

UM PERCURSO HISTÓRICO DO JORNALISMO DE DADOS

A HISTORICAL JOURNEY OF DATA JOURNALISM

UN RECORRIDO HISTÓRICO POR EL PERIODISMO DE DATOS

*Cláudia Miranda Rodrigues
Leonel Azevedo de Aguiar*

Resumo: A proposta deste trabalho é apresentar uma linha histórica que traça a constituição contemporânea do jornalismo de dados a partir do uso, no século 19, de pesquisas empíricas oriundas das ciências sociais por organizações sociais e pela imprensa daquela época. O artigo, que utiliza o método da pesquisa bibliográfica, busca delinear os pontos que tangenciam a emergência do jornalismo de dados, desenvolvendo o percurso trilhado pelas práticas profissionais em nomenclaturas como jornalismo de precisão e reportagem assistida por computador (RAC). É possível concluir que o rigor no relato dos acontecimentos e a transparência no levantamento das informações perpassam a evolução das técnicas de apuração relacionadas ao jornalismo de dados. Além disso, a datificação da sociedade e a publicização de informações digitalizadas em bases de dados servem de estímulo para impulsionar as diversas nomenclaturas existentes nesta especialidade profissional.

Palavras-chave: Teorias do Jornalismo. Jornalismo. Jornalismo de Dados. Jornalismo de Precisão. Reportagem com Auxílio de Computador.

Abstract: The purpose of this work is to present a historical line that traces the contemporary constitution of data journalism based on the use, in the 19th century, of empirical research from the social sciences by social organizations and the press of the time. The article, which uses the bibliographical research method, seeks to outline the points that touch on the emergence of data journalism, developing the path taken by professional practices in nomenclatures such as precision journalism and computer-assisted reporting (CAR). It is possible to conclude that rigor in reporting events and transparency in gathering information permeate the evolution of investigative techniques related to data journalism. In addition, the datafication of society and the publication of digitized information in databases serve as stimuli to boost the various nomenclatures that exist in this professional specialty.

Keywords: Journalism theories. Journalism. Data Journalism. Precision Journalism. Computer-Assisted Reporting.

Resumen: La propuesta de este trabajo es presentar una línea histórica que remonta la constitución contemporánea del periodismo de datos a partir de la utilización, en el siglo XIX, de la investigación empírica de las ciencias sociales por las organizaciones de la sociedad y la prensa de la época. El artículo, que utiliza el método de investigación bibliográfica, pretende esbozar los puntos que tocan la emergencia del periodismo de datos, desarrollando el camino recorrido por las prácticas profesionales en nomenclaturas como periodismo de precisión y reportaje asistido por ordenador (RAC). Es posible concluir que el rigor del relato de los hechos y la transparencia en la recogida de información impregnan la evolución de las técnicas de investigación relacionadas con el periodismo de datos. Además, la datificación de la sociedad y la publicación de información digitalizada en bases de datos sirven de estímulo para potenciar las distintas nomenclaturas que existen en esta especialidad profesional.

Palabras clave: Teorías del Periodismo. Periodismo. Periodismo de Datos. Periodismo de Precisión. Reportaje Asistido por Ordenador.

1 INTRODUÇÃO

O jornalismo de dados resulta de um processo que se relaciona, de forma intrínseca, com a internet, a interação propiciada pela *web*, a cultura digital e o jornalismo *online*. Entretanto, ao considerar a emergência do jornalismo de dados, Anderson (2018) empreende uma investigação com base no método genealógico de Michel Foucault que localiza a origem da cultura contemporânea de dados no século XIX, imbricada com práticas das Ciências Sociais. Neste artigo, em diálogo com a obra de Anderson, pretendemos pontuar as clivagens que respaldaram a evolução de uma “especialidade” jornalística que mobiliza a atenção das pesquisas em Teorias do Jornalismo. A partir desse foco, vamos delinear os pontos que tangenciam a emergência do jornalismo de dados, para, por fim, apresentar nomenclaturas em torno dessa prática, que abarca o jornalismo de precisão, a RAC (Reportagem com Auxílio do Computador) e o jornalismo computacional, na perspectiva do contexto histórico e social. A metodologia utilizada está baseada na pesquisa bibliográfica da literatura especializada sobre o jornalismo contemporâneo, especialmente o jornalismo de dados.

Enquanto Fink e Anderson (2015, p. 470) defendem que “o jornalismo de dados é, em última instância, um termo profundamente contestado e simultaneamente difuso”, Hermida e Young (2019) apontam que sete termos foram nomeados em pesquisas recentes para essa prática: jornalismo guiado por dados; reportagem guiada por dados; reportagem com auxílio de computador; jornalismo de base de dados; jornalismo quantitativo; jornalismo computacional e jornalismo algorítmico. O que nos interessa destacar é que, para além do debate em torno das nomenclaturas, foi possível concluir que o rigor e a transparência perpassam a evolução das técnicas de apuração relacionadas ao jornalismo de dados, além da precisão ser um fator que contribui para elevar o grau de credibilidade (AGUIAR; RODRIGUES, 2019). Além disso, a datificação da sociedade e a publicização de informações digitalizadas em bases de dados servem de estímulos para as diversas

nomenclaturas existentes nesta especialidade profissional.

O artigo está dividido em três seções, além da introdução e da conclusão. A primeira seção remonta uma linha histórica da utilização de dados pelos movimentos sociais e a imprensa no século XIX. A seção seguinte aborda o jornalismo de precisão e a proposta de Philip Meyer, um dos jornalistas pioneiros no uso dos métodos das Ciências Sociais e do computador na produção de reportagem. A terceira seção discorre sobre o jornalismo de dados, suas práticas e exemplos notáveis.

2 PERSPECTIVA HISTÓRICA

No livro *Apostles of certainty: data journalism and the politics of doubt*, Anderson (2018) mergulha na investigação dos movimentos da Era Progressista nos Estados Unidos, especialmente na produção de trabalhos sociais do *Men of Religion Forward Movement* (MRFM) que, no século XIX, adotou práticas de coletar informações quantitativas sob o âmbito de um abrangente programa de pesquisa social.

Através do método foucaultiano, Anderson (2018) propõe uma genealogia que constrói uma ponte entre o passado, marcado pela visão positivista, e o presente, no qual se consolida o jornalismo de dados: se este termo possui uma conotação contemporânea, o uso de dados no jornalismo remonta o final do século XIX e, em diferentes momentos, está atrelado à tentativa de construir narrativas verdadeiras. Portanto, a relação entre jornalismo e uso de dados está fundada em um modo normativo: a cultura da verdade – que rege o campo jornalístico – foi impactada pelo trabalho social no momento em que a Sociologia estabeleceu seus métodos.

Técnicas parecidas com as jornalísticas, que Anderson (2018) denomina *social survey reportage* – reportagem para pesquisa social –, embasaram levantamentos sociais realizados em centros urbanos no início dos anos 1900. O método de coleta de dados e estatísticas se constituiu em uma lente própria que permitia a compreensão da sociedade. O interesse crescente em entender

as causas da pobreza estimulou movimentos progressistas na busca dos motivos para a promoção das mudanças sociais necessárias. O *Social Survey Movement* (SSM), do qual fazia parte o MRFM – um movimento religioso engajado nas reformas sociais –, acreditava na força de veracidade dos dados quantitativos e no alcance e impacto social da visualização.

A pesquisa social (...) é um fenômeno que se encontra na encruzilhada de uma variedade de caminhos epistemológicos importantes: um impulso progressivo e reformador em segmentos da cultura política dos Estados Unidos, uma crença na importância da coleta de evidências empíricas *in loco*, e uma obsessão por contar e com o poder dos números para iluminar importantes verdades. No momento em que surgia a reportagem para pesquisa social, as fronteiras de uma variedade de ocupações que geram conhecimento também estavam em fluxo (...). Por décadas, em outras palavras, as linhas divisórias entre jornalismo, reforma social, coleta de dados e ciências sociais eram muito menos sólidas do que eram na parte final do século XX (ANDERSON, 2018, p. 21).

Dentro dessa ótica, o MRFM vinculou dados quantitativos a métodos jornalísticos com a intenção de aprimorar sua credibilidade a partir da difusão de relatórios nos jornais americanos. O trabalho de coleta e visualização de dados a partir de “técnicas empíricas, gráficos e mapas” visava a ampliar a mobilização e conquistou significativo espaço na imprensa, em jornais como *The New York Times*, *El Paso Herald* e *San Francisco Call*. Entretanto, Anderson (2018, p. 42) observa que, mais do que publicar visualizações, os jornais as relacionavam com o conteúdo em texto.

Em 1895, edições do *New York Times* publicavam matérias com nomes e números em listas em visualização gráfica rudimentar, mas de extrema clareza, que se sofisticaria nas décadas seguintes (ANDERSON, 2018, p. 76). Na Inglaterra, em 1821, o *Manchester Guardian* – embrião do *The Guardian* – já havia publicado reportagem, com uma tabela de dados em formato parecido com a do jornal americano, que revelou o número preciso de crianças atendidas por escolas públicas nas cidades de Manchester e Salford, no norte da Inglaterra – o que dimensionava o tamanho da pobreza (ROGERS, 2013, p. 60).

Duas publicações emblemáticas no percurso genealógico traçado por Anderson (2018) são o *Survey Graphic* (1923-1952), originalmente destinado a movimentos sociais, que buscou nas visualizações uma forma de desencadear mobilização e ação social – e a revista *Fortune*, considerada precursora do jornalismo de dados, ao construir reportagens dentro de um contexto interpretativo que buscava responder a questões além da retórica do *lead*, relacionando o porquê e o significado do acontecimento. Os formatos vinculados ao jornalismo de dados, entretanto, só foram abraçados de forma mais abrangente pelas organizações jornalísticas a partir de 1980 e 1990 com a evolução tecnológica que favoreceu a publicação de *design* gráfico mais sofisticado.

Sabemos que, com o surgimento da *penny press*, no final do século XIX – que se vincula à emergência do jornalismo de informação –, houve uma mudança no foco do trabalho jornalístico, que passa do registro para a reportagem. Neste sentido, acontece a valorização do relato em detrimento do registro, o que favoreceu o desenvolvimento de técnicas como o testemunho ocular e a entrevista. Alimentada pela demanda de notícias sobre a Guerra Civil nos Estados Unidos, essa mudança de *ethos* marca a invenção da reportagem a partir de coleta de evidências. “A *penny press* contribuiu para consolidar o novo conceito de jornalismo, que separou e valorizou o fato em detrimento da opinião, o que ajudou a efetuar a passagem de um jornalismo de opinião para um jornalismo de informação” (AGUIAR, 2008, p. 18). Entretanto, o modelo de imprensa orientada por documentos – que atendia a um público urbano interessado em negócios e finanças – nunca deixou de existir no universo da mídia norte-americana.

No processo de profissionalização, que irrompeu com o advento da Modernidade, diferente do jornalismo, as Ciências Sociais firmaram sua legitimidade enquanto ciência, buscando foco nas evidências. Neste sentido, Anderson (2018) observa como o jornalismo de precisão proposto por Philip Meyer, na virada da década de 1960 para a de 1970, marca uma proposta de

reaproximação das duas profissões. Aplicar métodos quantitativos das Ciências Sociais nas rotinas jornalísticas resume a proposição de Meyer, considerado pedra basilar para a reportagem com auxílio do computador (RAC), o jornalismo computacional, o jornalismo de dados e o jornalismo guiado por dados (CODDINGTON, 2018; PARASIE; DAGIRAL, 2013).

3 JORNALISMO DE PRECISÃO

Com o recurso de coleta de dados, amostragem e análise, em 1967, Meyer investigou conflitos ocorridos em Detroit com a intenção de identificar o perfil dos manifestantes e suas motivações. A reportagem, publicada no *Detroit Free Press*, apontou que negros do Sul dos Estados Unidos não eram maioria nos protestos – como aludiam editoriais – e que a falta de emprego, condições precárias de moradia e brutalidade policial se somaram como estopim de revolta e indignação (HOWARD, 2014, p. 11).

Na concepção de Meyer (1973/2002, p. 231-232), o jornalismo de precisão “envolve pouco mais do que a manutenção do papel tradicional do jornalista com apenas uma modesta melhora quantitativa na velocidade e precisão”. Na década de 1970, o jornalista americano já enxergava o potencial das bases de dados e do *spreadsheets* (planilhas) nas rotinas de produção, a partir de programas como *Lotus* e *Excel*, do ponto de vista da sistematização, cruzamento e recuperação de informação. Reconhecer padrões e evidências documentais e utilizar entrevistas apenas como complemento constituem o modelo criado por Meyer.

Schudson (2010) observa a influência dessa objetividade científica na emergência da era do repórter, no final do século XIX, quando o jornalismo é influenciado pelos ideais da Filosofia positivista, muito antes da invenção do computador. Em 1922, Lippman já havia expressado a relevância das “máquinas de registro” utilizadas na Bolsa de Valores que asseguravam divulgação de notícias com “precisão confiável”. A preocupação de Lippman com a ingerência da propaganda na veracidade das notícias se equipara a de

Meyer com a falta de objetividade e uma forma de verdade subjetiva que desponta com o jornalismo interpretativo (ou literário). Neste sentido, Anderson (2018, p. 120) destaca que, em oposição a esse estilo conduzido por jornalistas como Tom Wolfe e Gay Talese, o jornalismo de precisão tem “a pretensão ao rigor epistemológico” baseado em “um processo transparente extraído das ciências sociais com hipóteses testáveis”.

É o conceito do jornalismo de precisão que embasa a reportagem com auxílio do computador (RAC). A prática emergiu no final dos anos 1960 a partir de pioneiros como Philip Meyer (1973/2002) e Eliot Jaspin, que trabalharam nos jornais americanos *Postville* e no *Republican Herald*, na Pennsylvania. Com apoio do computador e o recurso da base de dados, os jornalistas podiam cruzar dados acessados publicamente para apurar reportagens. O CAR (*Computer-assisted reporting*) tornou-se ferramenta nas mãos de jornalistas investigativos (PARASIE; DAGIRAL, 2013; COHEN *et al.*, 2011; HOWARD, 2014; FLEW *et al.*, 2012; GRAY *et al.*, 2012).

O conceito da RAC foi adotado pela primeira vez através do uso de computador em reportagens realizadas em 1952, pela rede de televisão norte-americana CBS, para prever o resultado das eleições nos Estados Unidos. A ubiquidade conquistada pelo computador a partir dos anos 1980 mudou a forma de conduzir investigações e resultou em denúncias de grande repercussão a partir de processamento de dados públicos (GRAY *et al.*, 2102). O próprio Meyer (1999) reconhece um certo anacronismo na nomenclatura e tece ironias: “alguém lembra do primeiro repórter que usou telefone? Ele era chamado de repórter com auxílio de telefone”? (MIELNICZUK; GEHRKE, 2017).

O jornalista, então, já enxergava o jornalismo em uma “batalha pela sobrevivência” e enfatizava o uso do computador enquanto método:

O que os praticantes do CAR estão buscando - conscientemente ou não - é um padrão mais alto para falar a verdade. Nossa resposta à era da informação tem sido aprender a gerenciar corpos maiores de informação com ferramentas analíticas cada vez mais poderosas, levando a

uma definição mais exata da verdade. Acontece que um computador é útil para fazer isso. Mas o computador em si não é o objetivo, nem define o que estamos tentando fazer. Estamos tentando empurrar o jornalismo na direção da ciência (MEYER, 1999).

A RAC situa-se dentro do âmbito do jornalismo de precisão e do jornalismo guiado por dados (HOWARD, 2014, p. 11). Na conferência anual do NICAR (*National Institute for Computer-Assisted Reporting*) realizada em Baltimore, em 2014, que reuniu cerca de mil profissionais, David Kaplan, um expoente do jornalismo investigativo, adotou o termo “jornalista de dados” como identificação por vários participantes. O programa *NICAR* foi criado em 1989 pelo *Investigative Reporters Editors* (IRE) com a intenção de fomentar jornalismo de excelência, especialmente no que diz respeito ao jornalismo de dados¹.

A agilidade, flexibilidade e autonomia propiciadas pela RAC vieram agregar precisão ao jornalismo investigativo. Cabe ressaltar que a epistemologia dessa prática prevê o relato de fatos sólidos com base em informações minuciosas a fim de confrontar a realidade de forma mais direta e completa (ETTEMA; GLASSER, 1998, p. 13). Uma das características da apuração investigativa – que alguns estudiosos consideram basicamente jornalismo produzido corretamente – é buscar múltiplas fontes e preservar a autonomia, ou seja, não reportar essencialmente com informação oficial (ETTEMA; GLASSER, 1998; AGUIAR, 2006; LOPES; PROENÇA, 2003; SEQUEIRA, 2005).

Uma das etapas do processo de apuração é definida por Ettema e Glasser (1998) pela categoria “pesando as evidências” que são obtidas a partir de documentos, gravações e de “qualquer material que capture e certifique os fatos” (ETTEMAN; GLASSER, 1985, p. 194). Um exemplo desse peso foi o *Pentagon Papers* – denúncia oriunda de um funcionário do Pentágono, Daniel Elsberg, que desmascarou ações do governo norte-americano durante a Guerra do Vietnã. O material fornecido à equipe do jornal *New York Times* envolveu um total de 7 mil cópias de um relatório composto por 14 mil páginas. Isso mostra claramente como as práticas investigativas, impulsionadas pela

RAC e pelo jornalismo de dados, facilitaram sobremaneira o trabalho jornalístico (COHEN *et al.*, 2011; GRAY, 2012; HOWARD, 2014; AGUIAR; ANDRADE, 2020).

4 JORNALISMO DE DADOS

Ao traçar a linha do tempo do jornalismo de dados, autores convergem na percepção da importância do trabalho de Adrian Holovaty, jornalista e programador americano que aplicou, no jornalismo, os dados estruturados e ajudou a difundir a nomenclatura (CODDINGTON, 2015). Fundador da *EveryBlock* – uma plataforma que conecta vizinhanças –, Holovaty inspirou a criação do *PolitiFact*, que faz *fact-checking*, e ganhou o prêmio Pulitzer em 2009. Nesse mesmo ano, a partir de dados estruturados extraídos do site do Parlamento do Reino Unido, o jornal *The Guardian* criou o *Datablog*², uma página destinada a reportagens de dados, ao publicar matéria em que a visualização revelava despesas de ministros britânicos. Em 2010, o jornal britânico iniciou a publicação do *War Logs*, a partir do vazamento de dados secretos sobre as guerras no Iraque e Afeganistão e de documentos diplomáticos americanos (*Cablegate*) liderado pelo site *Wikileaks*, do hacker e jornalista Julian Assange (LEIGH; HARDING, 2011).

Quatro anos depois, o site *FiveThirtyEight*, de Nate Silver, ganhou grande notoriedade ao produzir reportagens de dados quando jornais tradicionais como o britânico *The Guardian*, os americanos *New York Times* e *Washington Post*, o canadense *The Globe* e o argentino *La Nación* já investiam neste segmento (HOWARD, 2014; GRAY, 2012). Organizações sem fins lucrativos, como a americana *ProPublica* e o britânico *The Bureau of Investigative Journalism*, surgiram nessa esteira com propostas de dados abertos. Howard (2014, p. 8) pontua que o crescimento exponencial da dataficação na sociedade e o surgimento de novas ferramentas expandiram a adoção do jornalismo guiado por dados nas redações.

No Brasil, os jornais *Estado de S. Paulo* e *Zero Hora* criaram, em 2012,

respectivamente, o *Estadão Dados* e o *ZH Dados*. O *Globo* mantém uma seção que reúne infográficos, com alguns dados abertos, depois de descontinuar o projeto *Na Base de Dados*³. Gray, Bounegru e Chambers (2012) focalizam o trabalho do *Amigos de Januária*, realizado por um grupo de cidadãos que praticam atos de jornalismo, moradores da cidade brasileira no norte do país, que treinaram e aprenderam a fazer jornalismo de dados. *Agência Pública* e *Gênero e Número* surgiram na esteira dessa onda de popularização. Fundada em 2015, a *Fiquem Sabendo* é uma agência de dados independentes especializada na Lei de Acesso à Informação (LAI), que recebeu o *Prêmio Claudio Weber Abramo de Jornalismo de Dados*⁴ na categoria que teve como foco os trabalhos jornalísticos realizados graças à Lei de Acesso à Informação ou abertura das bases de dados públicas que prezam pela publicação do código-fonte e da metodologia adotada.

A LAI e sua versão estrangeira *Freedom of Information Act – FOIA*, aqui expressa na forma simplificada *FOI*⁵ – também constituem um fator que contribuiu para a propagação do jornalismo de dados. Muitas características inerentes ao uso da LAI – confecção de planilhas e análise de padrões e estatísticas – se entrelaçam com as práticas do jornalismo de dados e inspiraram o pensamento computacional na cultura jornalística (HEWETT, 2017, p. 11). Premiada com o *Data Journalism Award do Global Editors Network* em 2017, a reportagem do *The Globe Unfounded* é um exemplo. Foi feita a partir de 250 requerimentos via *FOI*, que somaram dados de 873 jurisdições policiais canadenses e apontaram a forma como a polícia do Canadá descartava uma entre cada cinco alegações de abuso sexual sob o rótulo de “infundada” (HERMIDA; YOUNG, 2019, p. 50).

Na materialização do jornalismo de dados enquanto prática, compreendemos que o armazenamento de informação oficial digitalizada em bancos de dados, e publicada em portais de transparência, é outro elemento constituinte dessa guinada computacional, uma visão compartilhada por Anderson (2018) e Rogers (2013). A abertura de dados oficiais em páginas como *data.gov*, *gov.br*

ou *gov.uk* se mostra de suma relevância na alavancagem desse processo. Simon Rogers (2013, p. 30) sublinha a importância do lançamento, em 2009, do portal de transparência norte-americano pelo então presidente Barack Obama com 378.529 dados brutos.

Ao discorrer sobre o pensamento computacional, Flew *et al.* (2012, p. 158) estabelece que essa perspectiva presume métodos baseados mais “na conceituação do que na programação”. É a capacidade de abstração em diferentes níveis que permite solucionar problemas; na prática, a inteligência computacional desenvolvida pela ciência da computação aprimora gestões de coleta e criação de conhecimento dentro de sistemas mais complexos. Por isso, Flew *et al.* (2012, p. 159-160) destacam que as técnicas computacionais melhoram a qualidade do processo de investigação e possibilitam identificação de padrões com mais precisão a partir do cruzamento de dados obtidos com maior velocidade de busca e recuperação.

A transformação em notícia do vazamento promovido pelo *Wikileaks* não seria possível fora dessa perspectiva. O *War Logs* revelou 109.032 mortes registradas desde 2003, entre as quais 66.081 civis (LEIGH; HARDING, 2011, p. 130)⁶. Apenas sobre incidentes ocorridos na guerra no Iraque, foram identificados 391 mil documentos entre 2004 e 2009 (ROGERS, 2013). A cobertura – uma parceria entre *The Guardian*, *Der Spiegel*, *New York Times*, *El País*, *Le Monde* e *Wikileaks* – começou com uma pergunta: “Você é bom com planilhas, não é?” (ROGERS, 2013, p. 70). Uma equipe que incluiu desenvolvedores criou um banco de dados simples a partir do qual os repórteres facilmente acessavam informação para criar histórias. No livro *Wikileaks: inside Julian Assange’s War on Secrecy*, Luke Harding e David Leigh do *The Guardian*⁷ – que contaram com Nick Davies na cobertura do caso – descrevem a dimensão do vazamento que funcionou como um divisor de águas:

O tamanho absoluto era avassalador. Se o minúsculo *stick* de memória (...) fosse um conjunto de textos impressos, teria constituído uma biblioteca com mais de 2.000 livros de

tamanho considerável. Nenhum humano teria tentado escrever tanto antes do advento da era digital: nenhum espião teria sido capaz de despejar cópias de tantos papéis sem usar um caminhão, e nenhum humano seria capaz posteriormente de analisá-los sem gastar meia vida na tarefa (LEIGH; HARDING, 2011, p. 140).

A atuação do *hacker* Julian Assange no vazamento do *Wikileaks War Logs* revela outro fator preponderante na expansão do jornalismo de dados: a influência do movimento do *software livre* e da cultura *hacker* na constituição de *ethos* que se associa aos jornalistas de dados. Uma experiência precursora foi a do grupo norte-americano *Hacks/Hackers*⁸, fundado em 2009 por jornalistas que se autodenominavam uma comunidade digital interessada em reunir *hackers* capazes de filtrar e visualizar informações. O movimento angariou apoio de instituições como a *Columbia University* e o *National Institute for Computer Assisted Reporting* (NICA). De acordo com Lewis e Usher (2014), o crescente interesse em criar narrativas a partir de dados gerou uma demanda por expertise tecnológica e iniciativas, como o *Hack Days*, organizado pelos jornais *New York Times*, *Guardian* e *The Globe*, e resultou em conceitos como *hacking social* – acesso a informações restritas ou a um espaço físico sem a devida permissão.

Com o objetivo de estimular o trabalho colaborativo, os encontros do grupo promoviam interlocução sobre rotinas como mineração de dados entre jornalistas, desenvolvedores de *software*, designers e outros. Em abril de 2011, o *Hacks/Hackers* de Buenos Aires, na Argentina, organizou uma maratona de 30 horas em Rosário com a intenção de desenvolver um mapa interativo com dados relevantes sobre a ditadura militar no país (GRAY, 2012).

O *The Bureau of Investigative Journalism* (TBIJ) também organizou este tipo de evento ao redor do Reino Unido: o *The Big Council Budget Hack*⁹, que aconteceu no dia 3 de fevereiro de 2018, mobilizou profissionais multimídia para elaborar estratégias de coleta de dados sobre o orçamento dos conselhos estaduais britânicos.

Vale ressaltar que, quando inventou a *World Wide Web* (*WWW*) em 1992, o

físico Tim Bernes-Lee se ancorou no programa *Enquire*, encorajado pela parte da comunidade da internet que partilhava aspirações da contracultura. Depois, a versão do sistema de hipertexto WWW, que permitia obter e acrescentar informação para qualquer computador conectado à internet, sofreu modificações de *hackers*¹⁰ (WU, 2012). O surgimento, no final dos anos 1990, de programas de compartilhamento de arquivos *peer-to-peer* (P2P), como *Napster*, *Gnutella* e *KaZaa*, aprimorou os modelos de distribuição de conteúdo em rede. Em 2010, o próprio Berners-Lee chegou a afirmar que o futuro estava no jornalismo de dados (GRAY *et al.*, 2012; HOWARD, 2014).

A filosofia do código aberto e da partilha está na base da cultura *hacker* e do movimento do *software livre*. Essa mentalidade está, a nosso ver, intrinsecamente embutida no ideário da partilha de dados com o leitor. A cultura *hacker* nasce da autonomia propiciada pela interconexão de computadores e está associada a um conjunto de valores e crenças que despontaram no trabalho de colaboração criativa entre programadores de computador. A pesquisa etnográfica realizada por Träsel (2014, p. 182) com jornalistas de dados do *Estadão Dados* revelou profissionais que acreditam no “cooperativismo e na tecnofilia típicos da cibercultura”, mas identificados com o *ethos* jornalístico convencional calcado na “valorização da objetividade”.

Um momento emblemático de partilha profissional foi o *Panama Papers*, reportagem colaborativa conduzida pelo ICIJ (*International Consortium of Investigative Journalism*) que contou com a participação de mais de uma centena de profissionais de 107 veículos de comunicação e trouxe à tona operações ilícitas em paraísos fiscais que envolviam políticas e figuras públicas em países como Brasil, China, Reino Unido, França e México. A cobertura a partir do vazamento de 11,5 milhões de registros financeiros levou à renúncia dos primeiros-ministros da Islândia e Afeganistão. Nas primeiras semanas, 72 reportagens produzidas pelo ICIJ foram acessadas por mais de 80 milhões de pessoas espalhadas em 200 países. Em 2021, outra empreitada envolveu 600 profissionais no vazamento batizado de *Pandora Papers*, também conduzido

pelo *ICIJ*. Os documentos secretos atuais e ex-líderes mundiais e pelo menos 330 políticos mantêm operações ilegais em empresas de fundo secretas. O trabalho colaborativo em ambas as coberturas aponta uma das características que perpassam a rotina do jornalismo de dados e evidencia seu potencial em promover transparência.

5 CONSIDERAÇÕES FINAIS

Se a Contemporaneidade produziu forte impacto em conceitos como objetividade, Deuze (2003) atrela esse ambiente de contestações introduzido pela “modernidade líquida” à fluidez na classificação do que é jornalismo de dados. Coddington (2015, p. 337) tentou esgotar a imprecisão a partir de uma proposta que dissocia a RAC de outras denominações contemporâneas. Enquanto considera a reportagem com auxílio do computador intrinsecamente ligada às Ciências Sociais, Coddington (2015; 2018) funda o Jornalismo Computacional em processos de abstração e automação da informação e caracteriza o jornalismo de dados a partir da prática da análise de dados e apresentação dessa análise. No Jornalismo Guiado por Dados, por sua vez, a pauta surge da entrevista aos dados (PARASIE, 2015; TRÄSEL, 2014). Mas é possível que essas derivações de nomenclaturas confundam os próprios jornalistas.

A reportagem *As cidades que mais matam mulheres no Brasil*¹¹, da *Agência Pública*, feita por Andrea Dip e Bruno Fonseca, por exemplo, não está na seção *Dados*. Embora tenha partido de uma consolidação de dados e apresente vários gráficos e nota metodológica, foi classificada e inserida na seção *Reportagem*. Entretanto, Andrea Dip assina outra matéria sobre o assunto, que foi apurada *in loco* em Ananindeua, no Pará – a cidade mais violenta para mulheres – e publicada no típico formato da reportagem. Muitas matérias publicadas na seção *Dados* da *Pública* são complementadas com entrevistas realizadas com especialistas e personagens, um critério que poderia sugerir a diferença, mas na prática mostra uniformidade de método em boa parte do material publicado nas duas seções.

Enquanto alguns teóricos, como vimos, acreditam que o pensamento computacional significa pensar como um cientista da computação, nossa compreensão é a de que o jornalismo de dados requer uma prática profissional voltada para diferentes processos baseados em um amálgama entre jornalismo e tecnologia. Mas, em seu cerne, mostra-se fortemente entrelaçado à rotina da investigação que busca trazer à tona relatos precisos. Essa prática profissional está, a nosso ver, intrinsecamente embutida no ideário da partilha de dados com o leitor presente na cultura *hacker* que prima pela proposta da transparência.

No percurso genealógico que realizamos a partir de Anderson (2018), destacamos nossa percepção de que o rigor e a transparência perpassam a evolução das técnicas de apuração relacionadas ao jornalismo de dados. A datificação da sociedade e a publicização de informações digitalizadas em bases de dados servem de combustível que estimula as diversas nomenclaturas existentes neste campo profissional.

REFERÊNCIAS

AGUIAR, Leonel; RODRIGUES, Cláudia. Precisão e credibilidade: agências independentes de jornalismo e o uso do *big data*. **Esferas**, Brasília, n. 14, p. 109-120, jan./abr. 2019.

AGUIAR, Leonel; ANDRADE, Ana P. G. de. Novas interfaces nas rotinas produtivas e credibilidade jornalística: uma contribuição aos estudos da profissão. **Mediação**, Belo Horizonte, v. 22, n. 31, p. 67-79, jul./dez. 2020.

AGUIAR, Leonel. Entretenimento: valor-notícia fundamental. **Estudos em Jornalismo e Mídia**, Florianópolis, v. 5, n. 1, p. 13-23, jan./jun. 2008.

AGUIAR, Leonel. O jornalismo investigativo e seus critérios de noticiabilidade: notas introdutórias. **Alceu - Revista de Comunicação, Cultura e Política**, Rio de Janeiro, v. 7, n. 13, p. 73-84, jul./dez. 2006.

ANDERSON, Chris W. **Apostles of certainty**: Data Journalism and politics of doubt. Oxford: Oxford University Press, 2018.

CODDINGTON, Mark. Building frames link by link: the linking practices of blogs and news sites. **International Journal of Communication**, Rio de Janeiro, v. 6, p. 2007-2026, 2012.

CODDINGTON, Mark. Clarifying journalism's quantitative turn: a typology for

evaluating Data Journalism, Computational Journalism, and Computer-assisted Reporting. **Digital Journalism**, Londres, v. 3, n. 3, p. 331-348, 2015.

CODDINGTON, Mark. Defining and mapping Data Journalism and Computational Journalism. In: ELDRIDGE, Scott; FRANKLIN, Bob (eds.). **The Routledge Handbook of developments in digital journalism studies**. New York: Routledge, 2018. p. 225-236.

COHEN, Sarah; HAMILTON, James T.; TURNER, Fred. Computational Journalism: how computer scientists can empower journalists, democracy's watchdogs, in the production of news in the public interest. **Communications of the ACM**, Nova York, v. 54, n. 10, p. 66-71, 2011.

DEUZE, Mark. The web and its journalisms: considering the consequences of different types of news media online. **New Media & Society**, [S.l.], v. 5, n. 2, p. 203-230, 2003.

ETTEMA, James; GLASSER, Theodore. On the epistemology of Investigative Journalism. **Communication**, [S.l.], v. 8, p. 183-206, 1985.

FINK, Katherine; ANDERSON, C. W. Data Journalism in the United States: beyond the "usual suspects". **Journalism Studies**, Londres, v. 16, n. 4, p. 467-481, 2015.

FLEW, Terry; SPURGEON, Christina; DANIEL, Anna; SWIFT, Adam. The promise of computational journalism. **Journalism Practice**, [S.l.], v. 6, n. 2, p. 157-171, 2012.

GRAY, J; BOUNEGRU, L.; CHAMBERS, L. **The Data Journalism handbook: how journalists can use data to improve the news**. Sebastopol: O'Reilly, 2012.

HERMIDA, Alfred; YOUNG, Mary Linn. **Data Journalism and the regeneration of the news**. New York: Routledge, 2019.

HOWARD, Alexander B. **The art and science of Data-driven Journalism**. New York: Tow Center for Digital Journalism, 2014.

LEIGH, David; HARDING, Luke. **Wikileaks: inside Julian Assange's war on secrecy**. London: Guardian Books, 2011.

MEYER, Philip. **Precision journalism: a reporter's introduction to Social Science methods**. Maryland: Rowman & Littlefield, 1973/2002.

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

MIELNICZUK, Luciana; GERKHE, Marília. Philip Meyer, the outsider who created Precision Journalism. **Intexto**, Porto Alegre, n. 39, p. 4-13, maio/ago. 2017.

PARASIE, Sylvain; DAGIRAL, Eric. Data-driven Journalism and the public good: computer-assisted reporters and programmer-journalists in Chicago. **New Media and Society**, Chicago, v. 15, n. 6, p. 853-871, 2013.

PARASIE, Sylvain. Data-driven revelation? Epistemological tensions in investigative journalism in the age of big data. **Digital Journalism**, [S.l.], v. 3, n. 3, p. 364-380 2015.

ROGERS, Simon. **Facts are sacred**. London: Faber and Faber, 2013.

SCHUDSON, Michael. **Descobrimos a notícia**: uma história social dos jornais nos Estados Unidos. Petrópolis: Vozes, 2010.

SCHUDSON, Michael. **The rise of the right to know**: politics and the culture of transparency, 1945–1975. Cambridge: Harvard University Press, 2015.

SEQUEIRA, Cleofe Monteiro de. **Jornalismo Investigativo**: o fato por trás da notícia. São Paulo: Summus Editorial, 2005

RÄSEL, Marcelo. **Entrevistando planilhas**: estudo das crenças e do ethos de um grupo de profissionais de jornalismo guiado por dados no Brasil. 2014. 314 f. Tese (Doutorado em Comunicação Social) — Faculdade de Comunicação Social, Pontifícia Universidade Católica do Rio Grande do Sul, 2014.

WU, Tim. **Império da comunicação** – do telefone à internet, da AT&T ao Google. Rio de Janeiro: Zahar, 2012.

NOTAS:

¹ Disponível em <https://www.ire.org/about-ire/>.

² A página atual é nomeada *Data* e traz reportagens com dados abertos. Disponível em: <https://www.theguardian.com/data>.

³ O *blog Na Base de Dados*, de *O Globo*, que oferecia acesso a dados de algumas reportagens de dados, foi descontinuado. Disponível em <https://blogs.oglobo.globo.com/na-base-dos-dados/>.

⁴ O *Prêmio Claudio Weber Abramo de Jornalismo de Dados* foi criado, em 2019, com o objetivo de “incentivar a excelência no uso de dados e fortalecer este campo de trabalho e a democracia brasileira”. Disponível em <https://premio.jornalismodedados.org/edicoes/2019/>.

⁵ O *Freedom of Information Act (FOIA)* foi aprovado, nos Estados Unidos, em 1966, e determina o fornecimento de dados públicos armazenados por órgãos públicos (SCHUDSON, 2015). No Reino Unido, o *FOIA* foi editado pela primeira vez em 2000. No Brasil, a LAI (Lei de Acesso à Informação) entrou em vigor em 18 de novembro de 2011 e assegura, portanto, transparência pública e prestação de contas (*accountability*). Qualquer cidadão, com respaldo nestas legislações, pode enviar pedidos de informação de dados que não estejam abertos ao público.

⁶ A agência *Bureau of Investigative Journalism* – que posteriormente mudaria o nome para *The Bureau of Investigative Journalism* – publicou extensa cobertura sobre o vazamento em colaboração com o *Wikileaks*. No Brasil, a *Agência Pública* também firmou parceria com o *Wikileaks* para publicação de uma série de reportagens – projeto que está na base da proposta de criação da organização, em 2011. Um ano antes, Natália Viana, diretora da *Pública*, fez parte da equipe que trabalhou com Assange, na Inglaterra, na sistematização dos dados coletados pelo vazamento.

⁷ Leigh foi um dos coordenadores da cobertura ao lado de Nick Davies, do jornal *The Guardian*. Enquanto *The Guardian* e *Der Spiegel* decidiram publicar o *link* para o *Wikileaks* no jornal, oferecendo acesso aos dados, o *New York Times* optou por manter a fonte no anonimato. Toda a informação permanece acessível no site do *Wikileaks* (LEIGH; HARDING, 2011). Atualmente, o *link* no *The Guardian* direciona o leitor para a página principal da série de reportagens.

⁸ O *Hacks/Hackers* tornou-se a maior organização do gênero. Em janeiro de 2014, possuía mais de 75 grupos baseados em cidades ao redor do mundo e mais de 23 mil membros conectados em plataformas digitais que se reuniam em *hack days* e compartilhavam conhecimento (LEWIS; USHER, 2014).

⁹ O termo *hacking* possui a conotação de acesso a dados de forma não autorizada, mas os *hackathons* buscavam desenvolver formas legais de coleta de dados com chamadas públicas (*TBI*).

¹⁰ Richard Stalman, fundador da FSF (*Free Software Foundation*) salienta que o termo *hacker*, originalmente, se referia àqueles que alteravam programas e sistemas com o único intuito de aperfeiçoá-los (MARTINS, 2014; GILLMOR, 2006). A conotação depreciativa decorreu de acordos sigilosos impostos aos usuários em função de termos de venda e comercialização de novas gerações de software; acordos que significavam, na sua essência, a exclusão da “comunidade cooperativa” por conta da imposição de softwares com fins lucrativos (JONES; SALTER, 2003).

¹¹ Disponível em <https://apublica.org/2017/10/as-cidades-que-mais-matam-mulheres-no-brasil/>.

SOBRE OS AUTORES:

Cláudia Miranda Rodrigues

Doutora e Mestre em Comunicação pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Jornalista diplomada pela Universidade Federal do Rio de Janeiro (UFRJ). Integrante do Gupo de Pesquisa Teorias do Jornalismo e Experiências Profissionais (CNPq/PUC-Rio).

Orcid: <https://orcid.org/0000-0003-1055-822X>

E-mail: claudiar63@gmail.com

Leonel Azevedo de Aguiar

Professor do Programa de Pós-graduação em Comunicação da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Coordenador do curso de Jornalismo do Departamento de Comunicação da PUC-Rio. Doutor e Mestre em Comunicação pela Universidade Federal do Rio de Janeiro (UFRJ). Jornalista diplomado pela Universidade Federal Fluminense (UFF). Líder do GP Teorias do Jornalismo e Experiências Profissionais (CNPq/PUC-Rio).

Orcid: <https://orcid.org/0000-0002-1220-2131>

E-mail: laaguiar@uol.com.br

Artigo recebido em: 09 nov. 2023. | Artigo aprovado em: 23 maio 2024.