

INTERPRETAÇÕES SOBRE ALGUMAS FRAGILIDADES DO BLOG *ESTADÃO DADOS*¹

Interpretations on some fragilities of the blog Estadão Dados

Interpretaciones sobre ciertas debilidades del blog Estadão Dados

Myrian Del Vecchio de Lima

Jornalista e pesquisadora. Universidade Federal do Paraná
myriandel@gmail.com

Patrícia Pivoto Specht

Jornalista e pesquisadora. Pontifícia Universidade Católica do Rio Grande do Sul

Resumo

Este texto objetivou interpretar postagens publicadas no *Blog Estadão Dados* (ED), núcleo profissional do jornal *O Estado de S. Paulo*. Apresentam-se aportes teóricos sobre o jornalismo de dados para ser possível refletir, enquanto problema, sobre algumas fragilidades desta forma de buscar informações e produzir notícias baseadas em estatísticas e em projetos de visualização de dados. De visada qualitativa, o trabalho localiza e interpreta, por meio de cinco categorias, seis publicações postadas em novembro de 2017. Mesmo cientes de que o blog ED não é planejado para matérias analíticas aprofundadas, registramos a falta de textos que deem voz a personagens, o que nos leva à reflexão sobre as possibilidades de um jornalismo de dados que vá além do instrumental, apontando para matérias que abranjam aspectos sociais e humanos mais amplos.

Palavras-chave: Jornalismo guiado por dados. Jornalismo digital. *Blog Estadão Dados*

Abstract

This text aimed to interpret some posts published in the *Blog Estadão Dados* (ED), professional nucleus of the newspaper *O Estado de S. Paulo*. Theoretical contributions are presented on data journalism to be possible to reflect, as a problem, about some weaknesses in this way seeking information and producing news based on panting information and statistics and data visualization projects. Qualitatively, the paper locates and interprets, through five previous categories, six publications posted in November, 2017. Among the points analyzed, even aware that the ED blog is not planned for in-depth analytical materials, we verified the lack of texts that give voice to characters, which leads us to reflect on the possibilities of a data journalism that goes beyond the instrumental, pointing to matters that cover broader social and human aspects.

Key words: Data driven journalism. Digital journalism. *Blog Estadão Dados*.

Resumen

¹ Versão de trabalho apresentado no XIV Congreso de la Asociación Latinoamericana de Investigadores de la Comunicación (Alaic), na Universidad de Costa Rica, em julho de 2018.

Este texto ha pretendido interpretar algunos *posts* publicados en el *Blog de Estadão Dados* (ED), núcleo profesional del diario *O Estado de S. Paulo*². Las contribuciones teóricas se presentan sobre el periodismo de datos para ser posible reflexionar, como un problema, algunas debilidades de esta forma de buscar la información y producir noticias basadas en estadísticas y proyectos de visualización de datos. Cualitativamente, el documento localiza e interpreta, a través de cinco categorías, seis publicaciones, en noviembre de 2017. Entre los puntos analizados, incluso conscientes de que el blog del ED no está previsto para materiales analíticos en profundidad, registramos la falta de textos que dan voz a los personajes, lo que nos lleva a reflexionar sobre las posibilidades de un periodismo de datos que va más allá de lo instrumental, señalando asuntos que abarcan aspectos sociales y humanos más amplios.

Palabras clave: Periodismo basado en datos. Periodismo digital. Blog *Estadão Dados*.

1 INTRODUÇÃO

Entre as continuidades/descontinuidades do jornalismo atual marcadas pelos avanços das tecnologias digitais em rede, uma das instâncias que mais chama a atenção do mercado e da comunidade científica é a que passou a denominar-se jornalismo de dados. Assim como a checagem da informação, pedra angular do trabalho do repórter, que ao se utilizar de potencialidades tecnológicas passou a ser denominada pelo termo inglês *fact-checking*, seduzindo novas plateias, o jornalismo de dados também seduz novos ou experimentados profissionais e pesquisadores da área. Tornou-se, em anos recentes, um fenômeno midiático sobre o qual se passou a refletir e investigar.

Esse texto não pretende retomar discussões sobre os conceitos de dados, *big data*, algoritmos etc. — já há vasta bibliografia que remete às bases da informática que deram origem ao fenômeno. Adentramos diretamente à relação dados e jornalismo, no âmbito das redes digitais do terceiro milênio, levando em conta que “jornalismo de dados” ou “jornalismo guiado por dados” são expressões de difícil e variada definição. Entretanto, há certo consenso de que fatores como os avanços da tecnologia digital; a Lei de Acesso à Informação, no caso brasileiro; e a atitude investigativa dos jornalistas tornam possível e promovem o Jornalismo de Dados.

² “*O Estado de S. Paulo*, também conhecido como *Estadão*, é um jornal brasileiro publicado na cidade de São Paulo desde 1875. Ao lado de *O Globo*, *Folha de S. Paulo*, *Zero Hora*, *Correio Braziliense* e *Estado de Minas*, entre outros, forma o grupo dos principais jornais de referência do Brasil. Foi fundado com base nos ideais de um grupo de republicanos, em 4 de janeiro de 1875. Nessa época, o jornal se chamava *A Província de São Paulo* e foi o pioneiro em venda avulsa no país.” (Wikipedia, s.p.) In: https://pt.wikipedia.org/wiki/O_Estado_de_S._Paulo Acesso em agosto de 2020.

Uma definição clássica é a de Gray *et al* (2014), em manual elaborado em colaboração com especialistas de várias nacionalidades, que conceitua jornalismo de dados como um conjunto de “novas possibilidades que se abrem quando se combina o tradicional ‘faro jornalístico’ e a habilidade de contar uma história envolvente com a escala e o alcance absolutos da informação digital agora disponível” (GRAY *et al*, 2014, s.p.).

É importante ressaltar outras duas divisões conceituais: reportagens em que os dados guiam a construção da narrativa e são o próprio fundamento da pauta poderiam ser caracterizadas como Jornalismo de Dados; enquanto publicações que utilizam dados como ilustração ou apoio, sem que esses sejam o foco da narrativa, podem ser exemplos de Jornalismo com Dados (MANCINI; VASCONCELLOS, 2019, p. 69).

Este artigo toma como *pressuposto* que, na prática desta forma de se fazer pesquisa jornalística e organizar a informação, é preciso ir além da coleta e da raspagem de dados que vão permitir a construção de pautas e matérias informativas investigativas. Este “ir além” significa adentrar no campo da análise dos fatos e fenômenos sociais e humanos que emergem dos dados levantados, produzindo um jornalismo guiado por dados que se desdobre em histórias humanizadas e com implicações sociais.

Tal ação que requer, além da exposição dos dados, o exame atento e relacionado, não apenas com outros dados, mas com aspectos e fenômenos humanos e socioeconômicos, é uma premissa que julgamos necessária para que tais práticas possam ser jornalísticas, entendendo-se jornalismo a partir da visão de Traquina (2011), que o coloca como o relato à sociedade dos fatos de interesse público. Ao colocar as bases de dados no âmbito desse entendimento, pode-se refletir que, conforme Stray (2014, s.p.), “o jornalismo começa e termina no mundo da experiência humana, e a matemática é apenas uma parte no meio.”

Assim, perguntamos sobre algumas fragilidades desta prática jornalística, que mais do que um gênero, se tornou imprescindível no jornalismo atual, em especial uma destas debilidades: a dificuldade ainda observada de inserir o jornalismo mais humanizado na produção da área, em especial por meio da inserção de personagens, aqui entendidos como atores sociais consultados com falas inseridas nas matérias e que vão além das fontes tradicionais declaratórias, pois ilustram uma história na qual se inserem.

A partir do exposto, este texto — de visada qualitativa — trabalha interpretativamente algumas matérias de jornalismo de dados (JD), mais adiante definido, publicadas no blog do *Estadão Dados* (ED) (<https://www.estadao.com.br/infográficos>), núcleo do jornal *O Estado de São Paulo*, com o objetivo de observar os limites desta forma de garimpar informações

correlacionadas e apresentar notícias. No blog em questão, são publicados gráficos e animações sobre temas do noticiário do dia, além de cruzamentos de dados e análises feitas por uma equipe formada por jornalistas e programadores. Ao ser criado, a ideia era a de disseminar o jornalismo de dados em toda a redação do *Estadão*, colaborando com repórteres e editores em determinadas apurações, além de dar suporte no uso de softwares e ferramentas online. Para examinar a produção jornalística a partir dos dados publicados, elegemos como pontos centrais a acessibilidade para o leitor do jornal, a contextualização jornalística e o adensamento social/humano do conteúdo, entre outros pontos.

Ao final da análise, concluímos que blog poderia expandir suas matérias, avançando para reportagens humanizadas e socialmente valorizadas, indo além de servir como base para outras possíveis matérias das diferentes editorias do jornal que o abriga.

2 DO JORNALISMO DE PRECISÃO AO JORNALISMO GUIADO POR DADOS

Rodrigues (2015) lembra que, já no início dos anos 1970, nos Estados Unidos, o jornalista Philip Meyer (1973) desenvolveu um pensamento sobre o que chamou de “jornalismo de precisão”, ao enfatizar a necessidade de se usar na prática profissional métricas quantitativas baseadas em dados para que as informações das notícias fossem mais precisas e menos apoiadas em visões pessoais das fontes e repórteres. Evidentemente, tais dados não provinham de redes digitais, mas eram obtidos por meio de metodologias científicas. É recorrente, no campo profissional e no ensino, que ao se apresentar dados numéricos nas matérias jornalísticas obtém-se mais legitimidade aos olhos do leitor, em especial quando os dados advêm de fontes institucionais ou oficiais (RODRIGUES, 2015).

Atualmente, os dados, sejam aqueles incorporados a uma notícia ou dos quais decorre uma pauta para produzir uma reportagem, podem ser coletados/aspirados e “raspados”³, por meio de instrumental de pesquisa e utilização de softwares e técnicas digitais performáticas. Como as informações se disseminam por vários sites e em documentos PDF de empresas privadas, do governo e do terceiro setor, a raspagem de dados é muito oportuna para o

³A raspagem de dados é uma técnica em que um programa de computador extrai informações de uma interface feita para a leitura humana. No jornalismo, essa técnica é aplicada para extrair informações de páginas da Web e de documentos que não são organizados de forma estruturada. (ESCOLA DE DADOS, s.p).

jornalismo investigativo e para outras áreas jornalísticas especializadas: organizadas em formato estruturado, essas informações podem “descortinar possíveis irregularidades, correlações, levantar questões importantes e abrir caminhos para a produção de serviços e ferramentas úteis, educativas ou de entretenimento.” (ESCOLA DE DADOS, s.p.).

Muitos pesquisadores consideram a área uma nova forma de linguagem, uma estética de dados e suas formas de apresentação jornalística, tal qual se poderia falar de uma estética televisiva ou cinematográfica; neste sentido consultamos Barbosa (2008), que entende uma base de dados (BDs) a partir da visão de Lev Manovich, como uma forma cultural simbólica.

Coube a Lev Manovich (2001), reconhecer as bases de dados como uma forma cultural simbólica por ele denominada *computer era* ou *era do computador*. Ele foi um dos primeiros a sistematizar elaborações teóricas acerca da ubiquidade das BDs, percebendo-as como lógica e estrutura para forjar boa parte dos objetos da nova mídia. Ao argumentar em favor das BDs como uma forma cultural simbólica, Manovich a apresenta como aquela que está para a era do computador, assim como a novela e o cinema estiveram para a era moderna (BARBOSA, 2008, p.3).

Na corrente representada por Träsel (2014), as evoluções e apropriações tecnológicas têm ligação direta com o jornalismo de dados, cujas práticas investigativas estariam alicerçadas nas possibilidades instrumentais dos aparatos tecnológicos. Entretanto, o uso de uma ferramenta digital também está profundamente conectado às inserções culturais e cognitivas dos indivíduos, e não apenas à disponibilização do instrumental técnico ou, no caso do jornalismo, à disponibilização direta de dados:

Definir o JD pela dimensão tecnológica somente, por exemplo, implicaria em aceitar um argumento de natureza determinística: uma vez adotado o aparato técnico, estaríamos diante de uma necessária mudança no fazer jornalístico, no caso, uma mudança inexorável rumo ao JD. Nesse sentido, como a internet e os softwares tornaram-se aparatos acessíveis às redações, logo todo jornalismo com dados seria, necessariamente, Jornalismo de Dados, o que não nos parece correto (MANCINI, VASCONCELLOS, 2016, p. 71).

Há uma interpretação do senso comum em classificar o jornalismo de dados como uma especialização, na qual se requer do profissional jornalista habilidades híbridas entre seu campo original e as técnicas da informática, idealizando-se um profissional “completo” e, portanto, mais eficiente e hábil do ponto de vista técnico, mas também mais sobrecarregado ao assumir multitarefas que poderiam ser conduzidas em equipes multidisciplinares.

No Brasil, as pesquisas pioneiras de Elias Machado, publicadas entre 2003 e 2007, se tornaram referência na área com seu entendimento do jornalismo de dados como forma cultural típica da sociedade em redes, apresentando as funções de formatação para estruturar a informação, de suporte para sustentar narrativas jornalísticas e de memória, para resgate dos conteúdos publicados. (RODRIGUES, 2015).

Stray (2014) organiza a prática do jornalismo de dados em quatro estágios: quantificação, análise, comunicação e ação. No primeiro, como “os dados não são algo que existe na natureza” (STRAY, 2014, s.p.), é preciso transformar a realidade em *dados quantificáveis*, como, por exemplo, transformar o número de pessoas desempregadas em dados sobre desemprego. E esse processo de produção de dados, para o autor, envolve homens e máquinas decidindo/julgando sobre informações e categorias. No estágio seguinte, o da *análise*, os dados são transformados em conhecimento. Aqui é onde o jornalismo de dados se aproxima mais da ciência reprodutível e se baseia em matemática, estatística e lógica. Os jornalistas têm o dever de realizar análises corretas, o que pressupõe que o jornalismo de dados exija conhecimentos técnicos específicos.

A partir da análise e da produção de conhecimento, justifica Stray (2014), é preciso que haja uma *comunicação pública/social* para uma audiência, de forma que os dados extraídos da realidade e sua compreensão sejam capazes de atuar como agentes transformadores (*ação*). E essa comunicação é mais eficiente, muitas vezes, se os dados forem conectados a histórias humanas que trazem vida e profundidade aos números (STRAY, 2014).

2.1 Metodologia e análise qualitativa: um olhar sobre o *Estadão Dados*

Criado em maio de 2012, o blog *Estadão Dados* foi concebido para “capturar e tratar informações usando técnicas estatísticas, algoritmos e formas visuais de apresentação. Além de organizar os dados, o produto final permite, na internet e nos tablets, que os usuários façam suas próprias análises. Para o jornal impresso, gera infográficos de última geração” (ESTADÃO, 2012). Escolhemos o *Estadão Dados* pelo fato de, no período de realização da pesquisa, o núcleo mantinha uma equipe de Jornalismo Guiado por Dados (JGD) formalmente constituída, destinada ao tratamento e divulgação de conteúdo produzido a partir de dados, mesmo cientes de que não era objetivo do blog produzir reportagens com base em seus próprios dados, o que consideramos como um limite jornalístico intrínseco ao próprio blog.

Para compreender a apreensão e produção de conteúdo do blog, coletamos e analisamos material publicado no ambiente do ED em novembro de 2017 (material que continuava disponível em setembro de 2020). É importante registrar, a seguir, a parte descritiva do objeto empírico em exame, uma vez que esta etapa faz parte de qualquer procedimento metodológico. Em 9 de novembro de 2017, a *homepage* do canal apresentava oito seções, descritas no Quadro 1, a seguir:

Seções	DESCRITIVO
Página Inicial	Funciona como capa, com destaque para gráficos publicados em outras seções.
Sobre	Informações sobre a estrutura e missão do blog Estadão em Dados
Gráfico do Dia	Publicações de temáticas variadas: por exemplo, no período de coleta: um mapa dos estrangeirismos nos nomes de edifícios de São Paulo; um gráfico das mortes por câncer no Brasil, por região; tabelas de eleição.
Permanentes	Espaço para “séries estatísticas que serão constantemente atualizadas” (ESTADÃO, 2017): material datado no espaço e no tempo, como gráficos de intenção de voto e desempenho de candidatos de eleições, e dados sobre a avaliação do governo federal.
Séries	Focadas em temáticas específicas: gráfico sobre a mortalidade infantil no país, com explicação sobre as relações e as análises; além de mapas a respeito da realidade da cidade de São Paulo, com dados, por exemplo, dos bairros onde os moradores trabalham prioritariamente em casa, os locais com mais ou menos religiosos ou o gasto em deslocamento para o trabalho.
Eleições	No período da coleta, apresentava 10 conteúdos, a maioria sobre a eleição presidencial de 2014 e a disputa, no segundo turno, entre os candidatos Dilma Rousseff e Aécio Neves. Outros conteúdos exibiam mapas com o roteiro das viagens dos presidentes e como o projeto Bolsa Família influenciou a eleição de Dilma. Outros gráficos abordavam a intenção de voto para presidente, em 2015, e o encolhimento do PT nas eleições municipais de 2016.
96Xsp	O foco é exclusivamente os 96 distritos da cidade de São Paulo. Os mapas foram elaborados a partir de tabulações especiais censitárias feitas pelo Ibope em parceria com o <i>Estadão</i> , com títulos como: “Moto é mais comum nas zonas Sul e Norte” e “Posse de carro ainda é dobro nas áreas mais ricas”.
Projetos Especiais	Os principais projetos desenvolvidos pelo ED, com destaque para gráficos e mapas sobre política, cotas, religiões e biomas. Verificamos também, no período de coleta, uma visualização relativa a incêndio na boate Kiss ⁴ , ocorrido em Santa Maria (RS) 2013.

Quadro 1– Descritivo das seções do blog Estadão Dados (09/09/2017).

Fonte: Blog *Estadão Dados*, organizado pelas autoras.

⁴ O incêndio na boate Kiss matou 242 pessoas e feriu 680 outras numa discoteca de Santa Maria, no Rio Grande do Sul, Brasil, na madrugada de 27 de janeiro de 2013. A tragédia foi causada por um sinalizador disparado no palco em direção ao teto por integrante da banda que se apresentava.

A *Página Inicial*, que funciona como capa e destacava gráficos publicados em outras seções, e a sessão *Sobre*, com informações sobre a estrutura e missão do ED, não foram consideradas para análise. Os dados foram coletados das outras seis seções e, em função de o material disponível ser numeroso, foram analisadas as características da publicação mais recente de cada uma das seções, a saber: Seção *Gráfico do Dia*: “São Paulo, uma cidade dos anos 70” (16 abr. 2017). Seção *Permanentes*: “Veja o desempenho dos candidatos na média dos institutos” (20 ago. 2014). Seção *Séries*: “Uma análise visual do crescimento da economia” (30 ago. 2013) Seção *Eleições*: “A novidade mais relevante das eleições de 2016 até agora é o encolhimento do PT” (2 set. 2016). Seção *96Xs*: “Mais moto no bairro aumenta risco de morrer no trânsito” (30 jun. 2013). Seção *Projetos Especiais*: “Basômetro” (data de publicação não informada junto ao conteúdo). Cada um dos conteúdos foi analisado no seu aspecto global, ou seja, com todos os gráficos, comentários, links e textos relacionados ao material, a partir das categorizações propostas e outros aspectos que emergem para além desta classificação. Observe-se as diferenças de atualizações de algumas seções, sendo a publicação mais atualizada a da seção *Gráfico do Dia*; e a menos atualizada a seção *96Xs*.

Acentue-se que uma das pesquisadoras deste trabalho fez a observação das características de cada seção, visualizada por meio de uma navegação orientada no blog e coletou, entre os dias 13 e a 15 de novembro de 2017, o conteúdo então mais recente, com base em categorização prévia a ser explicada logo adiante. A seguir, a segunda pesquisadora deste trabalho, realizou as mesmas operações e checkou todas as informações coletadas para verificar algum possível erro ou omissão.

Para examinar a produção jornalística a partir dos dados publicados, a acessibilidade para o leitor do jornal, a contextualização jornalística e o adensamento social/humano do conteúdo, elegeram-se cinco categorias analíticas prévias (abreviadas por C), elaboradas com base na revisão teórica realizada e em uma atenta observação/leitura do blog à moda da chamada “leitura flutuante” (BARDIN, 2009), a saber:

C1) Reelaboração de dados: a metodologia usada para coletar, tratar e reelaborar os dados é fundamental para mostrar a transparência deste processo na produção jornalística. *Perguntamos*: há informação da metodologia utilizada para o tratamento/reelaboração dos dados?

C2) Estrutura do conteúdo: como o blog busca fornecer informações — dados tratados e reelaborados como gráficos ou outras formas de visualização —, para produção de

notícias em outros espaços do jornal, ou até mesmo em outros espaços de divulgação, *perguntamos*: o gráfico está acompanhado de texto ou *links* para material do site do veículo ou externo ao site do jornal?

C3) Interatividade: o processo interativo com o leitor — incluindo aí compartilhamento de conteúdo — é fundamental no processo de jornalismo de dados, assim como em outros procedimentos do jornalismo digital, *questionamos*: o gráfico ou mapa permite interação do leitor, ou seja, ele pode personalizar a navegação e gerar seus próprios dados? Há como comentar ou compartilhar o conteúdo?

C4) Contextualização/análise: jornalismo sem contextualização torna-se frágil e pouco relevante, assim como um agrupamento de dados sem análise, mesmo que em formato de gráficos, mapas ou infográficos; *indagamos*: há análise e contextualização dos dados feitos pela equipe do jornal ou por outra fonte consultada?

C5) Humanização: a humanização com personagens de uma situação ou história é um ponto recomendado e desejável na prática jornalística. *Perguntamos*: há personalização da reportagem por meio de histórias ilustrativas?

2.2 A análise: pouco contexto e nenhum personagem

A análise dos conteúdos publicados revelou, inicialmente, diversos níveis de complexidade no tratamento dos dados apresentados em forma de infográficos e mapas nas seis seções constituídas como *corpus* do trabalho. Para facilitar a compreensão dos resultados, os dados e suas categorizações foram agrupados a partir dos conteúdos das seis seções. Ao lado das observações está abreviada a categoria (C) à qual a análise interpretativa se refere: C1, C2, C3, C4 e C5.

1 - O primeiro conteúdo, “São Paulo, uma cidade dos anos 70”⁵ (16 abr. 2017 - Seção *Gráfico do Dia*), apresenta um mapa revelando a idade média das construções da cidade, com quatro cores representando as décadas de construção. *Não há informação acerca da metodologia utilizada no tratamento dos dados* (C1); em relação à segunda categoria, a da estrutura do conteúdo, destaca-se um link para a reportagem completa acerca do assunto, no portal do próprio *Estadão*. Na matéria, consta a informação da origem dos dados, ou seja, “levantamento inédito feito pelo *Estado* nos dados de cadastros da prefeitura para cobrança do IPTU”. Não há links para conteúdos externos ao jornal (C2). Com relação à interatividade,

⁵<https://sao-paulo.estadao.com.br/noticias/geral,sao-paulo-e-uma-cidade-dos-anos-70-14-dos-imoveis-e-daquela-decada,70001739800> (Acesso: agosto 2020).

o mapa permite que o leitor aproxime e afaste as imagens dos bairros por meio da ferramenta de *zoom*. O leitor pode também fazer comentários sobre o material, apesar de não haver nenhum postado. O usuário pode ainda compartilhar o conteúdo via Facebook, Twitter, Google +, LinkedIn ou então por e-mail, possibilidade existente em todos os conteúdos analisados (C3).

No que diz respeito à dimensão da interpretação ou contextualização dos dados contidos no mapa, verificou-se a presença de um texto, assinado por dois jornalistas do ED, com uma pequena análise dos números, concluindo que na década de 1970 São Paulo viveu seu primeiro grande *boom* de verticalização (C4). Na reportagem completa sobre o assunto aparecem fontes especializadas (urbanistas e arquitetos) e personagens, como comerciantes e moradores da cidade. Contempla-se aí, de forma bem clara, a humanização dos dados, com histórias pessoais ilustrando a realidade expressa em números no mapa (C5). Ressalta-se, porém, que a reportagem não se encontra no ambiente do blog, mas no portal do *Estadão*. O link do material está junto do mapa, abaixo do texto de apresentação.

2 - O segundo conteúdo “Veja o desempenho dos candidatos na média dos institutos” (20 ago. 2014 – seção *Permanentes*) apresenta quatro gráficos produzidos a partir de dados de pesquisas sobre a eleição presidencial de 2014. Assim como no primeiro conteúdo, não há informação sobre metodologia no tratamento dos dados, mas os institutos de pesquisa que geraram os dados primários são informados abaixo de cada gráfico, que podem ser visualizados pelo leitor — mas não há link para reportagens relacionadas (C1 e C2). O destaque fica por conta da categoria interatividade, já que os gráficos apresentaram o maior número de comentários dos leitores entre os conteúdos do *corpus* examinado. Há 19 comentários, sendo que três deles são respostas de Rodrigo Burgarelli, que integrava a equipe do ED, o que indica que a presença do jornalista interagindo com os leitores pode estimular a participação (C3) (Figura 1).

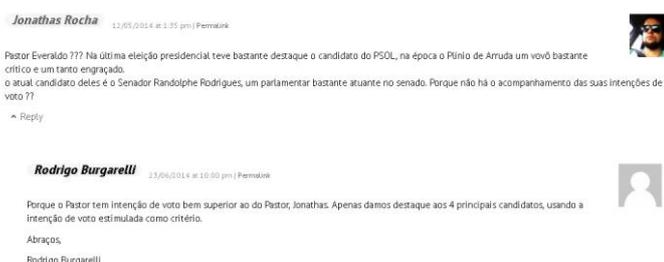


Figura 1: Jornalista integrante do ED interage com leitores no espaço dos comentários.

Fonte: blog *Estadão Dados*.

Em relação à contextualização e análise dos dados, foram encontradas apenas breves explicações, espécies de legendas, sobre alguns elementos dos gráficos. Não há textos ou explicações contextualizadas sobre o desempenho dos candidatos, refletido nos números apresentados (C4). Não há também nenhum tipo de história ou ilustração acompanhando o conteúdo (C5).

3 - A falta de contextualização também está presente no conteúdo “Uma análise visual do crescimento da economia” (30 ago. 2013- seção *Séries*), com cinco gráficos/tabelas sobre o desempenho da economia em 2013 (C2). Abaixo de cada gráfico há análises superficiais sobre as tendências de avanço ou retrocesso do crescimento, sem que se abordem as razões conjunturais das mudanças (C4). Assim como não há textos analíticos, também não há o uso de personagens ilustrativos (C5). A fonte dos dados primários é indicada como o Instituto Brasileiro de Geografia e Estatística (IBGE), com elaboração do ED. Também não se indicou a metodologia utilizada para manipulação dos dados e nem link para material externo (C1). No que diz respeito à interatividade, o único recurso disponível permite o leitor posicionar o *mouse* em determinado ponto do gráfico/tabela e ver surgir a informação estatística daquele local. Não há nenhum comentário postado sobre o assunto (C3).

4 - O material “A novidade mais relevante das eleições de 2016 é o encolhimento do PT” (2 set. 2016 - seção *Eleições*) é composto por dois mapas do Brasil e três gráficos, apresenta link de matéria da editoria de Política do site, que traz explicações sobre a redução de candidaturas do PT e análise do isolamento do partido nas disputas pelas prefeituras (C2 e C4). O maior investimento em análise e a contextualização dos dados estão novamente fora do ambiente do blog. Acima de cada um dos três gráficos, porém, há um texto resumido que dá uma breve explicação dos dados (C4). Em relação à interatividade, o leitor pode fazer buscas por cidades nos mapas e aumentar a visualização do material, que compara, a partir do uso de duas cores, as cidades onde o PT lançou candidato em 2012 e em 2016 (C3). Não há indicação de metodologia no tratamento dos dados e nem de fontes (C1). Também não há comentários de leitores (C3).

5 - No caso da seção *96xSão Paulo*, é preciso considerar que o mapa eleito para análise, “Mais moto no bairro aumenta o risco de morrer no trânsito” (30 jun. 2013), integra um conjunto de 25 mapas que abordam temas variados da realidade paulistana, como concentração de migrantes, estrangeiros, mães solteiras, bairros com maior e menor concentração de celulares etc. Consta a informação de que o material foi elaborado com base em tabulações especiais da amostra do Censo 2010 (Ibope em parceria com o ED) (C1). Não

há nenhuma reportagem analítica no espaço, nem vinculada ao material por meio de links (C2). Ao navegar pela seção, verificou-se que os mesmos recursos são usados em praticamente todos os mapas, com destaque para o uso de cores (tonalidades mais fortes indicam sempre números mais elevados) e o buscador de bairros. Outro recurso interativo é que, ao clicar nos bairros, é possível visualizar informações sobre o conteúdo do mapa. No caso do gráfico sobre as motos, ao clicar sobre o bairro Moema, por exemplo, aparecem na tela a taxa de mortes no trânsito e o número de casas com motos no local. Há três links que levam a gráficos sobre a mesma temática, motos e mortes no trânsito, todos gráficos do ED. Há dois comentários de leitores (C3). Como não há matérias elaboradas sobre o tema, também não como se apresentar fontes humanizadas (C4 e C5)

6 - Na seção “Projetos Especiais”, “O “Basômetro”” (data de publicação não informada no blog), destacou-se na análise de praticamente todas as categorias, pelas possibilidades de navegação (Figura 2) e pelo conteúdo extra disponível em forma de texto, vídeos e links com análises. Em relação à elaboração dos dados, o “Basômetro” constitui uma ferramenta interativa que permite medir o apoio dos parlamentares ao governo e acompanhar sua posição nas votações legislativas. Junto do material foi publicado o texto “Entenda como foi feito o “Basômetro”, que explica a origem do projeto e dados a serem obtidos, além de dicas de navegação. Explica-se no blog que o instrumento se inspirou no projeto *Camaraws*, do Grupo de Estudos de Software Livre da Poli-USP (PoliGNU), que usou os resultados de votações na Câmara dos Deputados para avaliar o grau de semelhança entre os partidos. Um link leva o leitor à página do PoliGNU. O texto informa que os votos dos parlamentares do “Basômetro” foram obtidos nos sites da Câmara dos Deputados e do Senado, considerando-se todas as votações nominiais, as únicas em que o voto individual é computado (C1 e C2).

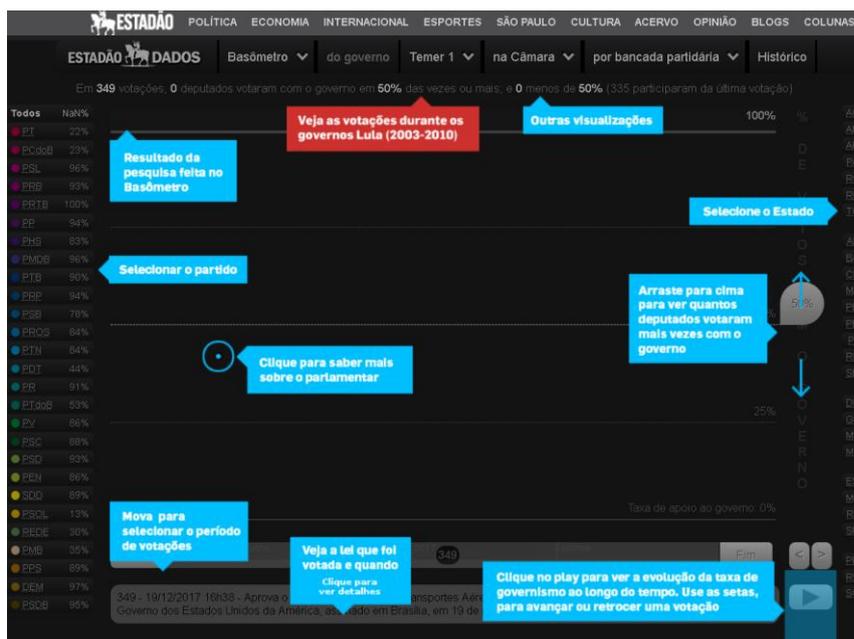


Figura 2: Recursos interativos do “Basômetro” .

Fonte: Blog *Estadão Dados*.

Em relação à interatividade, o “Basômetro” permite que o usuário faça análises próprias e compare a atuação das diferentes bancadas partidárias (C3). Destaca-se ainda a presença de 16 “Links relacionados”, a maioria com matérias analíticas geradas pelo cruzamento de dados do “Basômetro” (C4). Havia análise dos dados obtidos no “Basômetro” nesse material, sem, no entanto, a presença de personagens ilustrativos (C5). Não foi verificada a existência de nenhum link que levasse o leitor a conteúdo fora do ambiente do jornal. Criado em 2012, o instrumento estava atualizado, com dados de novembro de 2017. Um botão deslizável (*slider*) da taxa de governismo pode ser arrastado para cima (mais governista) ou para baixo (mais oposicionista). O título dinâmico no topo da tela aponta o número de parlamentares pró e contra o governo conforme a taxa e o período selecionados. Outro recurso interativo é que, além de visualizar os votos por partido, o usuário pode optar por ver os dados por votação. Como já dito, há vídeos tutoriais publicados no “Basômetro”. Um deles explica como usar os recursos da ferramenta e outro ensina o leitor a fazer um cruzamento de dados para analisar como o governo perde apoio de deputados na Câmara até 2012 (C3).

“No caso do ‘Basômetro’, houve uma preocupação nossa de não dar a ferramenta, simplesmente, sem o contexto. A gente fez várias reportagens em cima dele, então todo o contexto foi cuidado. O blog não, porque é algo mais dia a dia mesmo.” (BRAMATTI, in TRÄSEL, 2014, p. 223). Com base nesta fala de Bramatti, em 2014, enquanto editor do

“Basômetro” e do blog, assim como do que se depreende da observação dos seis conteúdos analisados do blog ED, constata-se que, o ambiente do ED efetivamente não é planejado para matérias analíticas aprofundadas. Entretanto o mesmo Bramatti ainda afirma sobre o “Basômetro”:

Mas nossos projetos mais de fôlego, apostas de reportagens mesmo, a gente não faz só o gráfico. O gráfico é uma das etapas da reportagem, que também tem texto, tem análise, tem interpretação. O contexto eu acho importante. Acho que, realmente, o jornalismo de dados não é só gráfico e visualização, é reportagem. É como entrevistar o número: tem que mostrar o que é importante, o que deve ir para o lide, o que deve ir pro sublide (BRAMATTI citado por TRÄSEL, 2014, p. 223).

As reportagens completas vinculadas aos conteúdos do blog, quando existentes, provavelmente contemplem o aspecto de ir além de uma breve interpretação sobre os gráficos, com contextualização, mas não foi objetivo desse trabalho analisá-las (C4). Entretanto, a constatação nos conduz à questão: para que manter o blog se não se usam, ou se seus dados são utilizados pouquíssimas vezes em matérias ou pelos leitores? Não parece fazer muita diferença, pelo baixo número de comentários ou de matérias internas do próprio *Estadão* ou externas de outras mídias jornalísticas, se o blog fosse apenas um serviço interno do jornal. Também não foi localizado no blog nenhum tipo de texto ou análise ilustrado por personagens (C5), o que leva novamente à observação feita no parágrafo acima.

Em relação às fontes dos dados manipulados pelo jornal para a produção dos gráficos e mapas, cinco dos seis materiais analisados revelaram de onde se originou o conteúdo bruto, ou seja, IBGE, outros institutos de pesquisa, prefeitura etc. (C1 e C2). O que não se fez foi publicar junto ao conteúdo o material bruto, prática seguida por alguns jornais com tradição em jornalismo guiado por dados (JGD), como o *Guardian Data Store* (<https://www.theguardian.com/data>), seção do website do jornal britânico *The Guardian*, na qual a redação oferece ao leitor a possibilidade de acessar e reutilizar as bases de dados completas usadas na produção do noticiário.

No caso do JGD, o acesso às bases de dados originais permitiria à audiência e a outros repórteres realizar suas próprias análises e julgar a validade das conclusões deles derivadas por uma determinada peça jornalística. Ao criar empecilhos à revisão por pares, estariam falhando em observar os princípios da objetividade mecânica e, principalmente, impedindo a construção de uma objetividade disciplinar. [...] a falha em incentivar a revisão por pares se mostra um fator de geração de tensões entre os profissionais envolvidos com essas práticas (TRASEL, 2014, p. 128).

Da mesma forma, a explicação da metodologia de busca, limpeza e tratamento dos dados é fundamental no processo de transparência jornalística para agregar credibilidade à prática (C1) — mas nenhuma das publicações analisadas mencionou estas etapas do processo do jornalismo de dados, anteriores à produção de gráficos e outras visualizações.

No que diz respeito às ferramentas de interatividade (C3), verificou-se que as possibilidades disponíveis ao leitor são variadas. Enquanto alguns conteúdos permitem que o usuário faça combinações e extraia informações inéditas, outros oferecem possibilidades mais limitadas, como aproximar e afastar a imagem do mapa e obter dados estatísticos, posicionando o *mouse* no local desejado. Mas, de modo geral, as formas de interatividade ainda são bastante limitadas, o que se reflete no baixo número de comentários dos leitores, com exceção do “Basômetro”, enquanto projeto especial que apresenta um diferencial perceptível no cuidado para gerar links, matérias e maior interação com o leitor, o que também se destaca na fala de Bramatti (in TRASEL, 2014) citada anteriormente.

3 CONSIDERAÇÕES

Não foi intenção deste trabalho abranger um período amplo de publicações do blog *Estadão Dados*. Mas, sim apontar, algumas possibilidades de caminhos para pesquisar características de espaços nas mídias jornalísticas dedicados ao Jornalismo Guiado por Dados. Assim, este estudo tentou marcar a importância do jornalismo contemporâneo “ir além” do instrumental que permite a obtenção e agrupamento de dados, ou mesmo de sua reelaboração, utilizando o potencial das tecnologias para oferecer melhor contextualização ao leitor sobre o fato em investigação; proporcionar clareza sobre como as informações foram obtidas e tratadas; fornecer possibilidades para o leitor aumentar sua interpretação sobre o assunto em pauta, cruzar os dados obtidos com outras fontes ou jornalistas; e, finalmente, chegar à humanização da reportagem, devolvendo a ela o caráter social, cultural e humano que toda boa estória jornalística merece.

Pode-se questionar que essa valoração que confere foco ao humano dificilmente caberia em narrativas mais curtas, no formato de notícias diárias. Por outro lado, não é possível se falar em jornalismo, do ponto de vista de sua função social, sem o colocar como expressão do humano, ao dar sentido a fatos e acontecimentos que permeiam a realidade social e o cotidiano das populações. E isso continua a valer mesmo em época da possibilidade de máquinas (robôs) produzirem notícias a partir de conteúdos, informações e dados disponíveis em rede.

Cientes de que a análise se restringiu a um núcleo de base de aproveitamento de dados, o blog *Estadão Dados*, que pode expandir a organização informativa para matérias mais aprofundadas e reportagens em outros espaços do jornal impresso ou do site *Estadão.com*, pudemos, entretanto, concluir sobre alguns pontos relativos ao uso da técnica de jornalismo de dados no blog em questão. Detectamos três fragilidades principais: falta de personagens; falta de adensamento humano e social; e ainda pouca clareza em termos de metodologia, elemento fundamental na prática do jornalismo. Estas três fragilidades nos fizeram finalizar a análise com a reflexão de que o *Estadão Dados*, na época em que foi examinado neste estudo, poderia ser apenas um espaço virtual destinado a aproveitamento interno entre os profissionais do próprio jornal, de forma a servir como base para gerar reportagens externas a ele mais aprofundadas e histórias humanizadas.

Em razão disso, não é possível deixar de refletir sobre a validade ou função jornalística de um repositório de conteúdo produzido a partir de dados que não fornece, em sua maior parte, elementos para a compreensão global do material informativo, como análises contextualizadas e personagens, além de muitas vezes não explicitar a metodologia utilizada. A função do ambiente, no período do estudo, pareceu ser muito mais a de reunir o material produzido pelo ED do que informar de maneira jornalisticamente adequada. Do ponto de vista de um jornalismo produzido de forma a gerar debate e a informar de maneira abrangente e crítica, ainda há um caminho a ser percorrido. Esse trabalho espera ter fornecido pistas nesse sentido e pretende avançar para o mesmo tipo de análise em reportagens realizadas com base em jornalismo de dados, assim como analisar outras formas de produção do jornalismo gerado por dados.

REFERÊNCIAS

ESTADÃO DADOS. **Estadão Dados une transparência e interatividade**. 12 maio 2012. Acesso em 29 out. 2017 e em 20 mai. 2021. Disponível em: <http://politica.estadao.com.br/noticias/geral,estadao-dados-une-transparencia-e-interatividade,872163>.

BARBOSA, Suzana. *Modelo JDBD e o ciberjornalismo de quarta geração*. **Grupo de Pesquisa em Jornalismo On-line**, FACOM-UFBA, 2008. Acesso em: 22 dez. 2017 e em 03 mai. 2021. Disponível em: https://www.facom.ufba.br/jol/pdf/2008_Barbosa_RedUCMx.pdf

BARDIN, Laurence. **Análise de conteúdo**. Portugal: Edições 70, 2009.

ESCOLA DE DADOS. Acesso em: 29 out. 2017. Disponível em: <https://escoladedados.org/manual/cursos/raspagem/raspagem-jornalismo-de-dados/>.



GRAY, Jonathan; BOUNEGRU, Liliana; CHAMBERS, Lucy (Orgs). **Manual de Jornalismo de Dados**, 2014. Acesso em: 18 nov. 2017. Disponível em: <http://datajournalismhandbook.org/pt/>.

MANCINI, Leonardo.; VASCONCELLOS, Fábio. *Jornalismo de Dados: conceito e categorias*. **Fronteiras**, São Leopoldo: Unisinos, n. 18, jan./abr. 2016. Acesso em: 31 out. 2017. Disponível em: <http://revistas.unisinos.br/index.php/fronteiras/article/viewFile/fem.2016.181.07/5300>.

RODRIGUES, Kelly De Conti. *Jornalismo de Dados: influência da construção narrativa no agendamento midiático*. In: BULHÕES, Marcelo.; MORAIS, Osvando J. **Ciências da Comunicação: Circularidades Teóricas e Práticas**. Sarapuí, SP: OJM Casa Editorial, 2015.

STRAY, Jay. *The Data Journalist's Eyes: an Introduction*. **Tow Center for Digital Journalism**, 18 set. 2014. Acesso em: 30 nov. 2017. Disponível em: <https://towcenter.org/the-data-journalists-eye-an-introduction>.

TRAQUINA, Nelson. **Teorias do jornalismo: a tribo jornalística – uma comunidade interpretativa transnacional**. Florianópolis: Insular, 2008.

TRÄSEL, Marcelo. **Entrevistando planilhas: estudo das crenças e do ethos de um grupo de profissionais de jornalismo guiado por dados no Brasil**. 2014. 315 f. Tese (Doutorado) – Faculdade de Comunicação Social, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brasil, 2014.

Original recebido em: 20 de setembro de 2018

Aceito para publicação em: 09 de agosto de 2020

Myrian Del Vecchio de Lima

Jornalista. Doutora em Meio Ambiente e Desenvolvimento pela Universidade Federal do Paraná (UFPR), com pós-doutoramento em Jornalismo Digital, realizado no Incom, da Université Lyon2, França. Professora do Programa de Pós-Graduação em Comunicação da UFPR.

Patrícia Pivoto Specht

Jornalista. Doutora em Comunicação Social pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS).



Esta obra está licenciada com uma Licença
Creative Commons Atribuição-NãoComercial-CompartilhaIgual 4.0 Internacional

