
Prêmio Cláudio Weber Abramo: Uma análise de conteúdo categorial das reportagens finalistas da premiação brasileira de jornalismo de dados¹

Giselia AMANYARA DE SOUZA²

Marcelo Alves DOS SANTOS JUNIOR³

Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, RJ

RESUMO

O jornalismo de dados é uma área que está em ascensão no Brasil e no mundo desde o século XX. Nacionalmente, no entanto, é um setor pouco explorado academicamente. O objetivo deste trabalho é analisar como se estabelecem as novas formas de manifestação da imprensa brasileira através do uso de dados. A pesquisa se propõe a responder quem são os principais produtores do jornalismo de dados no país e quais são as temáticas, fontes e tipos de visualização mais recorrentes nas reportagens de dados. Para isso, serão analisadas as reportagens finalistas das quatro edições do Prêmio Cláudio Weber Abramo de Jornalismo de Dados (2019-2022) através de uma análise de conteúdo categorial, com base nos processos metodológicos de Sampaio e Lycarião (2021) e as categorias de codificação trabalhadas por Loosen e colaboradores (2020).

PALAVRAS-CHAVE: Jornalismo de dados no Brasil; Prêmio Cláudio Weber Abramo; Premiação de jornalismo de dados; Análise de conteúdo categorial

O jornalismo de dados (JD) é um setor midiático que está em ascensão desde antes do estabelecimento da internet, e vem sendo promovido com mais intensidade diante da ampliação de movimentos por dados abertos (*open data*) e transparência pública (*open government*) mundo afora (LIMA, 2021). Com a digitalização dos meios, a imprensa precisou encontrar novos formatos: se historicamente o jornalismo usa dados de maneira ilustrativa em reportagens - como a quantidade de vezes que dois times de futebol se enfrentaram ao longo da década, no caso de matérias de esporte, ou percentuais de inflação e juros em pautas de economia -, atualmente eles são usados também de forma estruturante para o desenvolvimento de pautas específicas e mais complexas (MANCINI; VASCONCELLOS, 2016).

São poucos os estudos referentes ao JD no Brasil, se os compararmos ao volume de publicações sobre a prática de dados realizada internacionalmente. Assim, faz-se

¹ Trabalho apresentado no Intercom Júnior – IJ01 – Jornalismo do 46º Congresso Brasileiro de Ciências da Comunicação, realizado de 4 a 8 de setembro de 2023.

² Graduanda do Curso de Jornalismo da PUC-Rio, e-mail: [giseliaamanyara@yahoo.com](mailto:giseliamanyara@yahoo.com).

³ Orientador do trabalho. Professor do Curso de Estudos de Mídia da PUC-Rio, e-mail: marcelo_alves@puc-rio.br.

necessário analisar como se estabelecem as novas formas da imprensa de se manifestar nacionalmente, através do JD, a fim de torná-lo mais palpável e acessível aos jornalistas e leitores que pouco o conhecem. Este trabalho se propõe a estudar as reportagens finalistas das quatro edições do Prêmio Cláudio Weber Abramo de Jornalismo de Dados (2019-2022), uma iniciativa da Escola de Dados e Open Knowledge Brasil, instituições que buscam incentivar o conhecimento e o uso de dados para ampliar a excelência nas produções jornalísticas no país.

A partir disso, a pesquisa pretende esclarecer quais são as temáticas mais recorrentes no Jornalismo de Dados e como são produzidas suas reportagens - no contexto da premiação brasileira, que é pioneira no país, e de acordo com o modelo de avaliação do júri diante das matérias inscritas. Com isso, será possível avaliar se há assuntos mais favorecidos no setor e apontar ainda como o fazer jornalístico a partir de dados se estabelece no Brasil dentro da delimitação proposta.

DEFINIÇÕES E DISTINÇÕES DO JORNALISMO DE DADOS

Pensar e conceituar o JD é um esforço coletivo e relativamente recente. Não à toa, Bounegru e Gray (2021) questionam: “O que é jornalismo de dados? Para que serve? O que pode fazer? Quais limitações e oportunidades apresenta? Quem e o que estão envolvidos em sua criação e compreensão?” (BOUNEGRU; GRAY, 2021, p. 2). Os autores são responsáveis por organizar um livro sobre JD, que é um projeto colaborativo de 54 artigos científicos de profissionais referenciados do nicho e funciona como manual para interessados na área.

Não é incomum que o JD seja confundido com o jornalismo feito com dados. Considerando que ambos podem fazer uso de números e estatísticas como fonte, bem como criar visualizações através de infográficos, em especial em editorias como economia e política, é compreensível que a distinção entre um e outro seja sutil e até mesmo imperceptível em alguns casos. Não obstante, existem algumas sugestões de como os dois termos se distinguem, não apenas em nível conceitual, mas - e principalmente - em modelos de produção jornalística no dia a dia das redações.

Bazzo, Martins e Barbosa (2020) descrevem o JD como a modalidade em que se trabalha com uma quantidade de dados que só pode ser tratada por computadores (*big data*), e que tem como objetivo apresentar tais dados através da combinação da visualização com uma narrativa jornalística envolvente e de relevância social, que facilite

a compreensão das informações. O JD é tratado pelos autores como uma “ferramenta de promoção da cidadania e da democracia na era da internet” (BAZZO; MARTINS; BARBOSA, 2020, p. 281). Os pesquisadores constataram, inclusive, um aumento de iniciativas científicas em pesquisas de JD entre 2016 e 2017, possivelmente como resultado de uma demanda “por mais acesso à informação e mais transparência dos dados públicos” (ibid., p. 300).

No que tange a função social do JD, De Lima Santos (2019) descreve a prática de dados como um método de investigar e contar histórias a fim de engajar o público e ajudar a diminuir as barreiras à compreensão e à imersão em dados. Em sua revisão, o pesquisador aponta que “o jornalismo de dados representa, para muitos autores, a democratização de recursos, ferramentas, técnicas e métodos que antes eram restritos a poucos, como especialistas, governos, entre outros” (DE LIMA SANTOS, 2019, p. 16).

Já segundo Mancini e Vasconcellos (2016), as reportagens *com* dados se apropriam de informações numéricas de forma ilustrativa, a fim de reforçar a ideia construída na execução da pauta, enquanto as reportagens *de* dados utilizam os números como a razão da matéria. “No primeiro caso, o dado quantitativo teria o papel de auxiliar a ilustrar uma reportagem; no segundo, o dado seria o próprio fundamento da pauta e a história das relações entre os dados conduziria, neste caso, a reportagem” (MANCINI; VASCONCELLOS, 2016, p. 75).

Seriam então classificadas como JD aquelas matérias que atendessem aos seguintes critérios: (1) extração de dados, que avalia se a equipe responsável pela reportagem buscou ou produziu uma base de dados (primária, a partir de microdados; ou secundária, a partir de dados agregados) para a execução da pauta; (2) estruturação, a fim de examinar a organização dos dados na matéria; (3) análise, que julga se o texto oferece alguma interpretação dos dados para os leitores através de comparações estatísticas e apontamento de causas, consequências e implicações; e (4) visualização, indicando se a reportagem apresenta gráficos ou infográficos, sejam eles interativos ou não (MANCINI; VASCONCELLOS, 2016).

Tendo em vista que a prática de dados pode ser entendida como um método de investigação (DE LIMA SANTOS, 2019), o JD poderia ser, desta vez, confundido com o jornalismo investigativo. Ambas as práticas compartilham processos semelhantes de extração, estruturação, análise e, eventualmente, visualização de dados. Inclusive, a

Associação Brasileira de Jornalismo Investigativo (Abraji) é a principal instituição na promoção do desenvolvimento do JD no Brasil através de congressos e projetos na área.

Nesse sentido, o uso de dados no jornalismo se estabelecerá como uma atualização do trabalho que já era feito nas reportagens de investigação, mas com uma explícita distinção: a capacidade de programação. Por mais que as duas práticas se apropriem da disseminação de movimentos de dados abertos e transparência pública na web, somente o JD estaria apto a construir espaços de interatividade para que o leitor, presumindo o mínimo de letramento digital, possa por si só navegar pelas informações e formar ideias a respeito de um assunto (CHARBONNEAUX; GKOUSKOU-GIANNAKOU, 2015).

Moura (2018) obteve uma conclusão parecida, porém mais complexa. Na revisão do trabalho de Crucianelli (2013), a autora destaca que o JD não se define simples e exclusivamente como uma nova técnica que reproduz processos do jornalismo investigativo, mas se apresenta como uma combinação de diferentes práticas da imprensa. Entretanto, a “possibilidade de produzir visualizações de dados interativas e também a integração do trabalho do programador entre as práticas jornalísticas” (MOURA, 2018, p. 17) são acrescentadas pelo JD a esse conjunto de disciplinas.

BREVE CRONOLOGIA DO JORNALISMO DE DADOS

O *The Guardian* foi o veículo pioneiro no uso de dados em grande escala na imprensa ainda no século XIX. Em 1821, com o levantamento das escolas de Manchester na primeira edição do jornal inglês e, em 1858, com a publicação dos estudos de Florence Nightingale, que montou relatórios visuais com base em estatística aplicada para representar as causas de morte dos soldados britânicos na Guerra da Crimeia (ROGERS, 2010; 2011). No entanto, foi a partir da informatização das redações no fim da década de 1960 que as práticas jornalísticas começaram a ser reformuladas (BARBOSA, 2007).

Um dos precursores do JD, o jornalismo de precisão foi desenvolvido por Philip Meyer na década de 1960, quando Meyer era repórter do *Detroit Free Press* e apurou os conflitos ocorridos em Detroit em 1967. No episódio, a imprensa sugeriu que os tumultuadores seriam parte de grupos minoritários ou imigrantes do Sul dos EUA, mas o levantamento de Philip mostrou que não existia relação entre a participação nos motins e a condição econômica dos entrevistados, e que os nativos da cidade tinham três vezes mais propensão a participar de atos ilícitos do que os imigrantes sulistas. A prática foi

instituída com a publicação do livro *Precision Journalism* em 1973, em que Meyer explica como utilizar técnicas de ciências sociais e computação no jornalismo a fim de prevenir erros de apuração em reportagens.

A técnica da Reportagem Assistida por Computador (RAC) era pouco utilizada nas redações até 1989, quando o jornalista Elliot Jaspion fundou o Instituto Nacional para a Reportagem Assistida por Computador na Universidade de Missouri (TRÄSEL, 2014). O objetivo do grupo era treinar jornalistas para utilizar o computador na apuração de notícias. A RAC é uma alternativa às análises manuais do jornalismo investigativo, descrita como um “conjunto de técnicas que explora as buscas de dados online, a interpretação e análise de mapas, estatísticas e planilhas provenientes de dados públicos” (MOURA, 2018, p. 5-6).

A dificuldade de acesso a bancos de dados governamentais, devido à gravação das informações em fitas cassetes, por exemplo, atrasou o desenvolvimento da RAC. Porém, o surgimento da World Wide Web em 1991 e a difusão de serviços de computação a partir de 2000 possibilitou a popularização do acesso a planilhas e ferramentas de criação de visualização de dados, bem como permitiu que os jornalistas da RAC ampliassem a distribuição das técnicas de análises de dados para reportagens a profissionais de imprensa de todo o mundo.

Assim, o JD começa a se desenvolver em diferentes países, de maneira particular e em diferentes ritmos, não apenas devido ao contexto cultural, mas por questões econômicas e ausência de mão-de-obra (TRÄSEL, 2014). No Brasil, pode-se dizer que o JD foi construído com o desenvolvimento da *accountability* nacional, movimento que atua em prol da prestação de contas públicas e responsabilização por ações da parte de governantes e empresas privadas com influência na sociedade.

Na década de 1990, os jornalistas Fernando Rodrigues e José Roberto de Toledo adotaram a prática da RAC na *Folha de S. Paulo*. Tutores do Instituto Nacional para a Reportagem Assistida por Computador, da Universidade de Missouri, visitaram a redação paulista para ministrar cursos da técnica, que logo foi integrada ao programa de treinamento do jornal. Outro nome pioneiro do JD no país foi Cláudio Weber Abramo, matemático-jornalista com passagem por veículos como *Folha de S. Paulo* e *Valor Econômico*, que fundou em 2000 a organização Transparência Brasil, dedicada ao trabalho pela liberação de acesso a dados públicos e em prol do combate à corrupção.

A criação da Abraji em 2002 também ajudou a construir o JD no Brasil. A associação surgiu devido à iniciativa do à época diretor da *Folha de S. Paulo* na sucursal do Rio de Janeiro Marcelo Beraba que iniciou uma troca de e-mails com cerca de 45 outros jornalistas brasileiros, após a morte de Tim Lopes, para construir “uma organização que protegesse e representasse os jornalistas no país e também contribuísse para o desenvolvimento da imprensa” (ABRAJI, 2020). José Roberto de Toledo era um dos sócio-fundadores e vice-presidente da associação (TRÄSEL, 2014).

A partir disso, inicia-se o desenrolar do processo de *accountability* político brasileiro, entendido como um movimento em favor da prestação de contas públicas pela transparência governamental no país. A institucionalização das práticas de dados abertos e prestação de contas reflete sobre a performance das democracias e pode ser definida como um conjunto de “ações que geram fluxos e contrafluxos da temática da vida pública e ainda exige do campo político justificativas para suas decisões” (MANCINI; VASCONCELLOS, 2016, p. 70). A figura de Abramo e o agrupamento da Abraji foram fundamentais para a articulação da Lei de Acesso à Informação (LAI), que regulamenta a consulta de dados públicos de entidades municipais, estaduais ou federais por parte de qualquer pessoa, física ou jurídica, sem que seja necessário apresentar motivo. A Lei Federal nº 12.527/11 foi aprovada em 2011 e à ela estão submetidos:

I - os órgãos públicos integrantes da administração direta dos Poderes Executivo, Legislativo, incluindo as Cortes de Contas, e Judiciário e do Ministério Público; II - as autarquias, as fundações públicas, as empresas públicas, as sociedades de economia mista e demais entidades controladas direta ou indiretamente pela União, Estados, Distrito Federal e Municípios (BRASIL, 2011).

Em termos de redações especializadas em dados, o *Estadão* foi o pioneiro no Brasil. O *Estadão Dados* foi inaugurado em maio de 2012 sob a liderança de José Roberto de Toledo com o objetivo de “publicar peças centradas na análise de dados, oriundas do trabalho conjunto entre jornalistas, designers e programadores” (MOURA, 2018, p. 33). Logo, surgiram mais núcleos de dados em veículos de imprensa do país, como *Deltafolha* (*Folha de S. Paulo*) e *Na Base de Dados* (*O Globo*). A partir de 2015, começaram iniciativas independentes no uso de dados na imprensa, como o *Jornal Nexo* e a *Revista Gênero e Número*, e as agências de dados *Volt* e *iNFRA* (LIMA, 2021).

De Lima Santos (2019) mapeou pelo menos 52 veículos de imprensa que trabalhavam com JD no Brasil. Mastrella (2020) realizou um levantamento semelhante

ao de Santos (2019), conforme revisado por Lima (2021), que apontou 46 núcleos nacionais de JD. A cobertura da pandemia da Covid-19 foi um dos exemplos nacionais mais recentes e relevantes do setor. Os resultados da visualização de dados em gráficos, mapas ou tabelas seriam, para Dantas (2022, p. 27), “representações da realidade que ajudam o usuário a compreender e ver mentalmente um assunto denso”. Apesar do registro ainda incipiente de organizações de JD no Brasil, produções complexas de dados mostram o desenvolvimento da área no país e a importância de estudar os novos processos jornalísticos que se estabelecem na imprensa nacional.

METODOLOGIA

A metodologia escolhida para a realização desta pesquisa foi a análise de conteúdo categorial quantitativa dos trabalhos finalistas do Prêmio Cláudio Weber Abramo entre 2019 e 2022, com base nos métodos analíticos propostos por Sampaio e Lycarião (2021). Foi utilizada uma amostra de 39 reportagens, as quais foram analisadas através de uma categorização que utiliza seis variáveis, estabelecidas no trabalho de Loosen, Reimer e De Silva-Schmidt (2017, 2020): (1) tipo de mídia; (2) número de autores; (3) tópico; (4) fonte de dados; (5) acesso ao dados; e (6) visualização.

Os trabalhos finalistas do Prêmio Cláudio Weber Abramo foram escolhidos como *corpus* desta pesquisa, pois configuram um recorte de projetos de destaque no campo do JD no Brasil, considerados como produções de qualidade e dignos de reconhecimento através de protocolos avaliativos estabelecidos pela organização do prêmio, a Escola de Dados (Open Knowledge Brasil). A premiação é a primeira e única dedicada exclusivamente ao reconhecimento de produções de JD no Brasil e segue uma metodologia transparente, estabelecida no regulamento da Escola de Dados.

Os objetivos gerais da organização são estimular e incentivar o uso de dados e iniciativas de transparência governamental para tratar de temas de relevância social e interesse público no Brasil, bem como inspirar estudantes e jornalistas a trabalhar com dados e reconhecer produções de excelência no campo de JD no país (ESCOLA DE DADOS, 2022). Ainda, premiações são consideradas espaços de prestígio e capital social para celebrar trabalhos com mérito e potencial influência para a sociedade. Assim, a fim de evitar a composição do *corpus* de forma parcial e subjetiva, levamos em consideração os trabalhos que profissionais e jornalistas de dados avaliam relevantes.

O artigo de Loosen, Reimer e De Silva-Schmidt (2020) foi escolhido como referência para este trabalho com o objetivo de reproduzir semelhante análise no contexto brasileiro. Parte das categorias indicadas nas variáveis selecionadas para este trabalho foram adaptadas para este artigo. Tomamos como base as mesmas perguntas de pesquisa que os autores utilizaram na produção do artigo original na análise do prêmio internacional *Data Journalism Awards* (DJA) entre 2013 e 2016:

Q1. Quais são os atores que produzem reportagens de jornalismo de dados nomeadas para o DJA (organizações de mídia, equipes internas, parceiros externos)? **Q2.** Quais são os tópicos abordados nos projetos nomeados para o DJA? **Q3.** Em que dados (tipo, fonte, análise) as histórias nomeadas para o DJA são baseadas, e quais visualizações e recursos interativos elas fornecem? (LOOSEN; REIMER; DE SILVA-SCHMIDT, 2020, p. 1250).⁴

A análise de conteúdo foi realizada de forma individual em duas etapas: pré-análise de aproximadamente 15% da amostra ($n = 6$) e codificação final em duas fases. O teste de estabilidade individual cego, que verifica a confiabilidade em pesquisas individuais (SAMPAIO; LYCARIÃO, 2021) foi feito após o fim da segunda fase da codificação final através da ferramenta de confiabilidade no site do Professor Doutor Deen Freelon⁵ e atingiu um percentual de concordância de 78,42% e um coeficiente alfa de Krippendorff médio de 0.608. O *corpus* e a metodologia detalhada estão organizados nesta planilha⁶ aberta.

JORNALISMO DE DADOS NO BRASIL EM NÚMEROS

Os resultados a seguir buscam responder às perguntas propostas por Loosen, Reimer e De Silva-Schmidt (2020) e aqui replicadas. Vamos tentar entender em primeiro lugar quem são os atores do JD no Brasil que aparecem em destaque no Prêmio Cláudio Weber Abramo (CWA), ou seja, quais são as organizações de mídias e o tamanho das equipes responsáveis pelos trabalhos finalistas e vencedores. Em seguida, buscaremos compreender quais são os assuntos mais recorrentes nos projetos finalistas da premiação brasileira e quais são as fontes de dados e formas de acesso que os veículos utilizam para a construção de reportagens de JD no âmbito do prêmio.

⁴ Tradução livre.

⁵ Disponível em: <http://dfreelon.org/utills/recalfront/recal2/>. Acesso em: 26 jun. 2023.

⁶ <https://docs.google.com/spreadsheets/d/1La6LDmnW83PZLZVHNgynaCjU0oDvrWPPdzkeao7LLsc/>.

No contexto da premiação brasileira, as mídias nativas digitais são o principal *tipo de mídia* responsável pelos projetos finalistas do Prêmio CWA com 30,8% da amostra, e correspondem ainda a 46,2% de vencedores. Entendemos esse resultado como consequência do surgimento constante de novos portais de jornalismo on-line e atores independentes do JD na internet para além dos veículos tradicionais de mídia (LIMA, 2021). Dentre os trabalhos que correspondem à mídia nativa digital, no entanto, 25% ($n = 3$) são do portal *GI*, que faz parte do *Grupo Globo*, grande conglomerado de mídia no Brasil. Os jornais correspondem a 23,1% das reportagens finalistas da premiação, segundo maior percentual de projetos.

Tabela 1 - Resultados da variável *tipo de mídia*

Tipo de mídia	Finalistas	Vencedores
Mídia nativa digital	12 (30,8%)	6 (46,2%)
Jornal	9 (23,1%)	1 (7,7%)
Organização de jornalismo investigativo	8 (20,5%)	1 (7,7%)
Agência de dados	4 (10,3%)	2 (15,4%)
Revista impressa	3 (7,7%)	1 (7,7%)
Agência de notícias	3 (7,7%)	2 (15,4%)
Total	39 (100%)	13 (100%)

Fonte: autoria própria.

Organizações de jornalismo investigativo também são parte fundamental do JD no Brasil, considerando que são os principais incentivadores da prática no país, como *The Intercept Brasil* e *Agência Pública*, e representam 20,5% dos finalistas. As demais categorias que aparecem são: agência de dados (10,3%); revista impressa (7,7%); e agência de notícias (7,7%). Apesar de serem categorias trabalhadas por Loosen e colaboradores (2020), emissora pública; emissora privada; organização não-jornalística; e mídia universitária não tiveram nenhuma incidência na amostra.

No que diz respeito ao *número de autores*, a produção em equipe é a mais recorrente entre os trabalhos finalistas da premiação e representa 87,2% da amostragem de finalistas e 92,3% ($n = 12$) dos projetos vencedores do Prêmio CWA. As produções individuais representam 12,8% entre os finalistas e 7,7% dos vencedores. Os resultados reforçam o caráter colaborativo do JD, de acordo, inclusive, com a narrativa histórica do

desenvolvimento da prática no Brasil, que é fortalecida com a criação da Abraji e demais organizações de jornalistas brasileiros (TRÄSEL, 2014; LIMA, 2021).

A média de autores por projeto foi de onze pessoas (DV [desvio padrão] = 24,96), com base na lista de nomes que assinaram os trabalhos finalistas da premiação. Ao mesmo tempo, poucos trabalhos foram feitos em parceria com outros veículos ou autores externos: apenas 15,4% das reportagens finalistas foram realizadas em colaboração com outras redações. Dentre elas, 33,3% ($n = 2$) foram vencedores do Prêmio CWA. Isso pode ser considerado resultado da especialização de redações e profissionais em JD em técnicas de estatística e programação no Brasil, o que reduz a necessidade de atores externos às equipes compostas nos veículos (LIMA, 2021).

Quanto à variável *tópico*, política é a temática mais recorrente entre os trabalhos (30,8%), além de ser o principal assunto entre os projetos vencedores (41,7%), o que reforça a ideia de que o JD funciona como ferramenta democrática no Brasil, por mais transparência de dados públicos e de governantes no país (MANCINI; VASCONCELLOS, 2016; DE LIMA SANTOS, 2019; BAZZO; MARTINS; BARBOSA, 2020).

Tabela 2 - Resultados da variável *tópico*

Tópico	Finalistas	Vencedores
Política	12 (30,8%)	5 (38,5%)
Saúde e ciência	7 (17,9%)	3 (23,1%)
Segurança pública	6 (15,4%)	2 (15,4%)
Meio ambiente e sustentabilidade	6 (15,4%)	2 (15,4%)
Sociedade e comportamento	3 (7,7%)	0 (0%)
Economia e negócios	3 (7,7%)	0 (0%)
Cidade	2 (5,1%)	1 (7,7%)
Cultura, educação, esportes e tecnologia	0 (0%)	0 (0%)
Total	39 (100%)	13 (100%)

Fonte: autoria própria.

Saúde e ciência é a categoria que aparece em seguida, correspondendo a 17,9% dos trabalhos, dentre os quais 57,14% ($n = 4$) fazem referência à Covid-19. Isso reforça

a proposição da importância de representações visuais em momentos de crise para facilitar a compreensão e absorção de informações para o leitor (DANTAS, 2022).

Segurança pública (15,4%) e meio ambiente e sustentabilidade (15,4%) foram as categorias com a terceira maior incidência da análise. No caso da categoria de meio ambiente, cerca de 83,3% ($n = 5$) dos projetos falavam exclusivamente sobre os tipos de exploração indevida que ocorrem na região da Amazônia Legal, produzidas por organizações de jornalismo investigativo como a *Agência Pública* (em parceria com a *Comissão Pastoral da Terra*), *Ambiental Media* e *Repórter Brasil*. Sociedade e comportamento (7,7%), economia e negócios (7,7%) e cidade (5,1%) tiveram menor cobertura, enquanto cultura, educação, esportes e tecnologia não apareceram entre os finalistas da premiação.

Da variável *fonte de dados*, 76,9% dos trabalhos utilizaram dados de instituições oficiais como a primeira base de informações para a construção de reportagens, como Ministério da Saúde, Portal de Dados Abertos da Câmara dos Deputados e Conselho Nacional de Justiça. Esses resultados apontam para a relevância do movimento de *accountability* nacional pela transparência pública dos governos, alimentado por diversos setores do jornalismo em favor da manutenção da democracia (MANCINI; VASCONCELLOS, 2016).

Tabela 3 - Resultados da variável *fonte de dados*⁷

Fonte de dados	1	2	3
Instituição oficial	30 (76,9%)	13 (68,4%)	9 (90%)
Base de dados própria	5 (12,8%)	5 (26,3%)	1 (10%)
Não indicado	2 (5,1%)	1 (5,3%)	0 (0%)
Empresa privada	1 (2,6%)	0 (0%)	0 (0%)
Organização não-comercial	1 (2,6%)	0 (0%)	0 (0%)
Total	39 (100%)	19 (100%)	10 (100%)

Fonte: autoria própria.

⁷ Considerando que os projetos inscritos no Prêmio CWA poderiam apresentar mais de um tipo de fonte de dados, forma de acesso e visualização, as variáveis *fonte de dados*, *acesso aos dados* e *visualização* foram codificadas de uma a três vezes cada uma, a depender das especificidades de cada reportagem.

Ainda observando as primeiras fontes de dados mencionadas nos trabalhos, 12,8% deles fazem uso de base de dados própria, através de levantamento realizado em redes sociais, investigação em aplicativos de mensagens ou entrevistas, por exemplo, enquanto 5,2% correspondem a dados de empresa privada ou organização não-comercial (2,6% cada) e 5,1% não indicam a fonte utilizada.

Dados públicos foram o principal tipo de acesso aos primeiros dados citados em 46,2% dos trabalhos da amostra, seguido de solicitações via LAI (25,6%), o que demonstra não somente a fundamentalidade da transparência pública para investigações jornalísticas, como também a importância da articulação da Lei de Acesso à Informação, que desde 2013 contribui diretamente para a manutenção do sistema democrático (MANCINI; VASCONCELLOS, 2016).

Trabalhos sem indicação de forma de acesso (15,4%) ficaram à frente de coleta própria (12,8%). Raspagem e vazamentos de dados não foram opções explicitadas por nenhum dos finalistas. Já no que diz respeito ao *acesso aos dados 2* e *acesso aos dados 3*, dados públicos continuam a ser a opção mais recorrente com 57,9% e 80% da amostra de finalistas. Em *acesso aos dados 2*, coleta de dados própria (26,3%) e não indicado (10,5%) tiveram maior incidência do que solicitações via LAI (5,3%).

Tabela 4 - Resultados da variável *acesso aos dados*

Acesso aos dados	1	2	3
Dados públicos	18 (46,2%)	11 (57,9%)	8 (80%)
Via LAI (Lei de Acesso à Informação)	10 (25,6%)	1 (5,3%)	1 (10%)
Não indicado	6 (15,4%)	2 (10,5%)	1 (10%)
Coleta própria	5 (12,8%)	5 (26,3%)	0 (0%)
Raspagem e vazamento de dados	0 (0%)	0 (0%)	0 (0%)
Total	39 (100%)	19 (100%)	10 (100%)

Fonte: autoria própria.

Entre os trabalhos finalistas, 48,72% utilizam somente uma fonte de dados para o desenvolvimento do projeto, enquanto 25,64% fazem uso de três ou mais bases de dados diferentes. 20,51% consultam pelo menos duas fontes diferentes e 5,13% não indicam as bases de dados consultadas. No caso das reportagens sem indicação de fonte, uma delas aponta que a forma de acesso às informações foi realizada via LAI, no entanto.

Entendemos como importante mencionar um outro caso especial, em que o trabalho - *As pensões e os bilhões da família militar*⁸, da *piauí* - acessou dados públicos através do Portal da Transparência, dados esses que só foram disponibilizados publicamente após solicitação da *Fiquem Sabendo* via LAI.

Os elementos de *visualização* que apareceram com maior recorrência nos trabalhos finalistas do Prêmio CWA foram gráfico estático simples (64,1%, percentual total de aparição na análise), imagem (56,41%), mapa (25,64%) e painel interativo (25,64%). Vídeo (23,07%) e tabela (5,12%) são recursos visuais menos recorrentes entre os finalistas. Dentre os trabalhos analisados, 7,69% apresentam outro tipo de visualização, como ilustração ou inserção de tweet via HTML.

Tabela 5 - Resultados da variável *visualização*

Visualização	1	2	3
Gráfico estático simples	12 (30,8%)	11 (40,7%)	2 (12,5%)
Imagem	9 (23,1%)	7 (25,9%)	6 (37,5%)
Mapa	7 (17,9%)	1 (3,7%)	2 (12,5%)
Painel interativo	5 (12,8%)	4 (14,8%)	1 (6,3%)
Vídeo	4 (10,3%)	2 (7,4%)	3 (18,8%)
Sem visualização	1 (2,6%)	0 (0%)	0 (0%)
Outra visualização	1 (2,6%)	1 (3,7%)	1 (6,3%)
Tabela	0 (0%)	1 (3,7%)	1 (6,3%)
Gráfico estático combinado	0 (0%)	0 (0%)	0 (0%)
Total	39 (100%)	27 (100%)	16 (100%)

Fonte: autoria própria.

Somente um trabalho finalista não apresentou visualização (2,56%): a *Newsletter Don't LAI to me*⁹, da agência de dados *Fiquem Sabendo*, projeto vencedor na categoria *dados abertos* em 2019. Entre os vencedores, gráfico estático simples (23,7%, $n = 9$) e imagem (17,94%, $n = 7$) foram os recursos mais utilizados, enquanto mapa, vídeo e painel interativo empataram em 7,69% de incidência ($n = 3$).

⁸ Disponível em: <https://piaui.folha.uol.com.br/as-pensoes-e-os-bilhoes-da-familia-militar/>. Acesso em: 02 jul. 2023.

⁹ Disponível em: <https://fiquemsabendo.com.br/transparencia/newsletter-dont-lai-to-me/>. Acesso em: 22 jun. 2023.

CONSIDERAÇÕES FINAIS

O JD é uma das mais recentes vertentes do jornalismo tradicional, entendida como responsável por tratar grandes volumes de dados e por atuar como ferramenta em favor da cidadania e manutenção da democracia na era digital. Ao nosso ver, é o caráter colaborativo do JD que faz com que a área seja constantemente desenvolvida e explorada através de organizações não-comerciais e associações de jornalismo investigativo, que foram e ainda são os principais fomentadores desta prática que, se podemos dizer, ainda está sendo descoberta e gradativamente construída.

São poucos os estudos referentes ao JD no Brasil, se os compararmos ao volume de publicações sobre a prática de dados realizada internacionalmente. Portanto, a fim de perceber como o fazer jornalístico a partir de dados se estabelece em território brasileiro, considerando a função social do uso de dados para o jornalismo nacional, este trabalho teve como objetivo apresentar quem são os principais atores responsáveis pela produção do JD no país e quais são as temáticas mais recorrentemente trabalhadas, além de compreender quais são as fontes de dados mais utilizadas e os recursos visuais mais aplicados em reportagens de dados nacionais.

Os resultados levantados na análise mostram que produções de JD tendem a atender objetivos sociais, em prol da transparência pública governamental, através de produções realizadas em equipes multidisciplinares e que atualmente podem ser desenvolvidas por novos veículos, independentes e autônomos, para além dos grandes conglomerados de mídia. Com isso, percebemos o modelo de funcionamento do JD no Brasil - mesmo que de forma limitada diante do enxuto n amostral - e podemos concluir que a área acompanha a proposição teórica para o uso de dados no jornalismo brasileiro.

REFERÊNCIAS

BARBOSA, Suzana. **Jornalismo Digital em Base de Dados (JDBD): um paradigma para produtos jornalísticos digitais dinâmicos**. 2007. Tese (Doutorado) - Faculdade de Comunicação, Universidade Federal da Bahia, Bahia, 2007.

BAZZO, Jessica; MARTINS, Dalton Lopes; BARBOSA, Filipe Augusto Couto. O surgimento da pesquisa em Jornalismo de Dados no Brasil. **Intexto**, Porto Alegre, n. 50, p. 280-302, 2020.

BOUNEGRU, Liliana; GRAY, Jonathan. **Manual de Jornalismo de Dados: Rumo a uma prática crítica de dados**. Associação Brasileira de Jornalismo Investigativo (ABRAJI), ed. 2, p. 310-319, 2021. Disponível em: <http://datajournalismcom.s3.eu-central->

1.amazonaws.com/handbooks/The-Data-Journalism-Handbook-2_PT.pdf. Acesso em: 4 abr. 2023.

BRASIL. **Lei nº 12.527, de 18 de novembro de 2011**. Regulamenta o acesso à informações públicas por qualquer pessoa, física ou jurídica, de órgãos e entidades públicas municipais, estaduais ou federais. Brasília: Casa Civil, [2011]. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 14 mai. 2023.

CHARBONNEAUX, Juliette; GKOUSKOU-GIANNAKOU, Pergia. O Jornalismo de “Dados”, uma Prática de Investigação? Um olhar sobre os casos alemão e grego. **Brazilian Journalism Research**, vol. 11, ed. 2, p. 266–291, 2015.

DE LIMA SANTOS, Mathias Felipe. **Tem #DDJBR aqui? Mapeando a presença do jornalismo de dados no Brasil**. Trabalho de Conclusão de Curso (Graduação) - Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019.

ESCOLA DE DADOS. **Regulamento**. Brasil: Escola de Dados, 2022. Disponível em: <https://premio.jornalismodedados.org/regulamento/2022/>. Acesso em: 5 jun. 2022.

LIMA, Patrícia Medeiros de. **O micromundo dos jornalistas de dados no Brasil: carreira profissional e construção de identidade**. 2021. Dissertação (Mestrado) - Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Santa Catarina, 2021.

LOOSEN, Wiebke; REIMER, Julius; DE SILVA-SCHMIDT, Fenja. **Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the Data Journalism Awards 2013–2015**. Hamburg: Verlag Hans-Bredow-Institut, n. 41, 2017.

LOOSEN, Wiebke; REIMER, Julius; DE SILVA-SCHMIDT, Fenja. Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the Data Journalism Awards 2013–2015. **Journalism**, vol. 21, ed. 9, p. 1246–1263, 2020.

MANCINI, Leonardo; VASCONCELLOS, Fabio. Jornalismo de dados: conceito e categorias. **Revista Fronteiras**, Porto Alegre, n. 18, p. 69-82, 2016.

MOURA, Isabella Cristina. **O estado do Jornalismo de Dados no cenário luso-brasileiro**. 2018. Dissertação (Mestrado) - Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa, Lisboa, 2018.

ROGERS, Simon. Florence Nightingale, datajournalist: information has always been beautiful: May 5, 1821. **The Guardian**, 13 ago. 2010. Disponível em: <https://www.theguardian.com/news/datablog/2010/aug/13/florence-nightingale-graphics>. Acesso em: 16 maio. 2023.

ROGERS, Simon. The first Guardian data journalism: May 5, 1821. **The Guardian**, 26 set. 2011. Disponível em: <http://www.theguardian.com/news/datablog/2011/sep/26/data-journalism-guardian>. Acesso em: 16 maio. 2023.

SAMPAIO, Rafael Cardoso; LYCARIÃO, Diógenes. **Análise de conteúdo categorial: manual de aplicação**. Brasília: Enap, 2021.

TRÄSEL, Marcelo Ruschel. **Entrevistando planilhas: estudo das crenças e do ethos de um grupo de profissionais de jornalismo guiado por dados no Brasil**. 2014. Tese (Doutorado) - Faculdade de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014.