma visualização simples, como um gráfico da arrecadação prevista, um comparativo do desempenho das ações das petroleiras ou uma linha do tempo da vigilância do MP, que ajudariam na leitura rápida e compreensão do impacto econômico (UA11).

A notícia intitulada “**BC reabre busca por dinheiro esquecido em instituições financeiras**” (<https://tinyurl.com/ypk5aw9e>), publicada pelo *Correio Braziliense* no dia 1º de março de 2023, revela que o título não apresenta números ou percentuais, embora trate de um tema com alto potencial quantitativo. Ele se limita a anunciar a ação do Banco Central (BC), sem destacar valores ou estatísticas (UA01). A matéria apresenta dados relevantes, como estimativa total de valores esquecidos (R\$ 6 bilhões), total de beneficiários (38 milhões de CPFs e 2 milhões de CNPJs), além de dados-chave como início das consultas (28/02) e início dos pagamentos (03/07/2023) (UA02). Não há evidência de uso de bases abertas ou ativadas de dados. A apuração está ancorada em informações oficiais e centralizadas do BC, divulgadas via comunicado ou coletiva de imprensa (UA03). O conteúdo se limita a reproduzir as estimativas e instruções fornecidas pelo BC, sem aplicar análise crítica, segmentação de dados por perfil, região ou grupo. O uso de dados é passivo e funcional (UA04). Apesar da utilidade pública da notícia, ela carece de contextualização histórica, análise crítica, aprofundamento social e explicação sobre como os valores ficaram “esquecidos”. Essas lacunas indicam potencial inexplorado de aprofundamento na cobertura (UA05). Os dados são coerentes com a fonte oficial, que é o BC, sem indícios de erros, distorções ou manipulações (UA06). A fonte dos dados é única e oficial. Isso confere legitimidade, mas limita a diversidade de vozes e interpretações, sem especialistas independentes, consumidores ou entidades financeiras opinando sobre a reabertura do Sistema de Valores a Receber (SVR) (UA07). Não há iniciativa investigativa, sendo a cobertura reativa a uma agenda pública, baseada em declarações institucionais com dados (UA08). A LAI não é mencionada, tampouco sugerida como instrumento de aprofundamento ou fiscalização do sistema (UA09). Apesar de o assunto tratar de acesso a informações pessoais e financeiras, não há menção à LGPD (UA10). A ausência de

visualizações gráficas é uma limitação, considerando o alto volume de dados (R\$ 6 bilhões, milhões de CPFs/CNPJs), a diversidade de origens dos valores e o cronograma de funcionamento do sistema. A inserção de mapas, fluxogramas ou gráficos de distribuição poderia ampliar a compreensão e retenção das informações (UA11).

No dia 1º de março de 2023, a notícia intitulada “**Desemprego recua em oito das 27 unidades da Federação no 4º trimestre de 2022**” (<https://tinyurl.com/mveyh8jr>) revela que, apesar de tratar de dados estatísticos, o título não apresenta números específicos ou percentuais, limitando-se a uma descrição genérica da variação de desemprego em parte dos estados (UA01). O corpo da matéria é fortemente baseado em dados estatísticos, como variação nacional do desemprego de 8,7% para 7,9%, estados com maiores quedas – Mato Grosso do Sul (–1,8 pp), PE (–1,6 pp), Bahia (–1,6 pp), Amapá como única alta: +2,5 pp; menores taxas: Rondônia (3,1%), Santa Catarina (3,2%), MS (3,3%), MT (3,5%); maiores taxas: Bahia (13,5%), Amapá (13,3%) (UA02). A apuração se limita à divulgação oficial do IBGE, sem evidência de que um jornalista tenha consultado diretamente a base da PNAD Contínua ou feito outras filtragens personalizadas (UA03). Trata-se de uma cobertura dependente da agenda oficial, ou seja, reativa e declaratória. A reportagem reproduz dados da PNAD Contínua sem realizar transformações analíticas, cruzamentos ou interpretações além do que é fornecido pelo próprio IBGE (UA04). Não há análise socioeconômica ou interpretativa dos dados, como, por exemplo, informalidade, desigualdades estruturais, acesso por gênero e raça (UA05), e o conteúdo é baseado em material de divulgação oficial, sem indícios de desinformação (UA06). As fontes dos dados são oficiais, claramente identificadas (UA07). Não há pedidos formais de acesso à informação, uma vez que o conteúdo é baseado em material de divulgação (UA08). A LAI não é mencionada, nem relacionada à obtenção ou uso de dados (UA 09). Apesar do uso de dados agregados pela UF, não há menção à LGPD (UA10). A ausência de infográficos, tabelas ou mapas prejudica a clareza da comparação entre os estados e a visualização das tendências (UA11).

Na análise de conteúdo aplicada à notícia intitulada “**Na mira do Judiciário: 96% dos feminicidas foram condenados no DF**” (<https://tinyurl.com/4s8dhf89>), publicada no dia 09 de março de 2021 pelo *Correio Braziliense*, observamos que ela traz título com dado percentual (96%) (UA01). O texto contém diversos dados quantitativos, como o número de medidas protetivas, denúncias, percentuais de variação e dados históricos (UA02). A reportagem não indica que o jornalista fez busca ativa em bases públicas de dados, mas que estes provêm exclusivamente de fonte oficial (UA03). Os dados são reproduzidos a partir de levantamento do Ministério Público do Distrito Federal e do Fórum Brasileiro de Segurança Pública (UA04),

carecendo de interpretação crítica e contextualização. O próprio promotor citado na notícia menciona a necessidade de se amadurecer a leitura dos dados (UA05). A reportagem não manipula ou distorce os dados; ao contrário, problematiza e questiona os números (UA06). A matéria informa claramente a origem dos dados, que são o Ministério Público do Distrito Federal (MPDF) e o Fórum Brasileiro de Segurança Pública (UA07). Não há menção à obtenção dos dados por meio de pedido ativo de informação (UA08). A reportagem não faz referência à LAI (UA09) e também não há citação da LGPD (UA10). A matéria não apresenta elementos visuais de apoio, como gráficos ou tabelas (UA11).

Também publicada no dia 09 de março de 2023, a notícia intitulada “**37% das mulheres brasileiras investem na bolsa de valores**” (<https://tinyurl.com/kf9jck5w>) apresenta dado numérico central: “37% das mulheres brasileiras investem na bolsa” (UA01). A matéria contém diversos valores percentuais, como participação por gênero, classe social, idade e tipo de investimento (UA02). O texto não indica que a redação consultou diretamente bases de dados, que provêm de pesquisa consolidada pela Anbima/Datafolha (UA03). A reportagem apenas reproduz os dados apresentados pela pesquisa institucional (UA04). A matéria toca em pontos que pedem interpretação, como o recorte de gênero e as motivações para investir, porém não os aprofunda (UA05). Os dados são apresentados de forma clara, fiel e contextualizada (UA06). A pesquisa é claramente atribuída à Anbima, em parceria com o Datafolha (UA07), e não há evidência de que o veículo tenha solicitado os dados diretamente às fontes (UA08). A LAI não é citada (UA09) e também não há menção à LGPD (UA10). A notícia não traz apoio visual, como gráficos, tabelas ou visualizações dos dados (UA11).

Na análise de conteúdo da reportagem “**Mudanças climáticas ameaçam produção de café no Brasil**” (<https://tinyurl.com/mt2uaeam>), publicada no dia 09 de março de 2023 no *Correio Braziliense*, identificamos que o título não apresenta números ou percentuais (UA01). A matéria apresenta dados percentuais, como 98% de pequenos produtores, 75% a 80% da produção mundial de café, aumento de até 54% na produção, entre outros (UA02). Não há indício de que os dados tenham sido buscados ativamente em bases públicas. O conteúdo se baseia em entrevistas e dados institucionais (UA03). Os dados são apresentados conforme fornecidos pelas fontes, sendo elas a *International Coffee Partners* (ICP), uma análise de Opitz, Löfberg e Nabeiro (UA04). A reportagem requer interpretação sobre clima, sustentabilidade, segurança alimentar e estrutura fundiária brasileira (UA05). Ela utiliza dados reais, contextualizados e com atribuição clara, sem distorções (UA06). A fonte principal é a *International Coffee Partners* (ICP), citada nominalmente em vários trechos (UA07). Não há menção a pedido ativo de informações por parte do jornalista (UA08). A reportagem não

menção a LAI (UA09) e não há referência à LGPD (UA10). Apesar da densidade de dados, não há visualizações gráficas que auxiliem a leitura e a compreensão (UA11).

A análise de conteúdo da reportagem **“Dia da Mulher: mulheres são maioria em cursos profissionalizantes”** (<https://tinyurl.com/ytx9u5kh>), publicada em 09 de março de 2023, revela que o título destaca o protagonismo feminino, mas não apresenta dados numéricos (UA01). A reportagem contém dados estatísticos e percentuais, como, por exemplo, 58% de mulheres na EPT, 85% de presença em cursos de saúde, entre outros (UA02). Os dados vêm de estudos do Itaú Educação e Trabalho e da Fundação Roberto Marinho. Não há, portanto, busca ativa por parte do repórter (UA03). Os dados são transcritos conforme divulgados pelas instituições responsáveis (UA04). O tema envolve desigualdade de gênero, juventude, mercado de trabalho e economia do cuidado, exigindo análise multidimensional (UA05). Os dados são apresentados sem distorções (UA06) e as fontes são declaradas de forma clara e oficial (UA07). Não há evidência de busca autônoma de dados públicos pelo repórter (UA08). A LAI não é mencionada (UA09) e a LGPD também não é citada (UA10). Apesar de conter muitos dados percentuais, não há visualização gráfica que facilite a leitura e interpretação dos números (UA11), contrariando entendimento de Casadei, Rodrigues e Biernath (2016) de que os infográficos integram os principais meios expressivos do jornalismo de dados.

Na análise de conteúdo da notícia **“Brasília é a cidade em que mais se fala palavrões no Brasil; veja o ranking”** (<https://tinyurl.com/2paymwpu>), coletada no dia 17 de março de 2023, observamos que o título não traz números nem menções a dados quantitativos (UA01). A matéria, contudo, apresenta quantitativos como o número médio de palavrões por cidade, percentuais de locais onde são ditos palavrões, percentuais de quem evita falar palavrões em certas situações (UA02). A notícia se baseia em uma pesquisa privada encomendada pela empresa Preply, não havendo uso de banco público. Os dados são analisados e organizados por outra instituição e não pelo jornalista (UA03). O conteúdo reproduz integralmente os resultados fornecidos pela pesquisa, sem cruzamento com outras fontes (UA04). A matéria não propõe aprofundamento crítico, apenas relata dados descritivos e comportamentais (UA05), no entanto, a matéria mantém integridade na apresentação dos dados fornecidos pela fonte (UA06). Os dados são atribuídos à Preply, com metodologia citada e menção à empresa responsável pela coleta, que é a Censuswide (UA07). Os dados foram fornecidos pela empresa, sem ação ativa da reportagem (UA08). Não se trata de dados públicos obtidos por meio da LAI (UA09) e a matéria não aborda nem problematiza aspectos da LGPD (UA10). A reportagem é composta exclusivamente por texto, sem infografia (UA11).

Análise de conteúdo da notícia **“Agronegócio brasileiro exportou US\$ 9,89 bi em**

fevereiro” (<https://tinyurl.com/yf3m8fxr>), publicada em 17 de março de 2023 no *Correio Braziliense*, evidencia título que menciona diretamente o valor da exportação, de US\$ 9,89 bilhões, caracterizando uso de dado numérico (UA01). A matéria está repleta de números, como os valores de exportação, quedas percentuais, volumes por produto e comparações anuais (UA02). A reportagem não realiza consulta ativa a bancos públicos ou abertos, apenas utilizando dados divulgados em nota oficial (UA03), os quais foram reproduzidos de órgãos como Ministério da Agricultura e Pecuária (Mapa), Companhia Nacional de Abastecimento (Conab) e Organização Mundial de Saúde Animal (OMSA), sem reinterpretação ou cruzamento (UA04). A matéria levanta diversos fatores econômicos e climáticos, mas não aprofunda nem discute o impacto em regiões produtoras ou políticas públicas relacionadas ao tema (UA05). Não há evidência de manipulação ou uso incorreto dos dados, que são contextualizados e atribuídos (UA06). Fontes oficiais são identificadas, sendo elas o Mapa, a Conab e a OMSA (UA07). O jornalista não solicitou ou analisou dados próprios, apenas replicando nota oficial do Mapa (UA08). A matéria não menciona a LAI, nem depende dela para a obtenção dos dados (UA09). Não há qualquer relação com proteção de dados pessoais e, portanto, não se relaciona à LGPD (UA10). A matéria não utiliza recursos gráficos para facilitar a visualização dos dados apresentados (UA11).

A análise de conteúdo da reportagem **“Barulho do trânsito pode causar hipertensão, aponta pesquisa”** (<https://tinyurl.com/yf3m8fxr>), publicada no dia 25 de março de 2023, evidencia título que informa o achado da pesquisa, mas sem apresentar dado quantitativo (UA01). A matéria traz números absolutos, percentuais e faixas de decibéis, como 246 mil pessoas na amostra, 8,5% desenvolveram hipertensão, entre outros (UA02). Os dados foram extraídos de uma pesquisa científica publicada, sem indicação de apuração ativa em banco de dados (UA03). Os dados da pesquisa foram reproduzidos fielmente, com atribuição às fontes originais (UA04). A matéria envolve relações entre poluição sonora, doenças cardiovasculares e políticas públicas, exigindo interpretação multidisciplinar (UA05). Não há manipulação ou distorção dos dados apresentados, sendo que as informações são precisas e contextualizadas (UA06). A fonte é claramente atribuída ao estudo do American College of Cardiology, publicado no *JACC: Advances* (UA07). Não houve solicitação ativa de dados ou informações aos autores do estudo (UA08). A LAI não é mencionada (UA09) e a LGPD também não é citada (UA10).

Também em 25 de março de 2023, o *Correio Braziliense* publicou notícia intitulada **“Com recorde em número de casos, governo entra em alerta contra tuberculose”** (<https://tinyurl.com/27vf9xyy>). O título menciona “recorde em número de casos”, configurando

uma referência quantitativa direta (UA01). Apresenta dados numéricos e percentuais, como 5 mil mortes, 78 mil casos, cobertura vacinal inferior a 88%, incidência por 100 mil habitantes, entre outros (UA02). A matéria utiliza dados oficiais secundários, a exemplo de informações advindas da Organização Mundial de Saúde (OMS) e do Ministério da Saúde (MS), sem investigação direta em bancos públicos ou científicos (UA03). Os dados são claramente reproduzidos de fontes oficiais (UA04). O tema envolve políticas públicas, desigualdades sociais, imunização, prioridades globais de saúde e crise sanitária, mas sem aprofundamento (UA05). Os dados são citados corretamente e sem distorções (UA06). As fontes dos dados são devidamente atribuídas, sendo majoritariamente oficiais, além de especializadas (UA07). Não há evidência de solicitação de dados ou consulta ativa a bancos de dados públicos (UA08). A LAI não é mencionada (UA09) e a LGPD também não é citada (UA10). A matéria menciona mapa com os coeficientes de incidência por estado, embora ele não esteja reproduzido no corpo textual (UA11).

A análise de conteúdo da reportagem **“Entenda por que o diagnóstico tardio de autismo é mais comum em mulheres”** (<https://tinyurl.com/5n8m2yzh>), publicada no dia 02 de abril pelo *Correio Braziliense*, revela que o título é informativo, mas não apresenta dados quantitativos (UA01). A matéria cita dados como 1 em 36 crianças é autista, estimativa de 6 milhões no Brasil, desigualdade racial no diagnóstico, entre outros (UA02). A reportagem faz análise contextual de dados secundários, inclusive aplicando proporção dos Estados Unidos à realidade brasileira (UA03). Os dados são reproduzidos de fontes como o Centro de Controle de Prevenção e Doenças (CDC) e de especialistas (UA04). O texto aborda aspectos sociais, raciais, culturais e psicológicos do Transtorno do Espectro Autista (TEA) em mulheres, exigindo análise interseccional e contextual (UA05). Os dados estão contextualizados sem distorções (UA06). As fontes dos dados são citadas, como o CDC estadunidense e especialistas em neurologia e psicologia (UA07). Não há menção a pedidos formais de acesso à informação pública (UA08). A LAI não é mencionada (UA09) e a LGPD também não é citada (UA10). A matéria não contém infográficos, mapas, ou tabelas visuais (UA11).

Em notícia publicada no dia 02 de abril de 2023, com o título **“Sumiço sem resposta: Correio conta a história de meninas desaparecidas no DF”** (<https://tinyurl.com/mr9znyp5>), verificamos que o título é descritivo, sem dados quantitativos (UA01). A matéria menciona a existência de 6 meninas desaparecidas há pelo menos 15 anos, com nome, idade e local (UA02). Utiliza informações do banco de desaparecidos da Polícia Civil do Distrito Federal e do Laboratório de Representação Facial Humana (UA03). A reportagem reproduz dados públicos disponíveis no *site* oficial da Polícia Civil (UA04), mas carece de análise sobre memória pública

de casos antigos, desigualdade investigativa e desatenção institucional (UA05). Os dados são usados em contexto, sem manipulação ou distorção (UA06), sendo citada como fonte a Polícia Civil do Distrito Federal e seu laboratório oficial (UA07). Não há menção a uso de mecanismos de acesso à informação (UA08). A LAI não é citada (UA09) e a LGPD também não aparece (UA10). A matéria não traz visualizações gráficas ou recursos visuais de dados, como linha do tempo ou mapas (UA11).

Análise de conteúdo da reportagem **“Plataforma alerta para a vulnerabilidade de prevenção às tragédias no Brasil”**, publicada no dia 02 de abril de 2023 (<https://tinyurl.com/2xhkftdr>), com base nas unidades de análise propostas, revela título com caráter interpretativo, que não apresenta valores ou indicadores numéricos (UA01). A matéria traz dados como 86 cidades brasileiras reportaram dados espontaneamente, 18 mil ocorrências de desastres, 6 milhões de desabrigados, entre outros (UA02). Os dados vêm de plataformas e relatórios prontos, sem indícios de que o repórter acessou bancos de dados diretamente (UA03). Há reprodução de dados de fontes como o *Carbon Disclosure Project* (CDP) e o Painel Intergovernamental sobre Mudanças Climáticas (IPCC) (UA04). Requer análise sobre gestão de riscos climáticos, ausência de políticas públicas eficazes e transparência de dados ambientais (UA 05), embora os dados sejam usados corretamente, com fontes e contexto adequados (UA06), provenientes de fontes institucionais como o CDP e o IPCC (UA07). Não há menção a solicitações ou métodos de apuração direta via órgãos públicos (UA08). A LAI não é citada (UA09) e a LGPD também não é abordada (UA10). A matéria não apresenta gráficos, mapas ou visualizações interativas (UA11).

Uma análise de conteúdo jornalístico da matéria **“Despesas com saúde sobem rápido e chegam a 9,6% do PIB brasileiro”** (<https://tinyurl.com/3nav69yb>), publicada em 02 de abril de 2023, revela dado percentual (9,6% do PIB) como elemento central do título (UA01). A matéria é fundamentada em dados numéricos, como valores em bilhões de reais, percentuais, variação de emprego e inflação (UA02). Não há indicação de que o veículo tenha acessado diretamente bancos de dados primários; o conteúdo se baseia em relatórios prontos (UA03). O texto reproduz dados do Conta-Satélite de Saúde do Instituto Brasileiro de Geografia e Estatística (IBGE), Instituto de Pesquisa Econômica Aplicada (Ipea), Ministério da Saúde (MS) e Fiocruz (UA04). A matéria aborda fatores estruturais, demográficos e econômicos que explicam o crescimento dos gastos em saúde (UA05). Os dados são apresentados com coerência e atribuição de fontes (UA06). As fontes são citadas de forma clara, predominantemente institucionais, como o Ministério da Saúde, Ipea, IBGE, Fiocruz e Agência Nacional de Saúde (ANS) (UA07). Não há indícios de apuração direta junto a bancos de dados públicos (UA08).

A LAI não é citada (UA09) e a matéria não trata de dados pessoais, não mencionando, portanto, a LGPD (UA10). Apesar da riqueza de dados, não há visualizações gráficas, tabelas ou infográficos no conteúdo apresentado (UA11).

Com base nas unidades de análise indicadas, identificamos, na notícia **“5,5 milhões de bolsas de estudo em cursos de TI”** (<https://tinyurl.com/cd6c24s5>), publicada em 10 de abril de 2023, que o título inclui valor numérico (UA01). São mencionados ainda diversos números, como 5,5 milhões de bolsas, 134 cursos, 159 mil vagas por ano previstas até 2025 e 53 mil formandos anuais (UA02). Não há menção a pesquisa ativa ou cruzamento com bancos de dados públicos ou privados (UA03), sendo reproduzidos dados fornecidos por órgãos oficiais, como o Ministério do Trabalho e Emprego (MTE) e estudo de empresa privada de recrutamento (UA04). Embora a matéria traga dados quantitativos e fontes, não aprofunda questões estruturais como a desigualdade de acesso, barreiras digitais ou efetividade da educação à distância (EAD) para populações vulneráveis (UA05). As informações são claras, contextualizadas e oriundas de fontes devidamente identificadas (UA06). As fontes são claramente atribuídas, sendo elas o MTE, a Microsoft, especialistas do governo, além de referência a estudo do setor privado (UA07). Não há evidência de que os dados tenham sido solicitados diretamente pelo veículo (UA08). A notícia não trata da LAI (UA09) e não há nenhuma referência à LGPD, mesmo considerando a coleta de dados dos participantes (UA10). A notícia não traz infográficos, gráficos ou visualizações de dados, apenas texto e imagem genérica de banco de imagens (UA11).

A matéria publicada no dia 18 de abril de 2023, com o título **“Redes sociais derrubaram 756 perfis após violência nas escolas, diz Flávio Dino”** (<https://tinyurl.com/bdf88za7>), traz “756 perfis” como dado numérico central da informação (UA01). A notícia apresenta outros quantitativos, como 756 perfis suspensos, 100 pedidos de preservação de conteúdo, 225 prisões e apreensões, 7.473 denúncias e 377 casos em investigação (UA02). Não há indicação de uso de bancos públicos, cruzamento de dados ou solicitação ativa de acesso (UA03). A jornalista reproduz informações fornecidas por autoridades, como os ministros Flávio Dino e Rosa Weber (UA04). Apesar do volume de dados, a cobertura não analisa a eficácia das remoções, a origem da desinformação e nem o papel das plataformas na contenção do problema (UA05). A matéria não contém distorções ou uso inadequado de dados (UA06). Todas as informações são atribuídas a fontes oficiais, como o Ministério da Justiça, o Supremo Tribunal Federal (STF) e a reunião com o presidente da República (UA07). Não há evidência de busca investigativa de informações (UA08). A LAI não é citada, apesar do volume de dados envolvidos (UA09). A LGPD também não é

mencionada, embora a temática envolva perfis e conteúdos preservados para investigação (UA10). A notícia não apresenta infográficos ou recursos visuais para facilitar a leitura dos dados (UA11).

Análise de conteúdo da notícia publicada no dia 18 de abril de 2023, no *Correio Braziliense*, com o título “**Pesquisa mostra que remoção de ovários nem sempre compensa em casos de câncer**” (<https://tinyurl.com/5nyh5jke>), revela título informativo sem valores numéricos (UA01). A matéria traz o número total de mulheres analisadas (142.985) e grupos específicos dentro da amostra, sendo 22.974 delas com ovários removidos. Também apresenta dados sobre riscos e consequências a longo prazo (UA02). A pesquisa científica analisada é baseada em banco de dados nacional dinamarquês e inclui registros médicos e demográficos. Além disso, há uso de dados secundários qualificados (UA03). A jornalista apenas reproduz os dados e conclusões do estudo científico (UA04). A complexidade dos achados requer interpretação por envolver múltiplas variáveis, como idade, estado menopausal e risco genético (UA05). Os dados são apresentados alinhados à fonte original e com contextualização médica (UA06). As fontes são citadas explicitamente, incluindo cientistas da Sociedade Dinamarquesa do Câncer, revista *Annals of Internal Medicine* e médica brasileira especialista em ginecologia (UA07). Os dados foram extraídos de fontes científicas, não obtidos por iniciativa do veículo de imprensa (UA08). Por não se tratar de pesquisa brasileira, não há menção à LAI (UA09). Mesmo tratando de dados de saúde, a reportagem não discute o uso, proteção ou anonimização desses dados e não cita a LGPD (UA10). Não há gráficos ou visualizações que facilitem a leitura das informações complexas (UA11).

A análise da notícia “**Catadores de materiais recicláveis são agentes da saúde e do meio ambiente**” (<https://tinyurl.com/4hwar5jk>), publicada em 18 de abril de 2023, revela que o título é descritivo, sem apresentar valores numéricos (UA01). Há vários dados relevantes, como a citação de 1.513 catadores no Distrito Federal, 42.505 toneladas processadas em 2022, 46% de aproveitamento, entre outros (UA02). Os dados são apresentados como fornecidos por órgãos oficiais e fontes institucionais, sem investigação própria (UA03), tendo sido reproduzidos de fontes como o Serviço de Limpeza Urbana do Distrito Federal (SLU/DF), sem análise técnica ou cruzamento com outras bases (UA04). A matéria levanta questões estruturais importantes como o preconceito, a informalidade, a remuneração e a segregação, mas não aprofunda com dados comparativos, políticas públicas ou estatísticas mais abrangentes (UA05). Os dados e as informações são contextualizados e atribuídos a fontes (UA06), vindo de setores oficiais como o Serviço de Limpeza Urbana do Distrito Federal (SLU/DF), especialistas e depoimentos de catadores (UA07). Não há evidência de consulta a bancos de dados públicos

(UA08). Não menciona a LAI (UA09) e, apesar de abordar dados de trabalhadores e cooperativas, não há menção à proteção de dados pessoais relacionada à LGPD (UA10). Nenhuma representação visual acompanha o conteúdo, apesar da riqueza de dados (UA11).

Análise de conteúdo da notícia “**Sem aprovação do arcabouço fiscal, governo só tem recurso para pagar o básico**” (<https://tinyurl.com/mtuhr5ub>), publicada em 18 de abril de 2023, revela que o título é contextual, mas não apresenta valores numéricos (UA01). A matéria utiliza dados como R\$ 24,4 bilhões (sem arcabouço) e R\$ 172 bilhões (com arcabouço), além de citar gastos públicos (UA02). Os dados são públicos e apresentados pelas autoridades, como a ministra do Planejamento e Orçamento, Simone Tebet (UA03). A reportagem reproduz informações do Projeto de Lei de Diretrizes Orçamentárias (PLDO) e falas oficiais sem cruzamento ou análise própria (UA04) e, apesar de introduzir um tema relevante, não aprofunda os efeitos orçamentários por setor, nem compara com anos anteriores (UA05). As informações são lastreadas em fonte oficial, que é o governo federal (UA06). Cita o PLDO e a ministra do Planejamento e Orçamento, Simone Tebet (UA07). Não há indícios de busca de informações ou questionamentos complementares por parte do veículo (UA08). Não menciona a LAI (UA09), nem a LGPD (UA10). Apesar da presença de dados numéricos, não há representação visual ou quadro comparativo (UA11).

Análise de conteúdo da matéria “**Calor extremo: estudo mapeia áreas de risco no planeta**” (<https://tinyurl.com/4u9dph7t>), publicada no dia 26 de abril de 2023, revela que o título é interpretativo e não apresenta valores numéricos (UA01). A reportagem apresenta percentuais como 31% e dados geográficos baseados em modelos (UA02) e a jornalista se baseia exclusivamente em informações científica já publicadas, como estudo da *Nature Communications* (UA03). Os dados científicos do artigo são reproduzidos de forma didática e contextualizada (UA04). O tema é tratado com certa profundidade, mas há margem para explorar mais dados específicos por país ou região (UA05). A matéria é fiel às fontes científicas e evita generalizações (UA06), citando fontes dos dados, sendo elas a Universidade de Bristol e a revista científica *Nature Communications* (UA07). Não há evidência de mobilização por dados ou informações adicionais junto às instituições ou aos autores (UA08). Não menciona LAI (UA09), nem LGPD (UA10). Apesar de mencionar mapa, o conteúdo não apresenta visualizações na versão analisada (UA11).

A análise de conteúdo da notícia “**PF e CGU investigam 3º registro de vacina falsa de Bolsonaro, 2º em SP**” (<https://tinyurl.com/e2ytddf8>), publicada no dia 04 de maio de 2023, revela que o título não contém dados quantitativos, sendo informativo, com foco na investigação (UA 01). Há referências a datas, locais e uso de aplicativos públicos com registro

de ações (UA02). Os dados são oriundos de documentos e investigações da Polícia Federal (PF) e da Controladoria Geral da União (CGU), sem aprofundamento adicional da reportagem (UA03). A matéria reproduz trechos do relatório da PF e eventos com precisão documental (UA04), mas tem potencial de aprofundamento, como implicações jurídicas, contexto político e possíveis desdobramentos do caso (UA05). A reportagem é informativa, com dados verificados e fontes oficiais (UA06), apontando a PF, a CGU, o sistema ConecteSUS e documentos de investigação (UA07). A matéria não menciona pedidos diretos por informações ou investigação jornalística própria (UA08). Não cita a LAI (UA09) e, embora faça referência a dados pessoais, como Cadastro de Pessoa Física (CPF), Protocolo de Internet (IP) e certidões, a LGPD não é citada (UA10). Não há mapas, gráficos ou linha do tempo visual para facilitar a compreensão (UA11).

Publicada no dia 04 de maio de 2023, a matéria intitulada “**PF prende 47 pessoas com armas de fogo em situação irregular, diz Dino**” (<https://tinyurl.com/58m7ez6h>) apresenta título factual e direto, com dados numéricos (UA01). O texto também traz números, como 47 prisões, mais de 939 mil armas cadastradas e 6.168 armas irregulares (UA02). Não há, contudo, menção explícita à consulta a bancos de dados públicos ou estruturados (UA03). Os dados são reproduzidos a partir de falas de fontes oficiais, como o ministro da Justiça e a Polícia Federal (PF) (UA04). Apesar da riqueza de números, faltam contexto histórico, perfil dos presos e implicações políticas mais amplas (UA05). Os dados são apresentados com fontes oficiais e sem distorções aparentes (UA06), com as fontes dos dados citadas explicitamente, como o Ministério da Justiça, a PF e coletiva de imprensa (UA07). A matéria não resulta de solicitação direta à fonte de informação (UA08). A LAI não é citada (UA09) e a LGPD também não é mencionada (UA10). A reportagem é puramente textual, sem apoio visual para os dados (UA11).

A notícia intitulada “**Alzheimer: medicamento retarda em 35% o declínio cognitivo**” (<https://tinyurl.com/yc8d4hsm>), publicada no dia 04 de maio de 2023, apresenta título destacando um dado impactante: “retarda em 35% o declínio cognitivo” (UA01). Há dados específicos sobre eficácia do medicamento em questão (35%, 40%, 39%) e efeitos colaterais (24%, 31,4%) no corpo da matéria (UA02). Os dados vêm de uma pesquisa clínica privada e não de bancos de dados públicos ou estruturados (UA03). Os dados são reproduzidos de comunicado oficial da farmacêutica Eli Lilly, sem tratamento adicional ou checagem independente (UA04). A notícia não aprofunda questões sobre custo do tratamento, viabilidade de aplicação em larga escala ou comparação com outros métodos (UA05). Os dados são apresentados com clareza e suas limitações, como mortes e efeitos colaterais, também são

citadas (UA06). A fonte principal é empresarial (comunicado da Eli Lily), além de comentários de especialistas (UA07). Não há pedidos de dados adicionais (UA08). A LAI não é utilizada ou mencionada (UA09) e a matéria não trata de privacidade de dados pessoais ou sensíveis, não citando a LGPD (UA10). O conteúdo é inteiramente textual, sem visualizações gráficas que ilustrem os dados (UA11).

De publicação no dia 12 de maio de 2023, a notícia **“Lei sancionada prevê multa de até R\$ 500 mil para agressores de mulheres”** (<https://tinyurl.com/4y9nm5k2>) apresenta título destacando valor de multa de “até R\$ 500 mil”, atraindo atenção com dado numérico relevante (UA01). O texto apresenta dados quantitativos, como valores de multas, frações de agravamento e reincidência (UA02). A matéria não utiliza ou menciona bancos de dados públicos para aprofundamento ou checagem (UA03). Os dados são transcritos da legislação publicada no Diário Oficial do Distrito Federal sem tratamento adicional (UA04). Faltam contexto estatístico, por exemplo, número de casos de reincidência ou valores médios de ressarcimento, além de impacto prático da lei (UA05). As informações são claras, objetivas e fundamentadas em fonte oficial (UA06). A fonte da informação é institucional: Diário Oficial do Distrito Federal e Câmara Legislativa (UA07). A reportagem não menciona pedido de acesso a dados adicionais (UA08). A Lei de Acesso à Informação (LAI) não é citada (UA09). A matéria não aborda privacidade de dados pessoais e não faz referência à Lei Geral de Proteção de Dados (LGPD) (UA10). A reportagem é apenas textual, sem visualizações gráficas de apoio (UA11).

Outra notícia do dia 12 de maio de 2023, intitulada **“Ibama já aplicou R\$ 1,51 milhão em multas em operação na terra Karipuna”** (<https://tinyurl.com/3rdch9bm>), apresenta título com valor de multas aplicadas (R\$ 1,51 milhão), criando impacto informativo imediato (UA01). A matéria apresenta números detalhados, como 7,4 mil m³ de madeira, R\$ 1,2 milhão em valor de madeira apreendida, 12 madeireiros indiciados, 14 empresas autuadas e 80 agentes (UA02). A reportagem não utiliza ou cita dados abertos adicionais de bases do Instituto Brasileiro de Meio Ambiente e dos Recursos Naturais Renováveis (Ibama), Instituto Brasileiro de Geografia e Estatística (IBGE), MapBiomas ou plataformas de desmatamento (UA03). Os dados apresentados são reproduzidos de notas da Polícia Federal (PF), do Ibama e da Fundação Nacional dos Povos Indígenas (Funai), sem análise complementar ou checagem cruzada (UA04). Faltam informações sobre reincidência, impacto ambiental direto, prejuízos à comunidade indígena e envolvimento empresarial ou político (UA05). Os dados são contextualizados pelas autoridades citadas (UA06) e as fontes são oficiais, como PF, Ibama e Funai (UA07). Não há menção a solicitação de acesso a informações governamentais (UA08).

A reportagem não cita a LAI (UA09) e a LGPD também não é mencionada (UA10). A matéria não apresenta gráficos, mapas ou imagens visuais que auxiliem na compreensão dos dados (UA11).

Ainda no dia 12 de maio de 2023, a notícia **“Desmatamento na Amazônia cai 67,9% em abril, 1ª grande queda sob governo Lula”** (<https://tinyurl.com/2akw6bn5>) destaca logo no título uma queda percentual exata (67,9%), dando ênfase à comparação temporal (UA01). O texto traz valores absolutos (km²), percentuais e comparações com anos anteriores (UA02). A reportagem usa dados prontos de instituições, como o Instituto Nacional de Pesquisas Espaciais (Inpe) e o Imazon, mas não acessa diretamente bancos públicos (UA03), de forma que os dados são reproduzidos de relatórios e comunicados oficiais (UA04). Faltam análise de causas da redução, checagem da qualidade dos dados, impacto real da queda e projeções futuras (UA05). Os dados são contextualizados e atribuídos sem distorções (UA06). As fontes são diversificadas, incluindo o Inpe como fonte governamental e o Imazon como institucional especializada (UA07). A reportagem não menciona pedido de dados por canais de acesso à informação (UA08). A LAI não é citada (UA09), nem a LGPD (UA10). O texto é exclusivamente narrativo, sem visualizações que apoiem a interpretação dos dados (UA11).

Publicada no dia 20 de maio de 2023, a notícia **“Mulheres ocupam 34% dos cargos em altas lideranças na administração federal”** (<https://tinyurl.com/3m8uxsz6>) traz título informando percentual objetivo (34%) sobre representatividade feminina (UA01). A matéria apresenta dados percentuais (29% para 34%) e evolução histórica (26% em 2019) (UA02), e cita o “Observatório de Pessoal”, mas não acessa nem explora banco de dados diretamente (UA03). Os dados são replicados de fonte oficial (a ministra), sem tratamento analítico (UA04). A notícia não analisa desigualdades interseccionais, como raça, território e áreas do governo, nem contextualiza metas ou planos futuros (UA05). Os dados são apresentados com fonte explícita e sem distorções (UA06). A fonte dos dados citada é oficial, a ministra de Gestão e Inovação em Serviços Públicos (UA07). Não há solicitação de dados estruturados (UA08), bem como a reportagem não aborda acesso à informação pública por meio da LAI (UA09). A LGPD também não é mencionada (UA10). Não há elementos gráficos que auxiliem a visualização dos dados informados (UA11).

A reportagem **“Temperaturas extremas demandam mudanças arquitetônicas para proteger humanos”** (<https://tinyurl.com/26x4tabt>), publicada no dia 28 de maio de 2023, traz título descritivo sem dados numéricos (UA01). O texto apresenta dados como 85% da população afetada por mudanças climáticas, 30% de redução de emissões de CO₂, entre outros (UA02). A matéria utiliza estudos científicos, mas não há exploração ativa de bases de dados

públicas (UA03), sendo que os dados são reproduzidos de pesquisas acadêmicas e científicas internacionais (UA04). O texto apresenta múltiplas frentes, como energia, arquitetura, urbanismo, mas requer contextualização à realidade local brasileira (UA05) e os dados apresentados são atribuídos a fontes claramente identificadas (UA06). A matéria cita fontes de dados variadas, com predominância de fontes científicas/especializadas, incluindo estudos acadêmicos e especialistas como urbanistas, engenheiros e cientistas (UA07). Não há investigação jornalística ou coleta direta de dados (UA08). Também não há menção à LAI (UA09). O tema não envolve dados pessoais e não cita a LGPD (UA10). Não há apoio visual ou gráfico para ilustrar os dados apresentados (UA11).

Outra notícia publicada no dia 28 de maio de 2023, com título **“IRPF 2023: mais de 7 milhões não enviaram a declaração; prazo termina quarta”** (<https://tinyurl.com/32twc8hn>), traz título destacando os “mais de 7 milhões” de contribuintes que não enviaram a declaração (UA01), apresentando diversos dados quantitativos, como número de declarações, percentuais, prazos, valores de rendimentos e multas (UA02). A matéria não demonstra extração ativa de dados públicos, apenas reproduz informações da Receita Federal (UA03). Os dados são, portanto, simplesmente reproduzidos de fonte oficial, que é a Receita propriamente dita (UA04). O texto traz informações técnicas, mas carece de interpretação ampliada, como perfil dos atrasados ou impacto econômico dos atrasos (UA05). Não há evidência de manipulação ou uso incorreto dos dados (UA06). Conforme apontamos, a fonte é mencionada claramente (UA07). Os dados não foram obtidos por apuração própria ou pedido formal à fonte. (UA08). Não há menção à LAI (UA09) e, embora a reportagem mencione riscos de golpes e vazamentos, não há referência direta à LGPD (UA10). Não há representação visual dos dados, como gráficos ou tabelas (UA11).

A notícia intitulada **“Consórcio de serviços: conheça o crédito adotado para realizar pequenos sonhos”** (<https://tinyurl.com/y8ky2r8m>), publicada no dia 28 de maio de 2023, apresenta título descritivo e sem números (UA01). A reportagem traz valores como número de consorciados, tíquete médio, taxas e percentuais (UA02). A reportagem utiliza dados fornecidos por entidades, como a Associação Brasileira das Administradoras de Consórcios (Abac), sem exploração ativa de bancos públicos (UA03). Há reprodução de dados oficiais da Abac e de especialistas do setor (UA04). O texto aborda aspectos do produto financeiro, mas há margem para aprofundar impactos sociais, perfil dos consorciados e regulação (UA05). Não há uso incorreto de dados e as informações são contextualizadas e alertam sobre fraudes (UA06). Há a fonte oficial dos dados, a Abac, além das especializadas (UA07). A apuração se baseia em dados disponibilizados por fontes primárias (UA08). Não há referência à LAI (UA09) e, apesar

de tratar de serviços digitais, não aborda proteção de dados e não cita a LGPD (UA10). A matéria inclui material visual, produzido por editoria de arte, para ilustrar o tema (UA11).

A notícia intitulada “**FAB realiza mais de 5 mil atendimentos a ribeirinhos no interior do Amazonas**”, publicada no dia 05 de junho de 2023 (<https://tinyurl.com/bdfshxae>), traz título com dado numérico central (5 mil atendimentos) (UA01). A matéria apresenta números de atendimentos, medicamentos, militares envolvidos e toneladas transportadas (UA02). Não houve investigação ativa de bases públicas de dados (UA03), os quais foram repassados diretamente pela Força Aérea Brasileira (FAB) e apenas reproduzidos pelo webjornal (UA04). Embora descritiva, a matéria demanda aprofundamento dos dados e contextualização sobre a precariedade estrutural do Sistema Único de Saúde (SUS) na Amazônia (UA05). Não há distorção ou manipulação de dados; os números são oficiais e contextualizados (UA06). A reportagem informa que a fonte das informações é oficial, no caso, a FAB (UA07). O jornalista não solicitou formalmente dados (UA08) e a matéria não cita a LAI (UA09). Não há discussão sobre a LGPD (UA10). A matéria não apresenta visualizações gráficas, apenas texto e foto (UA11).

Análise de conteúdo da notícia “**“Macaco Lula”: leia a íntegra do e-mail que ameaça brasileiros em Portugal**” (<https://tinyurl.com/msdzp7s>), publicada no dia 21 de junho de 2023, mostra que traz título sensacionalista, centrado em um insulto racial, sem apresentar dados quantitativos (UA01). A íntegra do *e-mail* criminoso é apresentada, além de informações quantitativas sobre brasileiros em Portugal (UA02). A matéria não faz uso de bancos de dados externos, apenas cita documentos oficiais, como o *e-mail* em questão (UA03), cujo conteúdo é reproduzido integralmente (UA04). O tema exige contextualização histórica, legal e sociológica, como debate acerca de nazismo, xenofobia e imigração (UA05). A notícia não veicula informações enganosas; baseia-se em documentos oficiais (UA06), sendo a fonte primária um documento da Embaixada Brasileira em Portugal (UA07). Não se evidencia uma solicitação ativa, embora o documento pareça obtido por apuração própria (UA08). A LAI não é mencionada (UA09) e a LGPD também não é citada (UA10). A reportagem não apresenta elementos visuais como gráficos ou mapas (UA11).

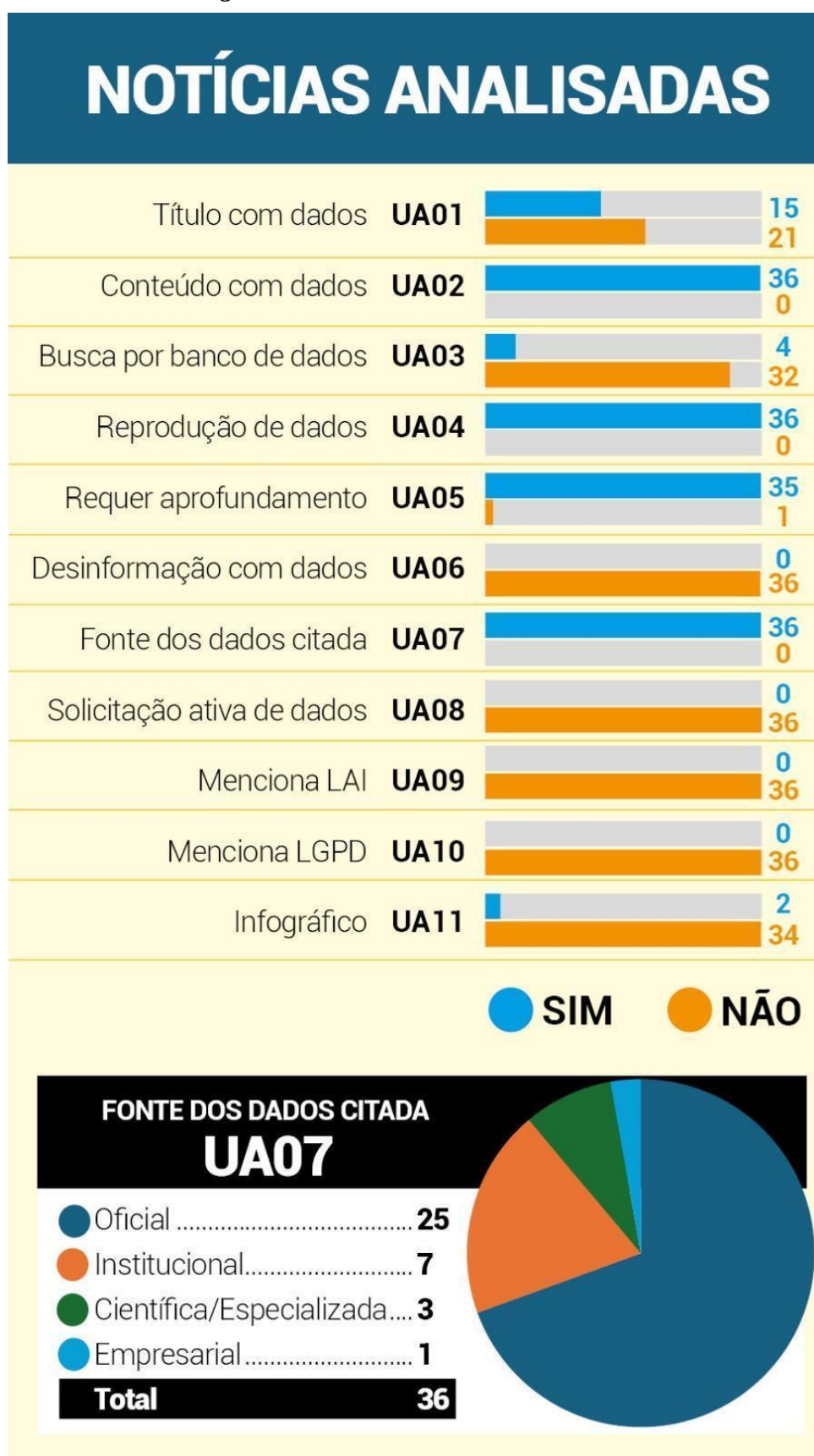
O Quadro 9 e a Figura 12 a seguir apresentam os resultados obtidos a partir das categorias definidas para análise de conteúdo das publicações do *Correio Braziliense* baseada em livro de códigos elaborado conforme preceitos teóricos de Bardin (2011) e Sampaio e Lycarião (2021):

Quadro 9 – Análise de conteúdo do *Correio Braziliense*

Notícias Analisadas		Sim	Não
UA01	Título com dados	15	21
UA02	Conteúdo com dados	36	0
UA03	Busca por banco de dados	4	32
UA04	Reprodução de dados	36	0
UA05	Requer aprofundamento	35	1
UA06	Desinformação com dados	0	36
UA07	Fonte dos dados citada	36	0
UA08	Solicitação ativa de dados	0	36
UA09	Menciona LAI	0	36
UA10	Menciona LGPD	0	36
UA11	Infográfico	2	34

Unidade de Análise	Total
UA07 Fonte dos Dados	36
Oficial	25
Institucional	7
Científica/Especializada	3
Empresarial	1

Fonte: elaborado pela autora (2025).

Figura 12 – Análise de conteúdo do *Correio Braziliense*

Fonte: elaborado pela autora (2025).

Com embasamento teórico alicerçado nos conceitos de Barbosa (2005; 2007; 2008; 2013), Mancini, Vasconcelos (2016) e Ferrari (2010) sobre Jornalismo de Dados e Jornalismo Digital e, ainda, Bardin (2011), Sampaio e Lycarião (2021), apontamos que foram coletadas 1.364 publicações noticiosas no *Correio Braziliense*, das quais 36 notícias e reportagens foram selecionadas para análise pelo aparente uso de dados. O exame aprofundado com aplicação do livro de códigos revelou que nenhuma das 36 notícias analisadas no webjornal *Correio Braziliense* reuniu o conjunto de atributos do JDBD, apesar de incluírem informações numéricas em todas (100%) as 36 publicações.

Em 15, ou 41,6% das 36 publicações analisadas, há citação de dados numéricos logo em seus títulos. A maioria dos conteúdos publicados, porém, tendo como referência 21, ou 58,3% das 36 notícias e reportagens analisadas do *Correio Braziliense*, não faz alusão a dados nos títulos.

Verificamos, a partir das categorizações estabelecidas para análise, que apenas 4, ou 11,1% do total de 36 notícias, foram sustentadas em levantamento simples de dados e com narrativa jornalística destacando tais dados, porém de forma alheia à vertente do JDBD. A maioria dos conteúdos noticiosos analisados – 32 publicações ou 88,8% do total de 36 – não foi ancorada e guiada por bancos de dados; tampouco foram incluídos recursos interativos em 34, ou 94,4% das 36 publicações.

A pesquisa ainda buscou mapear e evidenciar elementos do JDBD relacionados à origem e à pluralidade das fontes de dados. Constatamos que as publicações incluíram, majoritariamente, informações de canais oficiais, num total de 25 notícias, ou 69,4% de 36. Outras 7, ou 19,4% do total, citaram fontes institucionais; 3, ou 8,3%, com fontes científicas/especializadas e 1, ou 2,7%, com fonte empresarial.

Nenhuma das 36 notícias e reportagens analisadas que compõem o *corpus* foi produzida com dados obtidos por meio da LAI ou chegou a mencionar a LGPD, nem apresentou cruzamentos de bases de dados.

Além disso, apenas 2, ou 5,5% das 36 publicações noticiosas, incluiu infográficos estáticos, tendo a maioria delas – 34 ou 94,4% do total de 36 publicações – dispensado esse recurso. Os infográficos estão entre os principais meios de expressão do jornalismo de dados, apresentando características composicionais e, por consequência, efeitos de sentido, distintas em relação a outras formas de apresentação noticiosa, segundo Casadei, Rodrigues e Biernath (2016).

Embora a infografia não tenha sido um elemento muito presente, todas as 36 notícias e reportagens analisadas continham recursos visuais como foto ou ilustração; 2, ou 5,5% das

notícias, trouxeram recursos multimídia, sendo uma em vídeo (2,63%) e outra em áudio (2,63%).

5.3 Análise comparativa e resultado geral do uso de dados em notícias e reportagens do *Gazeta Digital* (MT) e *Correio Braziliense* (DF)

Do total de 1.959 publicações informativas coletadas dos webjornais *Gazeta Digital* e *Correio Braziliense*, somando textos jornalísticos diversos, como artigos de opinião, *releases* ou produzidos por agências de notícias, além de informes publicitários, 1.905 foram dispensadas da análise por não estarem classificadas como notícias e reportagens contendo dados numéricos, documentos digitais e infográficos, características constatadas após visualização e leitura. Restaram 54 notícias e reportagens que referenciavam uso de dados numéricos, documentos digitais acessados por computadores ou infografia.

Com o auxílio de categorias criadas conforme livro de códigos da Análise de Conteúdo (Bardin, 2011; Sampaio; Lycarião, 2021), divididas em 29 classificações, com 11 unidades de análise (UA) aplicadas às 54 notícias e reportagens coletadas da versão digital das publicações dentro do recorte temporal estabelecido neste estudo, verificamos que nenhuma das 54 peças analisadas nos webjornais *Gazeta Digital* e *Correio Braziliense* reuniram o conjunto de atributos do JDBD, como uso intensivo de bases de dados, algoritmos, linguagens de programação, metadados, mineração de dados, adoção de sistemas de gestão de conteúdo e *site* dinâmicos, automatização de produtos jornalísticos, uso de narrativas multimídias e infografias interativas, difusão de conteúdos via RSS (*Really Simple Syndication*), somada à mobilidade, variedade de formatos (texto, áudio, vídeo, infográficos, realidade virtual), interatividade, convergência e multimidialidade.

Mais da metade (29 ou 53,7%) das 54 notícias analisadas trouxeram dados numéricos logo em seus títulos, enquanto as outras 25, ou 46,3%, restantes não trouxeram tais dados nos títulos. Todas as 54 notícias incorporam dados como informação. Apenas 10, ou 18,5% das 54, demonstraram busca ativa por dados. As 44, ou 81,4%, publicações restantes não fizeram esse movimento de coleta. Além disso, nenhuma delas foi produzida com dados obtidos por meio da LAI e não houve qualquer menção à LGPD. Notamos que também não houve produção noticiosa com cruzamentos de bases de dados.

Todas as notícias coletadas apenas reproduziram dados, apresentando, predominantemente, características limitadas a um Jornalismo **com** Dados (Mancini; Vasconcellos, 2016), utilizando tais elementos quantificáveis sem que fossem o foco da

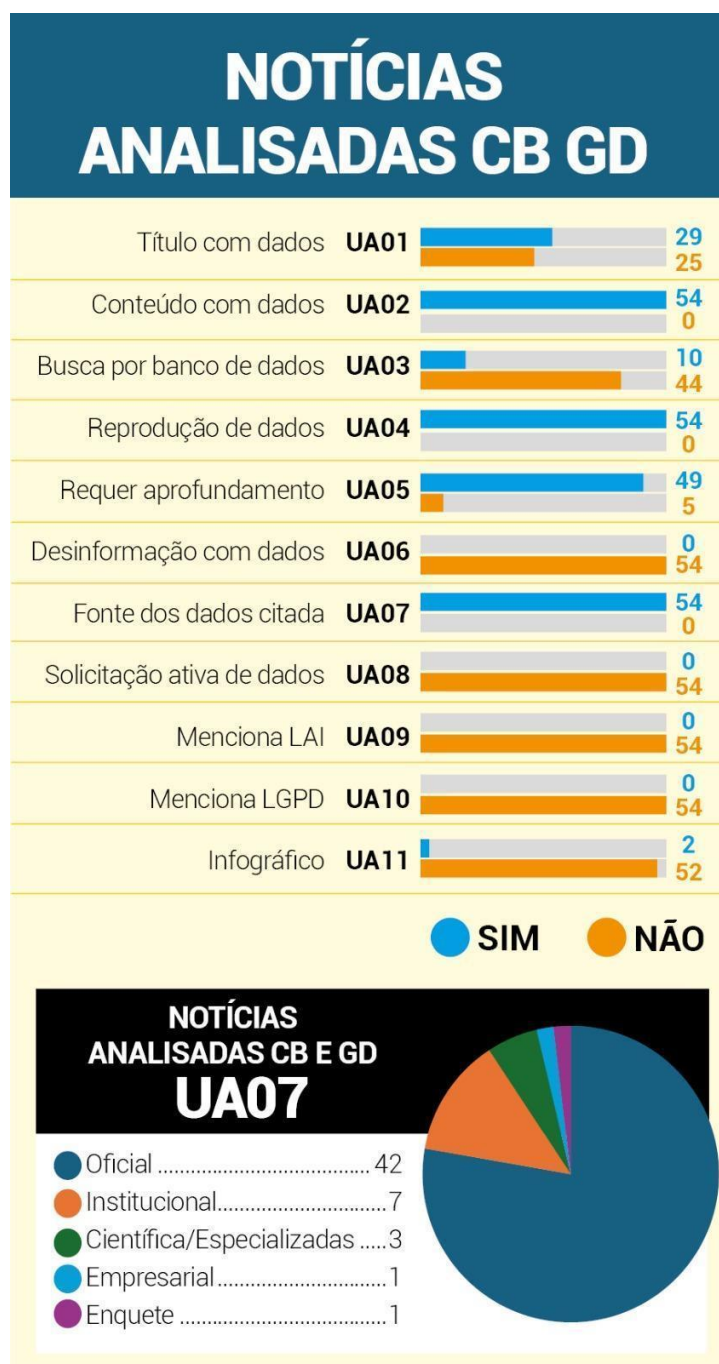
narrativa, configurando-se simplesmente em um apoio para ilustrar o material noticioso.

Conforme Mancini e Vasconcellos (2016), é possível distinguir dois tipos de produção jornalística quando se trata do uso de dados, sendo as reportagens **com** dados e reportagens **de** dados. No primeiro caso, os dados são usados apenas como complemento ou suporte visual à narrativa principal. Na segunda situação, porém, constituem o núcleo da matéria, sendo ponto de partida da apuração jornalística. Em outras palavras, enquanto nas reportagens com dados as informações quantitativas ilustram o conteúdo, nas reportagens de dados são elas que moldam a pauta e direcionam a construção da história.

A maioria das notícias e reportagens publicadas nos webjornais *Gazeta Digital* e *Correio Braziliense*, ou seja, 49, ou 90,7% do total, deixa lacuna para aprofundamento dos dados, condição não identificada em apenas 5, ou 9,3%, das notícias analisadas. Apesar disso, não foram percebidas marcas de desinformação em nenhuma das 54 notícias, como, por exemplo, ausência de citação de fontes ou das referências dos dados.

Apenas duas ou 3,70% das notícias incluíram gráfico do tipo estático para auxiliar na compreensão da informação. A maioria (52 ou 96,3%) não adotou essa ferramenta.

Este estudo também buscou identificar a origem e pluralidade de fontes de dados. A maioria, 42, ou 77,8%, de um total de 54 notícias e reportagens coletadas, incluiu informações de canais oficiais. Outras 7 delas, ou 13%, citam fontes institucionais; 3, ou 5,5%, citam fontes científicas/especializadas; além de uma, ou 1,85%, que menciona fonte empresarial; e uma, ou 1,85%, a referenciar enquête, conforme ilustrado no gráfico a seguir (Figura 13).

Figura 13 – Resultado da análise de conteúdo de *Gazeta Digital* e *Correio Braziliense*

Fonte: elaborado pela autora (2025).

Sobre o critério da interatividade apontado por Canavilhas (2003; 2007), Barbosa (2007) e Mielniczuk (2001; 2003; 2004), verificamos que a maioria dos atributos presentes nos ambientes digitais do *Gazeta Digital* e do *Correio Braziliense* ainda é trivial (Ferrari, 2010). Exemplos comuns incluem a navegação por *hiperlinks*, enquetes, testes e especiais multimídia, que reúnem elementos como áudio, vídeo, galerias de fotos e conteúdos complementares.

Recursos mais complexos, como fóruns nos quais os usuários podem expressar suas

opiniões sobre determinados temas e *chats*, são consideradas formas sofisticadas de interação. No entanto, mesmo essas iniciativas ainda representam um aproveitamento limitado do potencial interativo da internet, mantendo-se próximas aos formatos dos meios tradicionais e, muitas vezes, funcionando apenas como uma adaptação do conteúdo original para o ambiente *online*. Tal adequação com a incorporação destes elementos, somados à plataformização e mobilidade, observados nos dois webjornais *Gazeta Digital* e *Correio Braziliense*, caracterizam uma evolução da produção noticiosa para internet e hibridização, desde a primeira até a quinta geração do jornalismo online, conforme conceituação referenciada para esta pesquisa.

Comparando os resultados obtidos no *Gazeta Digital* e no *Correio Braziliense* evidenciam-se diferenças nas práticas de uso, menção e mediação de dados jornalísticos entre os dois veículos. Embora ambos apresentem índices de conteúdos com dados e de reprodução de dados, notamos variações no detalhamento das unidades de análise.

No que se refere à presença de dados nos títulos (UA01), observamos que o *Gazeta Digital* apresenta maior proporção de notícias com essa característica, somando 14 ocorrências positivas contra 4 negativas, enquanto o *Correio Braziliense* mostra um equilíbrio distinto, com apenas 15 notícias contendo dados no título frente a 21 que não apresentam essa inserção. Esse contraste sugere que o *Gazeta Digital* tende a enfatizar a quantificação logo no título, ao passo que o *Correio Braziliense* adota abordagem mais reservada nessa etapa inicial da notícia.

Quanto ao conteúdo com dados (UA02), ambos os jornais demonstram aderência total ou quase total à prática, sendo que o *Gazeta Digital* apresenta 18 ocorrências positivas, sem casos negativos, e o *Correio Braziliense* também demonstra presença marcante dessa característica, com 36 registros positivos. Essa convergência revela que a incorporação de dados no corpo do texto é uma prática consolidada em ambos os veículos, indicando valorização de elementos informativos baseados em evidências.

As diferenças tornam-se mais pronunciadas nas unidades relacionadas à busca por banco de dados (UA03). No *Gazeta Digital*, 6 notícias apresentam esse recurso, ao passo que 12 não o fazem. No *Correio Braziliense* esse número é mais reduzido, com apenas 4 notícias indicando a utilização ativa de bancos de dados contra 32 que não adotam essa prática. Assim, embora o índice seja baixo em ambos os casos, observa-se que o *Gazeta Digital* demonstrou maior propensão à consulta a bases informacionais externas.

No tocante à reprodução de dados (UA04), novamente ambos apresentam padrões similares, com 18 ocorrências positivas no *Gazeta Digital* e 36 no *Correio Braziliense*, com ausência total de casos negativos. Isso indica que a replicação de dados disponíveis é prática amplamente utilizada pelos dois veículos, ainda que não necessariamente acompanhada de

contextualização aprofundada.

A unidade de análise “Requer aprofundamento” (UA05) indica que o *Correio Braziliense* apresenta maior proporção de conteúdos cuja exploração poderia ser expandida, totalizando 35 ocorrências positivas, contrastando com apenas um caso que não demanda aprofundamento. O *Gazeta Digital*, por sua vez, mostra distribuição menos acentuada, com 14 notícias que requerem aprofundamento e 4 que não. Esses resultados sugerem que, embora ambos recorram a dados, o *Correio Braziliense* o faz frequentemente de modo que demanda maior contextualização, possivelmente devido ao volume superior de publicações.

As duas unidades seguintes (UA06 e UA08) revelam alinhamentos importantes, visto que nenhum dos veículos apresentou casos de desinformação com dados e ambos também não realizaram solicitações ativas de dados. Isso pode indicar ausência de práticas de jornalismo investigativo voltadas à requisição direta de informações, bem como inexistência de desinformação baseada em dados nos *corpora* analisados.

Quanto à citação de fontes de dados (UA07), observa-se diferença relevante no volume, mas não na prática, pois em ambos a totalidade das notícias que utilizam dados cita suas fontes, sendo 18 no *Gazeta Digital* e 36 no *Correio Braziliense*. Há, porém, disparidade significativa quanto ao tipo de fonte. O *Gazeta Digital* apresenta predominância de fontes oficiais, com 17 casos, e apenas uma ocorrência de enquete. O *Correio Braziliense*, por sua vez, apresenta diversidade maior, ainda que também com predominância oficial, incluindo 25 fontes oficiais, sete institucionais, três científicas ou especializadas e uma empresarial. Esse achado demonstra que o *Correio Braziliense* recorre a um leque mais variado de fontes, o que pode indicar maior esforço de checagem ou aprofundamento, mas também reflete sua estrutura editorial mais robusta.

Outro ponto de convergência é a ausência de menções à Lei de Acesso à Informação (LAI) e à Lei Geral de Proteção de Dados (LGPD) em ambos os veículos (UA09 e UA10). Tal ausência revela um distanciamento das discussões sobre transparência pública e proteção de dados pessoais no contexto das notícias analisadas, reforçando a percepção de que os dados são utilizados majoritariamente como insumos informativos, sem problematização jurídica ou ética.

Por fim, quanto ao uso de infográficos (UA11), verifica-se que o *Gazeta Digital* não apresentou nenhuma ocorrência, enquanto o *Correio Braziliense* registrou apenas duas. Ainda que a diferença exista, os números sugerem baixa utilização de recursos visuais de representação gráfica dos dados em ambos os veículos, o que indica potencial pouco explorado na apresentação visual de informações.

CONCLUSÃO

A presente dissertação teve como objetivo principal verificar de que maneira os webjornais *Gazeta Digital* e *Correio Braziliense*, representativos da região Centro-Oeste do Brasil, vêm se apropriando de bases de dados na produção jornalística digital, à luz do conceito de Jornalismo Digital em Base de Dados (JDBD), conforme delineado por Barbosa (2005; 2007; 2008; 2013). Considerando o avanço dos processos de digitalização e dataficação (Lemos, 2021; 2023; Segata; Rifiotis, 2021) que impactam profundamente as práticas comunicacionais contemporâneas, buscamos compreender em que medida essas transformações têm sido incorporadas por veículos regionais no contexto da produção de notícias.

A partir disso, nosso objetivo central consistiu em identificar como esses dois veículos, inseridos em contextos de localidades distintas - Mato Grosso e Distrito Federal -, utilizam bancos ou bases de dados na produção de notícias veiculadas em ambiente digital.

Como forma de alcançar esse objetivo, estabelecemos três metas específicas: (1) revisar a literatura sobre Jornalismo de Dados e o uso de bases de dados no jornalismo digital; (2) coletar e categorizar conteúdos informativos presentes nas capas dos webjornais selecionados, a fim de verificarmos o uso de dados, a pluralidade de fontes e a aplicação da Lei de Acesso à Informação (LAI); e (3) identificar de que maneira a incorporação de bases de dados pode impactar a qualidade da produção jornalística na região Centro-Oeste.

A escolha dos dois webjornais não se deu de forma aleatória, como explicamos na introdução a este estudo. Eles foram selecionados por sua relevância regional, pertencimento a conglomerados midiáticos e representatividade digital em seus respectivos estados. Baseamos a seleção em critérios de audiência, longevidade e mobilidade digital, identificados por meio de ferramentas como *Guia de Mídia* e *SimilarWeb*.

A metodologia adotada dividiu-se em três etapas. Inicialmente, realizamos uma pesquisa exploratória e revisão bibliográfica que contextualizou a evolução do Jornalismo Digital e do Jornalismo de Dados a partir de autores como Lévy (1998), Mielniczuk (2001, 2003, 2004), Canavilhas (2003; 2007), Meyer (1991), Träsel (2014; 2017) e Barbosa (2005; 2007; 2008; 2013). Em seguida, descrevemos o *corpus* empírico e os critérios de escolha dos veículos. Por fim, procedemos à coleta e análise dos conteúdos noticiosos publicados nas capas dos webjornais durante um período amostral de 15 dias, construído aleatoriamente entre 1º de março e 21 de junho de 2023, utilizando a ferramenta *Wayback Machine*.

Ao todo, selecionamos 595 publicações do *Gazeta Digital* e 1.364 do *Correio Braziliense*, após observação e leitura de cada conteúdo. A análise centrou-se exclusivamente

em notícias e reportagens que apresentassem evidências do uso de dados – por meio de bases oficiais, repositórios abertos ou vazamentos – e que incluíssem informações quantitativas. Foram desconsiderados gêneros como editoriais, crônicas, enquetes, artigos opinativos, conteúdos de agências e materiais publicitários.

Essa delimitação nos possibilitou investigar de forma mais precisa os traços do JDBD nos veículos analisados, contribuindo para a reflexão sobre a apropriação de dados na prática jornalística e os desafios e as potencialidades dessa abordagem para a produção de conteúdo informativo mais transparente, plural e ancorado em evidências, sobretudo no contexto regional do Centro-Oeste brasileiro.

Fundamentada na Análise de Conteúdo (Bardin, 2011; Sampaio; Lycarião, 2021), a metodologia utilizada permitiu examinar um total de 54 reportagens e notícias que apresentavam algum tipo de uso de dados numéricos, documentos digitais ou infográficos, a partir de um universo de 1.959 publicações.

A análise comparativa da cobertura jornalística baseada em dados realizada pelos webjornais *Gazeta Digital* e *Correio Braziliense* evidencia que, embora ambos utilizem dados regularmente, o fazem de maneira superficial. O *Gazeta Digital* recorre com frequência ao uso de dados nos títulos e consultas a bancos de dados, mas apresenta pouca diversidade de fontes e completo desuso de infográficos. O *Correio Braziliense* demonstrou aplicação de um volume robusto de dados, com fontes variadas, porém maior presença de conteúdos que requerem aprofundamento analítico. Essas distinções evidenciam não apenas escolhas editoriais, mas também desigualdades estruturais entre os veículos, sobretudo relacionadas à capacidade técnica e operacional para produzir jornalismo orientado por dados.

Constatamos que em nenhuma das matérias ou reportagens estão presentes os atributos essenciais do Jornalismo **de** Dados (Mancini; Vasconcellos, 2016; Bounegru e Gray, 2021; Parasie; Dagiral, 2013), Jornalismo Guiado por Dados (Matzat, 2011; Träsel, 2014; 2017) ou Jornalismo Orientado por Dados (Arrese, 2022), como coleta ativa de bases, limpeza, padronização ou cruzamentos de dados, uso de ferramentas de análise, equipes multidisciplinares, infográficos ou visualizações avançadas, elementos que, conforme os autores (Mancini; Vasconcelos, 2016; Bounegru e Gray, 2021; Parasie; Dagiral, 2013; Träsel, 2014; 2017; Arrese, 2022; Casadei, Rodrigues e Biernath, 2016), compõem a rotina produtiva que definem essas modalidades. Do mesmo modo, não foram identificadas práticas que caracterizam o Jornalismo Digital em Base de Dados (JDBD), tal como concebido por Barbosa (2005; 2007; 2008; 2013). A ausência de cruzamento de bases públicas e a não exploração de bases estruturadas contrariam as premissas do JDBD, que, segundo Barbosa (2005; 2007; 2008;

2013), deve apoiar-se em transparência, validação, contextualização, comparação, visualização informativa e produção orientada por dados.

A falta de mineração de dados, inexistência de infraestrutura tecnológica capaz de sustentar narrativas dinâmicas ou multimídia, de pedidos via Lei de Acesso à Informação (LAI) e não articulação com a Lei Geral de Proteção de Dados (LGPD) acentuam um distanciamento em relação ao modelo que integra algoritmos, metadados, linguagens de programação e processos automatizados em redações digitalmente estruturadas. Também permanecem ausentes os elementos da narrativa digital proposta por Bertocchi (2005; 2016), como a organização semântica de metadados, a construção de sistemas narrativos renderizados e a articulação entre bases informacionais e arquiteturas de sentido.

Os resultados mostram que os dois webjornais ainda operam em um modelo tradicional, com uso limitado de dados e dependência de conteúdo fornecido por fontes oficiais, distantes tanto do Jornalismo de Dados (Mancini; Vasconcellos, 2016; Bounegru e Gray, 2021; Parasie; Dagiral, 2013), Jornalismo Guiado por Dados (Träsel, 2014; 2017) ou Jornalismo Orientado por Dados (Arrese, 2022) quanto do Jornalismo Digital em Base de Dados (Barbosa, 2005; 2007; 2008; 2013).

A análise da cobertura jornalística com uso de dados pelos webjornais revelam, portanto, lacunas e omissões de propriedades do Jornalismo Digital em Base de Dados (Barbosa, 2005; 2007; 2008; 2013). Reforçamos que as reportagens analisadas não incorporaram os atributos essenciais do JDBD, tais como apropriação ativa de bases de dados, cruzamentos dessas informações, uso de ferramentas de análise de dados, infográficos interativos ou mineração de dados. Adicionalmente, observamos ausência de produção de conteúdo a partir da Lei de Acesso à Informação (LAI) ou de referências à Lei Geral de Proteção de Dados (LGPD), bem como baixa diversidade nas fontes de dados utilizadas, com predominância de fontes oficiais.

Concluimos, portanto, que os resultados evidenciaram predominância nas publicações analisadas de um uso superficial e limitado de dados, caracterizando-se predominantemente como Jornalismo **com** Dados (Mancini; Vasconcellos, 2016), em que tais elementos são apenas ilustrativos na narrativa, sem que orientem a investigação e produção jornalística.

Os achados desta pesquisa sugerem que, embora haja uma inserção gradual de elementos do jornalismo digital nas rotinas produtivas dos veículos estudados, o uso de dados se restringe a uma abordagem factual, não explorando plenamente o potencial investigativo, analítico e explicativo proporcionado pelo JDBD. Esse cenário pode ser atribuído a diversos fatores, como limitações técnicas e estruturais, escassez de tempo e recursos, ausência de cultura organizacional voltada à inovação, além da necessidade de formação específica dos

profissionais na área para lidar com ferramentas de coleta, análise e visualização de dados.

Diante do problema de pesquisa que buscou compreender de que maneira os webjornais *Gazeta Digital* e *Correio Braziliense* se apropriam de bancos ou bases de dados na produção jornalística digital, à luz do conceito de Jornalismo Digital em Base de Dados (Barbosa, 2005; 2007; 2008; 2013), este estudo apontou para a urgência de uma reflexão sobre a formação profissional dos jornalistas, especialmente no que tange à capacitação técnica e crítica para o uso de dados na construção de narrativas jornalísticas que contribuam para uma compreensão mais aprofundada da realidade social. Destacamos, ainda, a relevância de políticas públicas que promovam a ampliação da transparência, do acesso e da disponibilização de bases de dados, de modo a fomentar a produção de conteúdos jornalísticos mais qualificados e socialmente relevantes.

A análise aqui empreendida buscou contribuir para o campo dos estudos em jornalismo digital, em particular no que se refere ao jornalismo regional, oferecendo subsídios teóricos e empíricos para a compreensão das práticas atuais e das potencialidades do uso de dados no contexto jornalístico. A partir dos resultados obtidos, novas questões emergem e podem orientar futuras investigações: de que maneira o jornalismo regional pode evoluir na apropriação de bases de dados para além da factualidade? Em que medida os webjornais podem incorporar práticas investigativas baseadas em dados para ampliar o papel social do jornalismo? E, ainda, como o JDBD pode contribuir para a construção de uma inteligência coletiva (Lévy, 1998) mais robusta e participativa em contextos locais?

Esperamos, por fim, que este trabalho incentive novas reflexões, tanto no campo acadêmico quanto no profissional, sobre os desafios e possibilidades para a consolidação do Jornalismo Digital em Base de Dados no Brasil, especialmente em suas expressões regionais, onde as limitações estruturais convivem com oportunidades de inovação. Ao promover um jornalismo mais contextualizado, aprofundado e orientado por evidências, acreditamos que é possível contribuir para o fortalecimento da democracia, do acesso à informação, da participação cidadã e da leitura crítica da realidade social. Reafirmando, dessa forma, sua função como mediador qualificado entre dados e sociedade, especialmente diante do paradoxo sociológico da desinformação, quando conteúdos falsos tendem a gerar maior coesão e integração social do que informações verdadeiras, que, ao estimularem o debate racional, produzem dissenso (Bachur, 2021). Esse fenômeno reforça a urgência de pesquisas que aprofundem a compreensão sobre como os dados, a informação qualificada e o jornalismo podem atuar em defesa de uma esfera pública mais democrática.

REFERÊNCIAS

- ALVES, R. C. Jornalismo digital: Dez anos de web... e a revolução continua. **Comunicação e Sociedade**, Portugal, v. 9, n. 10, p. 93-102, 2006. Disponível em: <https://revistacomsoc.pt/article/view/1217>. Acesso em: 1 de set. 2024.
- ALVES, S. T. F.; GEHLEN, M. A. Jornalismo de Dados: revisitando conceitos em transformação. In: SOUSA, J. P. (Org.). **Jornalismo e Estudos Mediáticos** – Memória VI. Porto: Fundação Fernando Pessoa, 2023. p. 83-96. Disponível em: https://bdigital.ufp.pt/bitstream/10284/11959/1/ebook_mem%C3%B3ria_VI.pdf. Acesso em: 1 de set. 2024.
- AMADO, A. As metáforas do jornalismo: identidades implícitas para uma profissão mutante. **Lumina**, v. 13, n. 2, p. 11-23, 2019. Disponível em: <https://doi.org/10.34019/1981-4070.2019.v13.27737>. Acesso em: 1 de set. 2024.
- ARAÚJO, E. R.; SOUZA, E. C. **Obras jornalísticas** - Uma síntese. Brasília: Vestcon, 2003.
- ARRESE, Á. “In the Beginning Were the Data”: Economic Journalism as/and Data Journalism. **Journalism Studies**, v. 3, n. 4, p. 487-505, 2022. Disponível em: <https://doi.org/10.1080/1461670X.2022.2032803>. Acesso em: 1 de set. 2024.
- BACHUR, J. P. Desinformação política, mídias digitais e democracia: Como e por que as fake news funcionam? **Direito Público**, v. 18, n. 99, 2021. Disponível em: <https://doi.org/10.11117/rdp.v18i99.5939>. Acesso em: 1 de set. 2024.
- BARBOSA, S. Bases de dados e webjornalismo: em busca de novos conceitos. In: SOPCOM, 4., Aveiro, Universidade de Aveiro. **Livro de Actas**. Aveiro: Associação Portuguesa de Ciências da Comunicação, 2005. Disponível em: <https://arquivo.bocc.ubi.pt/pag/barbosa-suzana-bases-de-dados-webjornalismo.pdf>. Acesso em: 1 de set. 2024.
- BARBOSA, S. **Jornalismo Digital em Base de Dados (JDBD)** - Um paradigma para produtos jornalísticos digitais dinâmicos. 2007. Tese (Doutorado em Comunicação e Cultura Contemporânea) - Programa de Pós-Graduação em Comunicação e Cultura Contemporâneas da Universidade Federal da Bahia. Salvador, Bahia. 2007. Disponível em: <https://repositorio.ufba.br/handle/ri/11299>. Acesso em: 1 de set. 2024.
- BARBOSA, S. Modelo JDBD e o ciberjornalismo de quarta geração. **Grupo de Pesquisa em Jornalismo On-line**. Salvador, 2008. Disponível em: https://facom.ufba.br/jol/pdf/2008_Barbosa_RedUCMx.pdf. Acesso em: 1 de set. 2024.
- BARBOSA, S. Jornalismo convergente e continuum multimídia na quinta geração do jornalismo nas redes digitais. In: CANAVILHAS, J. (Org.). **Notícias e mobilidade: o jornalismo na era dos dispositivos móveis**. Covilhã: LabCom; UBI, 2013. p. 33-54. Disponível em: https://www.researchgate.net/publication/267569021_Noticias_e_Mobilidade_o_jornalismo_na_era_dos_dispositivos_moveis. Acesso em: 1 de set. 2024.
- BARBOSA, S.; TORRES, V. O paradigma “Jornalismo Digital em Base de Dados”: modos de narrar, formatos e visualização para conteúdos. **Galaxia**, São Paulo, v.13, n. 25, p. 152-

164, 2013. Disponível em: <https://www.scielo.br/j/gal/a/tmKP7Vf759RC7GjmRYNNf6n/>. Acesso em: 1 de set. 2024.

BARDIN, L. **Análise de Conteúdo**. São Paulo: Edições 70, 2011.

BARROS, Claudete. Os computadores invadem a redação. **A Gazeta**, Cuiabá, ano 1, n. 1, p. 2E, 23 de mai. 1990.

BASTOS, H. Ciberjornalismo, jornalismo e democracia. **Revista Media & Jornalismo**, n. especial, dez. 2015. Disponível em: <https://repositorio-aberto.up.pt/handle/10216/81651>. Acesso em: 1 de set. 2024

BENANTI, P. A necessidade de uma algorética para enfrentar o desafio da inteligência artificial. **Instituto Humanitas Unisinos**, 23 de dez. 2021. Disponível em: <https://www.ihu.unisinos.br/categorias/615549-a-necessidade-de-uma-algoretica-para-enfrentar-o-desafio-da-inteligencia-artificial-artigo-de-paolo-benanti>. Acesso em: 1 de set. 2024.

BERTOCCHI, Daniela. **Dos dados aos formatos: a construção de narrativas no jornalismo digital**. Curitiba: Appris, 2016.

BERTOCCHI, D. Gêneros jornalísticos em espaços digitais. In: SOPCOM, 4., Aveiro, Universidade de Aveiro. **Livro de Actas**. Aveiro: Associação Portuguesa de Ciências da Comunicação, 2005. Disponível em: <http://bocc.ufp.pt/pag/bertocchi-daniela-generos-jornalisticos-espacos-digitais.pdf>. Acesso em: 1 de set. 2024.

BOSELLI, M.A.; FERREIRA, F. V.; VARÃO, R. Breve história das origens da automação e do uso de inteligência artificial no jornalismo. In: ENCONTRO REGIONAL NORDESTE DE HISTÓRIA DA MÍDIA, 7., Nordeste. **Atas [...]**. Nordeste: Associação Brasileira de Pesquisadores de História da Mídia; UFMA; UFS, 2022. Disponível em: <https://redealcar.org/anais-eventos-regionais-nordeste-7o-encontro-2022/>. Acesso em: 1 de set. 2024.

BOUNEGRU, L.; GRAY, J. **Manual de Jornalismo de Dados: Rumo a uma prática crítica de dados**. 2. ed. São Paulo: Associação Brasileira de Jornalismo Investigativo (ABRAJI), 2021. Disponível em: <https://www.abraji.org.br/help-desk/manual-de-jornalismo-de-dados-2>. Acesso em: 1 de set. 2024.

BRADSHAW, P. The inverted pyramid of data journalism. **Online Journalism Blog**, 7 de jul. 2011. Disponível em: <https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>. Acesso em: 1 de set. 2024.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República; Casa Civil, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm. Acesso em: 25 de set. 2025.

BRASIL. Presidência da República. **Decreto nº 7.724, de 16 de maio de 2012**. Regulamenta a Lei nº 12.527, de 18 de novembro de 2011, que dispõe sobre o acesso a informações previsto no inciso XXXIII do caput do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art.

216 da Constituição. Brasília, DF: Presidência da República; Casa Civil, 2012. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/decreto/d7724.htm. Acesso em: 25 de set. 2025.

BRASIL. **Lei nº 12.527, de 18 de novembro de 2011.** Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal. Brasília, DF: Presidência da República; Casa Civil, 2011. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 18 de set. 2025.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018.** Lei Geral de Proteção de Dados Pessoais. Brasília, DF: Presidência da República; Casa Civil, 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 18 de set. 2025.

BRENOL, M. V. **Transparência digital e jornalismo:** modalidades comunicativas com uso de dados públicos. 2019. Tese (Doutorado em Comunicação e Informação) – Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019. Disponível em: <https://lume.ufrgs.br/handle/10183/196380>. Acesso em: 1 de set. 2024.

BRENOL, M. V.; BACCIN, A. Dez anos de LAI e a pesquisa sobre o acesso à informação pública pelo jornalismo. *Estudos em Jornalismo e Mídia* v. 19, n.2, jul.dez.2022. Disponível em: <https://doi.org/10.5007/1984-6924.2022.85925>. Acesso em: 1 de set. 2024

CANAVILHAS, J. Webjornalismo: considerações gerais sobre jornalismo na web. In: **Informação e comunicação online.** Covilhã: LabCom; UBI, 2003. p. 63-73. Disponível em: https://ubibliorum.ubi.pt/bitstream/10400.6/4358/1/CAP%C3%8DTULO_WebjornalismoConsidera%C3%A7%C3%B5esgerais.pdf. Acesso em: 1 de set. 2024.

CANAVILHAS, J. Webjornalismo: da pirâmide invertida à pirâmide deitada. In: BARBOSA, S. (Org.). **Jornalismo digital de terceira geração.** Covilhã: LabCom; UBI, 2007. p. 25-40. Disponível em: https://labcom.ubi.pt/ficheiros/20110824-barbosa_suzana_jornalismo_digital_terceira_geracao.pdf. Acesso em: 1 de set. 2024.

CASADEI, E. B.; RODRIGUES, K. de C.; BIERNATH, C. A. G. O não dito em infográficos: jornalismo de dados, discurso constituinte e silêncio como efeito de sentido. **Rev. Estud. Comun.**, Curitiba, v. 17, n. 43, p. 101-117, set.-dez. 2016. Disponível em: <https://periodicos.pucpr.br/estudosdecomunicacao/article/view/22555>. Acesso em: 1 de set 2024.

CESAR, P. P. **Correio Braziliense:** um jornal a serviço do leitor? 2006. Trabalho de Conclusão de Curso (Bacharelado em Jornalismo) – Faculdade de Comunicação Social do Centro Universitário de Brasília – UniCEUB, Brasília, 2006. Disponível em: <https://repositorio.uniceub.br/jspui/handle/123456789/1678>. Acesso em: 1 de set. 2024.

CHRISTOFOLETTI, R. Reafirmar o acesso para exigir garantias para uma cidadania informada. **Revista Eletrônica Internacional de Economia Política da Informação, da**

Comunicação e da Cultura - Revista Eptic, v. 21, n. 3, p. 190-193, set.-dez. 2019.

Disponível em:

<https://ufs.emnuvens.com.br/epitic/issue/view/revista%20epitic%20vol.21%283%292019/v.%2021%20n.%203%20%282019%29>

Acesso em: 1 de set. 2024

CIMADEVILLA, G. Sociedad digital, sociedad dual. **Signo e Pensamiento**, Bogotá, v. 28, n. 54, p. 68-81, 2009. Disponível: <http://www.scielo.org.co/pdf/signo/v28n54/v28n54a05.pdf>. Acesso em: 1 de set. 2024.

CORREIO Braziliense. **Edição do Dia**, Brasília, 21 de jun. 2023. Disponível em:

https://flip.correiobraziliense.com.br/index.php?cd_prded=&dtEdcaoIni=21%2F06%2F2023&dtEdcaoFim=21%2F06%2F2023&search=search. Acesso em: 13 de mai. 2024.

CÚPULA IBERO-AMERICANA. **Declaração de Bávaro**. Bávaro, Rep. Dominicana: XII Cimeira Ibero-Americana, nov. 2002. Disponível em: <https://www.segib.org/wp-content/uploads/DECLARASAOBAVARO.pdf>. Acesso em: 25 de set. 2025.

DALBEN, S. O uso de robôs no jornalismo brasileiro: três estudos de caso. In: SEMINÁRIO DE PESQUISA EM JORNALISMO INVESTIGATIVO, 6., São Paulo, Universidade Anhembi- Morumbi. **Atas** [...]. São Paulo: Abraji, 2019. Disponível em: https://seminario.abraji.org.br/wp-content/uploads/2023/04/SILVIA_DALBEN-O_uso_de_robos_no_jornalismo_brasileiro_tres_estudos_de_caso.pdf. Acesso em: 1 de set. 2024.

DA SILVA, M. A.; MEDEIROS, F. B.; CERETTA, K. A. C. Covid-19 and fake news: analysis of the verified news at the “Fact or fake” website. Chasqui – Revista Latinoamericana de Comunicación, Quito, n. 145, p. 119–136, 2020.

DE LIMA JR., R. A.; COELHO, T. F. Pandemia de Covid-19 como Acontecimento e a Desinformação nas Páginas do Jornal A Gazeta. In: CONGRESSO DE CIÊNCIAS DA COMUNICAÇÃO NA REGIÃO CENTRO-OESTE, 23., Campo Grande, Universidade Católica Dom Bosco. **Atas** [...]. Campo Grande: Intercom, 2023.

Disponível:

<https://sistemas.intercom.org.br/pdf/submissao/regional/9/05012023130546644fe35ae1e93.pdf>. Acesso em: 1 de set. 2024.

DIAKOPOULOS, N. Jornalismo aplicado a algoritmos: métodos e pontos de vista investigativos. **Manual de Jornalismo de Dados: Rumo a uma prática crítica de dados**. 2. ed. São Paulo: Associação Brasileira de Jornalismo Investigativo (ABRAJI), 2021.

Disponível em: <https://www.abraji.org.br/help-desk/manual-de-jornalismo-de-dados-2>.

Acesso em: 1 de set. 2024.

DONOHU, D. 50 Years of Data Science. **Journal of Computational and Graphical Statistics**, v. 26, n. 4, p. 745-766, 2017. Disponível em:

<https://doi.org/10.1080/10618600.2017.1384734>. Acesso em: 1 de set. 2024.

ERIKSEN L. B.; IHLSTRÖM C. Evolution of the Web News Genre - The Slow Move Beyond the Print Metaphor. In: ANNUAL HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, 33., Maui, HI. **Proceedings** [...], v. 2. [s.l.]: IEEE, 2000.

Disponível em: <https://ieeexplore.ieee.org/document/926697>. Acesso em: 1 de set. 2024.

FAVALESSA, M. **Jornalismo de Dados em veículos locais**: o uso de base de dados na cobertura da pandemia em Cuiabá (MT). 2023. Dissertação (Mestrado em Comunicação) Faculdade de Comunicação e Artes da Universidade Federal de Mato Grosso. Cuiabá, Mato Grosso, 2023.

FRANÇA, V. R. V. O objeto e a pesquisa em comunicação: uma abordagem relacional. Pesquisa em comunicação: metodologias e práticas acadêmicas. In: MOURA, C. P.; LOPES, M. I. V. (Orgs.). **Pesquisa em comunicação**: Metodologias e práticas acadêmicas. Porto Alegre: EdUPUCRS, 2016. p. 153-174. Disponível em: <https://repositorio.ufmg.br/bitstream/1843/39883/2/veraObjetoPesquisa.pdf>. Acesso em: 1 de set. 2024.

GAZETA, A. **Nova Gazeta**. Cuiabá, 23 de mai. 1998.

GAZETA Digital. Capa. **Web Archive**, 01 de abr. 2001. Disponível em: <https://web.archive.org/web/20010402012045/http://www.gazetadigital.com.br:80/>. Acesso em: 19 de mai. 2024.

GAZETA Digital. Capa. **Web Archive**, 21 de jun. 2023. Disponível em: <https://web.archive.org/web/20230621181135/https://www.gazetadigital.com.br/>. Acesso em: 19 de mai. 2024.

GEHRKE, M. **Os elementos de transparência no jornalismo guiado por dados**. 2021. Tese (Doutorado em Comunicação) – Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul, Porto Alegre, 2021. Disponível em: <https://lume.ufrgs.br/handle/10183/234821>. Acesso em: 1 de set. 2024.

HAN, B. C. **Infocracia** - Digitalização e a crise da democracia. Petrópolis: Vozes, 2022.

ITO, L. Credibilidade jornalística em debate: Entre o ataque e a mimetização como estratégias de desinformação. Revista Eco-Pós, v. 27, n. 1, p. 406-427, 2024.

KAUFMAN, D.; SANTAELLA, L. Os dados estão nos engolindo? **Civitas**: Revista de Ciências Sociais, São Paulo, v. 21, n. 2, p. 214-223, mai.-ago. 2021. Disponível em: <https://revistaseletronicas.pucrs.br/index.php/civitas/article/view/39640>. Acesso em: 1 de set. 2024.

KEMP, S. Digital 2024: Brazil. **DataReportal**. [S.l.]: Kepios, 2024. Disponível em: <https://datareportal.com/reports/digital-2024-brazil>. Acesso em: 1 de set. 2024.

LASWELL, H. **A linguagem da política**. Brasília: Editora Universidade de Brasília, 1982.

LEMO, A. Dataficação da vida. **Civitas**: Revista de Ciências Sociais, São Paulo, v. 21, n. 2, p. 193-202, mai.-ago. 2021. Disponível em: <https://doi.org/10.15448/1984-7289.2021.2.39638>. Acesso em: 1 de set. 2024.

LEMO, A. **Cibercultura**: tecnologia e vida social na cultura contemporânea. 9. ed. Porto Alegre: Sulina, 2023. Disponível em:

<https://www.editorasulina.com.br/img/sumarios/289.pdf>. Acesso em: 1 de set. 2024.

LÉVY, P. **A Inteligência Coletiva**: por uma antropologia do ciberespaço. São Paulo: Edições Loyola, 1998.

LÜDTKE, S.; SPAGNUOLO, S. (Coord.). **Atlas da Notícia**. 6. ed. São Paulo: Projor, 2023. Disponível em: <https://www.atlas.jor.br/>. Acesso em: 1 de set. 2024.

MANCINI, L.; VASCONCELLOS, F. Jornalismo de Dados: conceito e categorias. **Revista Fronteiras** – Estudos Midiáticos, v. 18, n. 1, p. 69-82, jan.-abr. 2016.

Disponível em:

<https://revistas.unisinos.br/index.php/fronteiras/article/view/fem.2016.181.07/5300>. Acesso em: 1 de set. 2024.

MATZAT, L. Datenjournalismus. **Bundeszentrale für politische Bildung**, 26 de out. 2011.

Disponível em: <https://www.bpb.de/themen/daten/opendata/64069/datenjournalismus/>.

Acesso em: 1 de set. 2024.

MAYER-SCHÖNBERGER, V.; CUKIER, K.; **Big Data**: A Revolution that Will Transform how We Live, Work, and Think. Boston: Houghton Mifflin Harcourt, 2013. Disponível em:

<https://books.google.com.br/books?hl=pt-BR&lr=&id=uy4lh-WEhhIC&oi=fnd&pg=PP1&ots=Jve8jiKRIL&sig=s9cEiTAEj-1PEhCBFIcCzChFGBc#v=onepage&q&f=false>.

Acesso em: 1 de set. 2024.

MCLUHAN, M. **Os meios de comunicação como extensão do homem**. 18. ed. São Paulo: Cultrix, 1964.

MCLUHAN, M. [1977]. O meio é a mensagem. **Monday Conference**. Entrevista concedida à ABC Radio National. You Tube, 15 de jun. 2020. Disponível em:

https://www.youtube.com/watch?v=RX-Gqa3_XpU.

Acesso em: 1 de set. 2024.

MELO, J. M.; ASSIS, F.; J. Gêneros e formatos jornalísticos: um modelo classificatório.

Intercom: Revista Brasileira de Ciências da Comunicação, v. 39, n. 1, p. 39-56, jan.-abr.

2016. Disponível em: <https://www.redalyc.org/articulo.oa?id=69870711003>. Acesso em: 1 de set. 2024.

MIELNICZUK, L. **Características e implicações do jornalismo na Web**. Trabalho

apresentado no II Congresso da SOPCOM. Lisboa, 2001. Disponível em:

https://facom.ufba.br/jol/pdf/2001_mielniczuk_caracteristicasimplicacoes.pdf. Acesso em: 1 de set. 2024.

MIELNICZUK, L. **Jornalismo na web**: uma contribuição para o estudo do formato da notícia na escrita hipertextual. 2003. Tese (Doutorado em Comunicação) – Faculdade de Comunicação da Universidade Federal da Bahia, Salvador, 2003. Disponível em:

<https://repositorio.ufba.br/handle/ri/6057>. Acesso em: 1 de set. 2024.

MIELNICZUK, L. Webjornalismo de terceira geração: continuidades e rupturas no jornalismo desenvolvido para a web. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA

COMUNICAÇÃO, 27., Porto Alegre. **Anais [...]**. Porto Alegre: Intercom, 2004. Disponível em:
<https://www.portcom.intercom.org.br/pdfs/33239839420892013900619660266793099419.pdf>
 f. Acesso em: 1 de set. 2024.

MINHA CONEXÃO. **Ranking Minha Conexão 2024**. 2024. Disponível em:
<https://www.minhaconexao.com.br/ranking>. Acesso em: 1 de set. 2024.

NUNES, A.C.B. Jornalismo digital de quinta geração: as publicações para tablets em diálogo com o desenvolvimento da web. **ALCEU**, v. 17, n. 33, p. 19-39, jul./dez. 2016. Disponível em: <https://doi.org/10.46391/ALCEU.v17.ed33.2016.149>. Acesso em: 1 de set. 2024.

PARASIE, S.; DAGIRAL, E. Jornalismo baseado em dados e o bem público: “repórteres assistidos por computador” e “jornalistas programadores” em Chicago. **New Media & Society**, v. 15, n. 6, p. 853-871, 2013. Disponível em: <https://doi.org/10.1177/1461444812463345>. Acesso em: 1 de set. 2024.

PERUZZO, C. M. K. Mídia regional e local: aspectos conceituais e tendências. **Comunicação & Sociedade**, São Bernardo do Campo, v. 26, n. 43, p. 67-84, 2005. Disponível em: <https://pdfs.semanticscholar.org/b85f/f21fe1af68936a0333f96599b612f30edca5.pdf>. Acesso em: 1 de set. 2024.

PINTO, M.C.; BARBOSA, S. Inteligências Artificiais (IAs) no Jornalismo Digital Brasileiro: Contexto Histórico de Processos Inovadores. In: CANAVILHAS, J.; RODRIGUES, C.; GIACOMELLI, F. (Orgs.). **Inteligência artificial e jornalismo móvel: Contextos, tendências, práticas e perspectivas**. Covilhã: LabCom; UBI, 2024. p. 57-90. Disponível em: https://www.researchgate.net/publication/380979121_INTELIGENCIAS_ARTIFICIAIS_IAS_NO_JORNALISMO_DIGITAL_BRASILEIRO_CONTEXTO_HISTORICO_E_PROCESSOS_INOVADORES. Acesso em: 1 de set. 2024.

PORTARI, R. D. L. **A construção da violência e da morte nas capas dos jornais Agora São Paulo e Folha de S. Paulo**. 2009. 134 f. Dissertação (mestrado em Comunicação Midiática) - Faculdade de Arquitetura, Artes e Comunicação da Universidade Estadual Paulista, 2009. Disponível em: <http://hdl.handle.net/11449/89436>. Acesso em: 1 de set. 2024.

PRIVACIDADE Hackeada. Produção de Karim Amer. Direção de Karim Amer e Jehane Noujaim. [s.l.]: Netflix, jul. 2019. Stream (110min.).

ROBERTO, M.; FIDALGO, A.; BUCKINGHAM, D. De que falamos quando falamos de infoexclusão e literacia digital? Perspectivas dos nativos digitais. **Observatório (OBS*) Journal**, v. 9, n. 1, p. 43-54, 2015. Disponível em: https://labcom.ubi.pt/publicacoes/201503021251-819_3189_1_pb.pdf. Acesso em: 1 de set. 2024.

RODRIGO, Pablo. Empresas suspeitas de cartel na Saúde mantêm contrato de R\$100 mi. **Gazeta Digital**, 02 de abr. 2023. Disponível em: <https://tinyurl.com/4s5vw95n>. Acesso em: 13 de mai. 2024.

ROUSSEAU, J.J. [1762]. **Do contrato social** - Ensaio sobre a origem das línguas. Vol. I. São Paulo: Nova Cultural, 1997.

SAMPAIO, R.C.; LYCARIÃO, D. **Análise de conteúdo categorial**: manual de aplicação. Brasília: Enap, 2021.

SANTAELLA, L. O livro como prótese reflexiva. **Matrizes**, São Paulo, v. 13, n. 3, p. 1-35, 2019. Disponível em: <https://www.redalyc.org/articulo.oa?id=143066289004>. Acesso em: 1 de set. 2024.

SANTOS, M. C. Existe jornalismo transmídia? Considerações sobre o reúso de conceitos. **Revista GEMINIS**, São Carlos, v. 8, n. 3, pp.136-149, set./dez. 2017. Disponível em: <https://www.revistageminis.ufscar.br/index.php/geminis/article/view/308/pdf>. Acesso em: 1 de set. 2024.

SCHWINGEL, C. Jornalismo digital de quarta geração: a emergência de sistemas automatizados para o processo de produção industrial no Jornalismo Digital. In: ENCONTRO ANUAL DA COMPÓS, 14., Niterói. **Anais eletrônicos** [...]. Campinas: Galoá, 2005. Disponível em: <https://proceedings.science/compos/compos-2005/trabalhos/jornalismo-digital-de-quarta-geracao-a-emergencia-de-sistemas-automatizados-para?lang=pt-br>. Acesso em: 1 de set. 2024.

SEGATA, J.; RIFIOTIS, T. Digitalização e dataficação da vida. **Civitas** - Revista de Ciências Sociais, São Paulo, v. 21, n. 2, p. 186-192, 2021. Disponível em: <http://dx.doi.org/10.15448/1984-7289.2021.2.40987>. Acesso em: 1 de set. 2024.

SPAGNUOLO, S. Manual de jornalismo de dados para redações. **Medium** – Volt Data Lab, 12 de nov. 2019. Disponível em: <https://medium.com/volt-data-lab/manual-de-jornalismo-de-dados-para-reda%C3%A7%C3%B5es-v-1-0-908485625e63>. Acesso em: 1 de set. 2024.

STOCHERO, D. **Privacidade de dados no marketing**: Estudo de caso da Cambridge Analytica na eleição presidencial nos Estados Unidos em 2016. 2020. Monografia (Bacharelado em Publicidade e Propaganda) – Faculdade de Comunicação Social da Universidade de Caxias do Sul (UCS), Caxias do Sul, Rio Grande do Sul, 2020. Disponível em: <https://repositorio.ucs.br/xmlui/bitstream/handle/11338/8976/TCC%20Daniel%20Stochero%20de%20Aguiar.pdf>. Acesso em: 1 de set. 2024.

TRÄSEL, M. Jornalismo guiado por dados: características definidoras e uma proposta de formulação do conceito. **Estudos em Jornalismo e Mídia**, v. 11, n. 1, jan.-jun. 2014. Disponível em: https://www.academia.edu/60143734/Jornalismo_guiado_por_dados_aproxima%C3%A7%C3%B5es_entre_a_identidade_jornal%C3%ADstica_e_a_cultura_hacker. Acesso em: 1 de set. 2024.

TRÄSEL, M. Tese. **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) – Programa de Pós-Graduação em Comunicação Social da Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2014. Disponível em: <https://repositorio.pucrs.br/dspace/handle/10923/6841>. Acesso em: 1 de set. 2024.

TRÄSEL, M.; FONTOURA, M. Microfinanciamento e pluralização: comparação das coberturas do webjornal participativo Spot.Us e da imprensa americana. **Intexto**, Porto Alegre, n. 27, p. 38-54, dez. 2012. Disponível em: <https://seer.ufrgs.br/index.php/intexto/article/view/30384/23477>. Acesso em: 1 de set. 2024.

VIEIRA, Alessandro. **Projeto de Lei n. 2.630/2020**. Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet. Brasília, DF: Senado Federal, 03 de jul. 2020. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2256735>. Acesso em: 19 de set. 2025.

VIEIRA, Luiz Fernando. Gazeta Digital completa 21 anos em constante evolução. **A Gazeta**, Vida, 13 de mai. 2024.

WARDLE, C.; DERA KHSHAN, H. Information Disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe, 2017. Disponível em: <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinaryframework-for-research-and-policy-making.html>. Acesso em: 19 de set. 2025.