
***The Pudding* como ponto de partida para aprender Jornalismo de Dados¹**

Julia BARROSO²

Daniela OLIVEIRA³

Universidade Veiga de Almeida, Rio de Janeiro, RJ

RESUMO

O artigo observou o veículo jornalístico *The Pudding*, criado nos Estados Unidos em 2017 e focado em criar narrativas baseadas em dados, e a possibilidade de aprender Jornalismo de Dados por meio da transparência metodológica. Com esse objetivo, foi realizada uma pesquisa bibliográfica sobre o Jornalismo de Dados, desde o contexto histórico até a dimensão ética e a relação do jornalismo com a cibercultura. Também foi realizado um estudo de caso sobre o *Pudding*, a fim de observar os conceitos teóricos aplicados na prática e como as matérias e conteúdos disponibilizados podem servir de inspiração para profissionais interessados na área.

PALAVRAS-CHAVE: Jornalismo de Dados; Ensino de Jornalismo; *The Pudding*.

INTRODUÇÃO

O jornalismo de dados, vindo do jornalismo de precisão e da Reportagem com Auxílio do Computador (RAC), não é uma prática nova. Construir matérias a partir de dados, porém, traz novos desafios diante da quantidade de informação disponível atualmente. Com isso, surgem ferramentas para estruturação e análise de dados, acompanhados pela velocidade de processamento dos computadores. O jornalista, então, precisa lidar com técnicas e tecnologias que não eram necessárias antes.

O impacto dessas mudanças afeta a vida cotidiana. “A convergência envolve uma transformação tanto na forma de produzir quanto na forma de consumir os meios de comunicação” (JENKINS, 2009, p.44). A cibercultura empresta ao jornalismo conceitos como transparência e colaboração, por meio de inteligência coletiva, transformando a relação entre veículos e profissionais, antes mais competitiva.

¹ Trabalho apresentado no IJ01 – Jornalismo, da Intercom Júnior – XVII Jornada de Iniciação Científica em Comunicação, evento componente do 44º Congresso Brasileiro de Ciências da Comunicação

² Graduada em 2021.1 no curso de Jornalismo da Universidade Veiga de Almeida, e-mail: barroso.julia@gmail.com

³ Orientadora do trabalho. Professora e pesquisadora do Curso de Jornalismo da Universidade Veiga de Almeida (Rio de Janeiro), integrante do NDE da mesma instituição. Mestre em Comunicação Social pela PUC-Rio, e-mail: daniela.oliveira@gmail.com

Além das práticas tradicionais do jornalismo, o profissional interessado na área de dados precisa aprender a utilizar ferramentas de análise e de visualização dos dados, bem como entender conceitos básicos de matemática. A consequência primária é a produção de um jornalismo mais confiável e, muitas vezes, transparente, mas esses conhecimentos também trazem oportunidades diferentes porque, mais do que uma nova área de trabalho dentro do jornalismo, estes jornalistas de dados já são procurados fora dos meios de comunicação (GRAY; BOUNEGRU; CHAMBERS, 2012).

Antes de o jornalismo de dados ser difundido pelas redações, é necessário formar profissionais capacitados. Diante da impossibilidade de explorar totalmente as técnicas durante a graduação, muitos alunos não conseguem se aproximar do uso de planilhas e dos conceitos estatísticos que possibilitam o trabalho com dados. Para Ventura e Paulino (2020), é cada vez mais necessário que os jornalistas percam a famosa pouca afinidade com os números.

Essa evolução tecnológica se reflete no processo de produção jornalístico, ao passo que a cada adesão de uma inovação é preciso reformular e adaptar técnicas e práticas. Logo, os cursos de graduação voltados para a área também precisam atualizar a oferta de alguns conteúdos para se manterem contemporâneos e permitir que o aluno possua um conhecimento que o prepare para a atuação profissional de forma satisfatória. A oferta de disciplinas destinadas a utilizar ferramentas digitais é uma das consequências disso. E entre elas, encontra-se as que se destinam a explorar o Jornalismo de Dados (JD) (VENTURA; PAULINO, 2020, p.247).

De acordo com Mielniczuk e Träsel (2017, p.624), “centrar o ensino nos processos cognitivos da apuração, redação e edição de notícias a partir de bases de dados pode ser mais produtivo e garantir maior flexibilidade a futuros repórteres”. Assim, observar o trabalho por trás de reportagens de dados pode se mostrar uma ferramenta interessante para captar e aprender sobre esses processos.

*The Pudding*⁴ é um veículo americano voltado para visualizações criativas e narrativas focadas em dados que se mostra útil nesse aprendizado. Devido à prática de transparência e à divulgação de metodologias – trazendo inclusive as bases de dados utilizadas nas matérias, tutoriais específicos para etapas do trabalho com dados e até mesmo vídeos no *YouTube* que explicam projetos –, este artigo tem como objetivo analisar o *Pudding* como veículo de jornalismo de dados e ferramenta de aprendizado. A pesquisa busca mostrar como a abordagem inovadora e as visualizações elaboradas do

⁴ Disponível em <https://pudding.cool>. Acesso em maio 2021.

Pudding são inspiração para profissionais interessados na área, e como a transparência colabora com a multiplicação de conhecimento, além de contribuir para a confiabilidade.

PANORAMA HISTÓRICO E O ENSINO DO JORNALISMO DE DADOS

O jornalismo de dados foi precedido pelo jornalismo de precisão e pela Reportagem com Auxílio do Computador (RAC). De acordo com Gray, Bounegru e Chambers (2012, p.30) no Manual do Jornalismo de Dados, a RAC foi usada pela primeira vez em 1952, enquanto o termo jornalismo de precisão foi cunhado por Philip Meyer na década de 1970 para se referir à apuração jornalística que empregava métodos de pesquisa das ciências sociais, como forma de alcançar a objetividade e a verdade buscadas pelos profissionais da área.

Em sua tese de doutorado, Marcelo Träsel (2014, p.27) ressalta que o uso de gráficos é observado desde a industrialização de redações no século XIX, devido às inovações em técnicas de impressão, mas que a manutenção de bancos de dados é uma prática mais antiga – os arquivos internos dos jornais serviam para recuperar informação, verificar antecedentes e como biografia para personagens das notícias. A utilização do computador, porém, adicionou novas variáveis ao processo: maior precisão, capacidade de armazenamento e acesso remoto.

Para Karam (2014, p.46), o surgimento e a implementação de novas tecnologias no processo jornalístico não alteram conceitos como interesse público, fato e acontecimento. A importância do jornalismo de dados, porém, é sua capacidade de reunir informações que, de outra forma, não poderiam ser transformadas em visualizações e talvez não tivessem sentido, além de democratizar “recursos, ferramentas, técnicas e métodos antes restritos aos especialistas” (GRAY; BOUNEGRU; CHAMBERS, 2012, p.34).

Apesar da longa tradição do uso de dados no jornalismo, as ferramentas e os métodos utilizados não são necessariamente intuitivos ou fáceis, e o profissional pode ter dificuldade na prática sem o treinamento apropriado. Assim, com o papel positivo do jornalismo de dados, se torna interessante que jornalistas aprendam como trabalhar com as novas tecnologias. Como destacado em Gray, Bounegru e Chambers (2012, p.40), o jornalismo de dados no Brasil ainda não está presente na maioria das redações, apesar de ser divulgado desde 2010 por associações profissionais internacionais, imprensa e

jornalistas. Com o processo de informatização, porém, tem sido adicionado à prática na cultura jornalística nacional de acordo com os autores.

Ventura e Paulino (2020, p.255) concluem que o Brasil é o país, entre os 12 analisados, com mais Instituições de Ensino Superior (IES) listadas na base de dados *Quacquarelli Symonds*, sendo que cinco das 11 instituições listadas oferecem alguma disciplina que atenda ao *corpus*, mas a área ainda é pouco lecionada no país. Elas ressaltam que “o levantamento feito serve justamente para demonstrar a ausência de um foco maior nessa prática pedagógica, que acontece ainda de forma diluída e esparsa.” (VENTURA; PAULINO, 2020, p.298)

Conforme Träsel (2014), o autodidatismo caracteriza o processo de aprendizado de técnicas de JGD entre os profissionais atuantes no mercado. Os jornalistas procuram recursos on-line como tutoriais em vídeo, fóruns de discussão e documentação de software para resolver problemas específicos de cada pauta. Publicações especializadas, portanto, são um importante vetor de difusão desse tipo de prática. (MIELNICZUK; TRÄSEL, 2017, p.622)

Berret e Phillips (2016, p.47) afirmam que alguns professores não tratam o jornalismo de dados como vital para a carreira, e isso se relaciona com o fato de que, geralmente, essa disciplina ainda é, muitas vezes, oferecida de forma eletiva, atraindo apenas alguns alunos – além das questões de capacidade e falta de visibilidade do tema dentro das universidades. Como consequência, muitos dependem de instruções fora da educação formal.

ÉTICA E CIBERCULTURA

Apesar das diferenças na prática, Araújo (2016, p.189) ressalta que o jornalista que trabalha com dados é profissional de imprensa e que é necessária a reflexão, por parte de quem faz notícia, sobre a forma como a mídia opera. Devido às várias etapas no trabalho com dados, cada uma dessas fases tem necessidades diferentes. Não só o jornalista deve estar consciente das possíveis falhas contidas nos dados, mas o público também precisa dessa informação para entender a abordagem feita.

Como destacado por Marília Gehrke (2020, p.3) no simpósio *Computation + Journalism*, o código de ética brasileiro foi escrito em 2007 e não especifica transparência como tópico de interesse, apenas pontua que o acesso à informação é fundamental. Ainda que não explicitado na deontologia jornalística, esse método de prestação de contas é válido para garantir interpretações corretas.

A transparência, de acordo com Gehrke (2020, p.1) é um importante valor jornalístico para combater desinformação e aumentar a credibilidade dos veículos de notícias. Ela menciona que o aumento de veículos praticando jornalismo de dados no Brasil acontece na época de queda da confiança nas mídias, e essa queda foi quantificada pelo *Digital News Report 2019*: menos 11 pontos percentuais no Brasil em comparação ao ano anterior.

A transparência metodológica se relaciona com os preceitos do Jornalismo de Precisão e “enseja a reprodutibilidade dos experimentos e, desta forma, sua validade pelos pares” (TRÄSEL, 2014, p.128). Outro fator importante a ser considerado é que, para Träsel (2014, p.235), a relação cooperativista de disponibilizar bases de dados geram fontes de valor-notícia por tempo indefinido, enquanto as matérias produzidas a partir delas podem ser perecíveis. Além disso, a pouca variedade de fontes foi destacada por Bounegru (2014, *apud* TRÄSEL, 2014, p.125), que menciona fontes como o Banco Mundial e as Nações Unidas, amplamente utilizadas devido à facilidade de acesso. Com isso, reforça-se a ideia de que as preocupações éticas do trabalho com dados são as mesmas da prática jornalística vista como tradicional, apenas acrescentando novos detalhes às questões anteriores.

A ausência de menções explícitas ao jornalismo de dados nos códigos de ética não isenta as considerações dessa prática no *ethos* dos profissionais, que Bonnewitz (2003, *apud* TRÄSEL, 2014, p.16) define como sendo “os princípios ou os valores em estado prático, a forma interiorizada e não-consciente da moral que regula a conduta cotidiana”. Träsel (2014, p.17) explica que o *ethos*, então, é um elemento válido para observar práticas profissionais relativamente novas, que ainda estão em processo de institucionalização e cuja deontologia está em elaboração.

Na Sociedade da Informação, segundo Karam (2014, p.137), e diante da convergência tecnológica, o *ethos* jornalístico tem maior responsabilidade para ter legitimidade social e ser portador de uma credibilidade que o qualifique para representação social. Ou seja, as mudanças na prática profissional e no consumo de produto jornalístico exigem que o trabalho seja feito de forma ética, dependendo mais do bom senso do que de manuais com regras estabelecidas.

De acordo com Rüdiger (2011, *apud* TRÄSEL, 2014, p.36), a cibercultura pode ser definida como uma formação histórica que surge a partir da tecnologia dos computadores e da rede de comunicação digital, junto com novas práticas sociais e de

imaginários, e que foi consolidada durante o século XX. Por sua vez, o *ethos* do jornalismo de dados “apresenta uma combinação de elementos da cibercultura, em especial da ética hacker, e da cultura profissional jornalística” (TRÄSEL, 2014, p.15).

Isso se justifica com a ideia de que a cibercultura está presente na rotina do homem comum, categoria em que os jornalistas se encaixam, e, por isso, passam a incorporar o pensamento tecnológico no desempenho de tarefas cotidianas, como a prática profissional (RÜDIGER, 2011, *apud* TRÄSEL, 2014, p.16). A tecnologia computacional está presente nas relações sociais, na autoidentidade e no sentido amplo de vida social, de acordo com Santaella (2010, p.105). A construção do produto jornalístico dentro do chamado ciberespaço possibilita uma composição pensada como enredo interativo que tenha começo, fluxo e fim, mesmo que incorpore alternativas de desenvolvimento (MACHADO, 2007, p.119). Partindo da ideia tradicional de jornalismo, é possível criar conteúdo com ramificações a serem exploradas de acordo com o interesse do público.

Agora, o jornalista depende do público não apenas como consumidor, mas cocriador do produto. Negroponte (1995, *apud* SANTAELLA, 2010, p.89) explica que, por isso, as redes de televisão e de computadores são quase opostas: aquela é baseada em uma hierarquia distributiva, conta com uma fonte e muitos receptores, enquanto esta é formada por processadores heterogêneos, todos atuando como fontes e escoadouros. E Santaella (2010, p.104) acrescenta que a cibercultura é heterogênea, descentralizada, reticulada, e não se dinamiza apenas quando os usuários ligam o computador.

O movimento da cibercultura se relaciona com a convergência. Mais do que uma questão tecnológica, a convergência “representa uma transformação cultural, à medida que consumidores são incentivados a procurar novas informações e fazer conexões em meio a conteúdos de mídia dispersos.” (JENKINS, 2009, p.30). A inteligência coletiva e as comunidades virtuais, mencionadas por Santaella (2010, p.105) como duas importantes consequências da cibercultura, também fazem parte do processo de convergência. Fundamental para a formação autodidata dos jornalistas de dados, a “inteligência coletiva refere-se a essa capacidade das comunidades virtuais de alavancar a expertise combinada de seus membros.” (JENKINS, 2009, p.56).

Além das trocas em grupos de profissionais, ofertas de iniciativas independentes buscam suprir a carência no ensino do jornalismo de dados dentro dos cursos de graduação, como os *Massive Open Online Course* (MOOC) oferecidos pelo Centro *Knight* para o Jornalismo nas Américas, cursos de capacitação da *Google News Initiative*

e eventos da Associação Brasileira de Jornalismo Investigativo (Abraji) (VENTURA; PAULINO, 2020, p.253), por exemplo.

De acordo com Jenkins (2009, p.331), a convergência também provoca flutuações na mídia e expande oportunidades para grupos alternativos reagirem aos meios de comunicação de massa. Ele acrescenta que a reação à dominação dos meios se faz necessária porque a concentração freia a competição e prioriza indústrias em detrimento às demandas dos consumidores, além de reduzir a diversidade, que é “importante em termos de cultura popular, essencial em termos de notícias” (JENKINS, 2009, p.331).

A diversidade na prática jornalística é relevante, proporcionando novos pontos de vista, mas os avanços tecnológicos não são necessariamente aproveitados por todos. A dificuldade para acesso ou entendimento das técnicas de jornalismo de dados pode ser um bloqueio para muitos profissionais. Porém, como destacado por Araújo (2016, p.185), muitas publicações buscam ensinar o jornalista a utilizar os recursos, principalmente os gratuitos, atualmente disponíveis. Essa independência no aprendizado tem relação com a multiplicidade de veículos e profissionais.

Com os métodos disponíveis, há uma abertura de mercado para veículos e profissionais que não estão subjugados aos grandes conglomerados de mídia. Também são abordadas novas pautas, que fogem do ciclo de notícias e dos possíveis agendamentos. Dessa forma, novos espaços podem ser ocupados, antes ignorados por serem custosos, inacessíveis ou irrelevantes para os meios tradicionais.

THE PUDDING

Enquanto alguns veículos utilizam jornalismo de dados para algumas reportagens, outros se dedicam totalmente a essa prática – é o caso do *Pudding*. Criado em janeiro de 2017, o veículo americano apresenta pautas inéditas e criativas, que abordam temas cotidianos. O *Pudding* busca abordar tópicos debatidos na cultura popular com as chamadas *visual essays* (ensaios visuais, em tradução livre), utilizando bases de dados originais e interatividade (THE PUDDING, 2017).

Para Matt Daniels, um dos fundadores do *Pudding*, o objetivo é um dia revelar verdades profundas sem a meia-vida de 24 horas – às vezes, ao olhar os trabalhos visuais do *New York Times* ou do *Washington Post*, ele pensa sobre a genialidade que foi transformada naquele produto e ainda assim o resultado é o lixo da *internet* (DANIELS, 2017). Na visão dele, os conteúdos de dados relacionados a pautas quentes, ainda que

exijam tempo e habilidade dos jornalistas, acabam sendo abandonados depois de um dia. O *Pudding* busca, então, publicar matérias de maior durabilidade, que sejam atemporais.

Estamos também tentando avançar a técnica. Jornalismo visual está ainda na infância. Não temos uma linguagem padrão estabelecida como encontramos em reportagens tradicionais. Às vezes arriscamos uma abordagem visual não-familiar – não porque é trabalho garantido, mas porque não saberemos até tentar. Raramente organizações têm a liberdade de se arriscar, mas nós ainda somos pequenos o suficiente para experimentar em busca de qualidade. (THE PUDDING, 2017, tradução nossa)⁵

A vontade de construir um veículo inovador se relaciona com o conceito do jornalista-engenheiro, abordado em um artigo publicado no *Medium* de Daniels anos antes do lançamento do *Pudding*. Ao notar que matérias bem-conceituadas do *New York Times* e do *Bloomberg* eram predominantemente *software* em vez de texto, ele também começou a experimentar essa nova abordagem ao contar histórias: em vez de apenas apresentar uma análise, publicar uma visualização que facilitasse aos leitores a exploração dos dados (DANIELS, 2015).

Ao verificar que há interesse nesse tipo de conteúdo, Daniels (2015) explicita um novo caminho para o jornalista, que traz novas possibilidades. Para os leitores, a construção da narrativa a partir da exploração de dados aumenta a autonomia; para os profissionais, apesar dos desafios de lidar com ferramentas mais avançadas ou com linguagens de programação, maior foco na personalização de visualizações. Ser um jornalista-engenheiro, então, significa possuir conhecimentos computacionais que acrescentem à prática jornalística, de forma que não seja necessário contar com profissionais de outras áreas, como *designer*, para chegar ao produto final.

De acordo com o guia de boas práticas escrito pelo jornalista de dados Kuang K. K. Ser (2018, p.4), a diversidade de habilidade importa mais do que o tamanho da equipe, e é importante não depender de uma pessoa no time para fazer as etapas do trabalho. Isso é refletido no trabalho dos jornalistas-engenheiros. Daniels (2017) afirma que o *Pudding* é um coletivo, cada membro precisa de uma identidade que seria cultivada independentemente, e acrescenta que eles são mais do que partes em uma máquina de ensaios visuais. Os sete jornalistas-engenheiros têm históricos bem diferentes de carreira,

⁵ “We’re also trying to advance the craft. Visual journalism is still in its infancy. We don’t have an established pattern language found in traditional reporting. Sometimes we’ll attempt an unfamiliar visual approach—not because it’s guaranteed to work, but because we won’t know until we try. Rarely do organizations have the liberty to take such risks, yet we’re small enough to experiment in the pursuit of quality.” (THE PUDDING, 2017)

desde biologia marinha à filosofia, incluindo ciência da computação e jornalismo, explicitando que é possível trabalhar na área mesmo sem anos de experiência.

O *Pudding* surgiu a partir do *Polygraph*, fundado dois anos antes. A principal diferença é que o *Polygraph* desenvolve visualizações para clientes produzindo peças comissionadas. Ainda que o *Polygraph* não tenha o direcionamento jornalístico do *Pudding*, suas visualizações seguem os mesmos princípios de uma matéria de dados. O foco em narrativas e abordagens criativas permanece, ainda que os resultados dependam do envolvimento do cliente. A divisão entre os dois *sites* é importante por evidenciar a diferença entre os conteúdos. Com mesma equipe de jornalistas-engenheiros, o *Pudding* serve como um portfólio para potenciais clientes do *Polygraph*, enquanto o *Polygraph* é essencial para a sobrevivência do *Pudding*.

Essa noção de unidade no local de trabalho é um diferencial do *Pudding* em relação a outros veículos. É frequente, na visão de Daniels (2017), que autores independentes renunciem a suas vozes individuais quando começam a trabalhar em uma agência, redação ou corporação. Por isso, em sua projeção do que esperava para o *Pudding* em 2018, ele expressa suas considerações sobre manter a equipe satisfeita. Contratar pessoas inteligentes que são apaixonadas pela missão se traduz em esforço e rápido aprendizado, então o *Pudding* precisa ser um ótimo lugar para trabalhar, uma melhoria em comparação a ser profissional independente ou fazer parte de uma outra redação (DANIELS, 2017).

Observar as etapas do trabalho com dados aplicadas a um veículo é importante para mostrar o que muda dentro da redação e como funciona a equipe focada nesse tipo de jornalismo. O *Pudding*, especificamente, ainda se diferencia devido à abordagem que prioriza a narrativa através das visualizações. O processo editorial é disponibilizado no *GitHub*, uma plataforma muito utilizada por programadores.

O *Pudding* divide seu processo editorial em quatro fases (GOLDENBERG, 2017): ideias – uma fase constante, que semanalmente conta com o *storytime*, um tipo de reunião de pauta; seleção – revisão do *backlog*, um registro de ideias, e escolha de uma história; exploração – a partir de resultados de pesquisa e da criação de um gancho, decidir de a ideia escolhida realmente será expandida para um artigo; apresentação – produção do artigo e, então, publicação.

Cada reunião do *storytime* é responsabilidade de um jornalista-engenheiro, que deve trazer novas ideias, mesmo que sejam apenas premissas. A atribuição de cada

storytime a um dos sete membros da equipe só é possível devido às habilidades generalistas dos jornalistas-engenheiros. O *backlog* também é fundamental para a estrutura do *Pudding*. Além dos registros individuais e compartilhados dentro da equipe, existe um *backlog* aberto ao público, com ideias abertas a sugestões dos leitores. Além disso, essas possíveis pautas estão disponíveis para colaboração com jornalistas externos.

Diferente de outros veículos, o *Pudding* traz matérias focadas em visualização, utilizando frequentemente a interatividade como parte da narrativa. O estilo *scrollytelling*, popularizado a partir da grande reportagem multimídia *Snow Fall: The Avalanche at Tunnel Creek*⁶, do *New York Times*, produzida em 2012, é frequentemente utilizado por oferecer uma forma imersiva de visualização. Os elementos textuais são pensados como complemento das partes gráficas, contextualizando o que não pode ser mostrado.

Destaca-se novamente o fato de serem jornalistas-engenheiros. Em redações que há separação entre as duas áreas e não há participação direta dos jornalistas na construção de visualizações, pode ocorrer uma drástica redução das possibilidades de narrativa segundo Machado (2007, p.120). Ainda assim, nem todos os jornalistas têm habilidades computacionais suficientes para essa independência. O *Pudding* é um veículo cuja equipe utiliza linguagens de programação, e essas são as principais na etapa da visualização: *HTML* é utilizada na construção de páginas *web*, *CSS* trata da estilização dessas páginas, e *JavaScript* é responsável pela interatividade. Para gráficos estáticos ou não-interativos, eles utilizam a linguagem *R*, mas também o *Figma*, uma plataforma de projetos de *design* com foco *web*, ou o *Flourish*, um *site* que permite a criação de gráficos a partir de modelos prontos (THE PUDDING, 2017).

A grande variedade de ferramentas disponíveis para a construção de visualizações e matérias de dados é reflexo do interesse nesse tipo de conteúdo. Matt Daniels, no artigo sobre o conceito de jornalistas-engenheiros, menciona que a comunidade *r/dataisbeautiful*⁷ na rede social *Reddit* tinha 4,3 milhões de inscritos, e isso era indicativo do potencial do mercado (DANIELS, 2015). Atualmente, a comunidade conta com 15,7 milhões de inscritos. Em postagem sobre o que era previsto para o *Pudding* em 2019, Daniels explica que o trabalho com visualizações tem impacto para futuros jornalistas:

⁶ Disponível em <https://www.nytimes.com/projects/2012/snow-fall/index.html>. Publicada em dez. 2012, acesso em maio 2021.

⁷ Disponível em <https://www.reddit.com/r/dataisbeautiful/>. Acesso em abr. 2021.

um leitor pode ser influenciado a se tornar jornalista de dados a partir da dedicação de um autor em contar uma história visualmente (DANIELS, 2019).

Como veículo nativo do digital, as narrativas baseadas em visualização do *Pudding* aproveitam as ferramentas disponíveis para o meio, desenvolvendo um tipo de conteúdo que não seria possível antes da cibercultura, considerando os recursos livres e *open source*. No cenário digital e com as novas possibilidades de interação e construção de narrativa, a dispersão do foco do leitor exige que a informação seja transmitida de forma interessante, e que os leitores sejam atraídos a, além de consumir o conteúdo, interagir com as visualizações. Como destacado por Garland (2018), o *Pudding* promove uma cultura de conhecimento compartilhado por meio de uma seção detalhada no site com tutoriais sobre *scrollytelling* e vídeos mostrando o processo criativo da equipe. No *Github*, a equipe disponibiliza informações gerais, como o processo editorial e arquivos descrevendo como é trabalhar com cada um dos jornalistas-engenheiros, códigos e bases de dados utilizadas nas matérias.

Para Jan Diehm, jornalista-engenheira do *Pudding*, dados são um recurso poderoso no jornalismo por representarem uma verdade objetiva, mas podem ser coletados e visualizados de tantas formas diferentes que se torna importante ser transparente sobre metodologia e alertar os leitores para possíveis erros (GARLAND, 2018). Além disso, Matt Daniels afirma que a melhor forma de aprender jornalismo de dados é por meio de projetos (DANIELS, 2019).

As práticas de transparência do *Pudding* ajudam a difundir o conhecimento sobre jornalismo de dados e podem ser um recurso valioso para quem pretende desenvolver projetos na área. Como parte da rotina de trabalho, eles utilizam o *Github* e se preocupam em documentar os códigos e bases de dados que utilizam – um procedimento fundamental para uma equipe integrada, mas que também pode servir de tutorial para outros profissionais e estudantes.

Para melhor contextualização, foram selecionadas três matérias que ilustram como a transparência metodológica é fonte de aprendizado. Ainda que alguns procedimentos sejam avançados ou dependam de conhecimentos de programação, é possível observar as etapas do trabalho com dados e adaptá-los de acordo com as capacidades individuais. Nota-se a abrangência do que pode ser construído no jornalismo de dados, tanto pela variedade dos temas quanto pelas diferentes ferramentas utilizadas.

*Women's Pockets are Inferior*⁸ traz a comparação entre bolsos frontais femininos e masculinos e uma análise visual do que cabe nesses bolsos. A matéria representa uma coleta de dados física, envolvendo medição de bolsos de calças *jeans*, e a pauta parte de uma premissa observável no dia a dia: bolsos frontais de calças *jeans* femininas são menores. Com os dados coletados, a questão é comprovada e quantificada, sendo possível notar a diferença independentemente da variedade entre marcas. Como consequência, menos itens cabem nesses bolsos – a segunda questão levantada é quais itens cabem. Após a exposição dos dados, há uma contextualização histórica e chega-se à conclusão de que a diferença entre tamanhos de bolsos tem relação com a dinâmica de poder entre homens e mulheres na sociedade.

*The Structure of Stand-Up Comedy*⁹ mostra, a partir de uma análise da apresentação da comedianta Ali Wong disponível na *Netflix*, a estrutura narrativa de *stand-up comedy*. Em um vídeo que detalha a metodologia, dois dos jornalistas-engenheiros comentam que a ideia surgiu nove ou dez meses antes da publicação a partir da observação de que uma apresentação de *stand-up* possui vários *callbacks*, que são referências a piadas anteriores. A equipe olhou quatro outras apresentações, obtiveram o arquivo de legendas da *Netflix* com os tempos e registraram manualmente os *callbacks*, segundos de risadas e níveis das piadas, uma métrica que criaram para determinar piadas dentro de outras, que seriam os arcos narrativos. A ideia não surgiu com o ângulo apresentado na matéria, e inicialmente eram consideradas múltiplas apresentações, mas notou-se que era possível construir a narrativa a partir de apenas uma análise.

*Life After Death on Wikipedia*¹⁰ é a matéria com metodologia mais detalhada pelo *Pudding*, apresentada em uma *playlist* com 51 vídeos¹¹ que mostram o processo completo de produção. A pauta considera como a morte de celebridades é refletida em acessos na Wikipédia, e a metodologia tem início com a exploração da *Wikipedia API*. Esse é um exemplo de como o trabalho com dados muitas vezes se difere das práticas usuais: foi preciso utilizar uma interface para desenvolvedores a fim de obter as informações necessárias. Como não é esperado que os jornalistas saibam utilizar este tipo de ferramenta, há a recomendação de buscar outros projetos que utilizaram esses recursos e

⁸ Disponível em <https://pudding.cool/2018/08/pockets>. Publicada em ago. 2018, acesso em set. 2021.

⁹ Disponível em <https://pudding.cool/2018/02/stand-up/>. Publicada em fev. 2018, acesso em set. 2021.

¹⁰ Disponível em <https://pudding.cool/2018/08/wiki-death/>. Publicada em ago. 2018, acesso em set. 2021.

¹¹ Disponível em <https://www.youtube.com/playlist?list=PLsuhXm2zs07JuSfrNentA3DxAbaFO7ay2>. Publicada em jul. 2018, acesso em set. 2021.

que possuem metodologia detalhada, e a equipe de jornalistas-engenheiros realizou essa etapa de pesquisa. A metodologia também destaca a importância da análise exploratória dos dados devido ao grande volume de informação disponível, bem como a contextualização para que os leitores entendam a dimensão do que é apresentado.

A partir das matérias observadas, reforça-se a ideia da transparência metodológica como ferramenta de aprendizado, visto que são detalhadas as etapas do trabalho com dados. Outro ponto importante é a variedade de temas, que serve de estímulo para que se busquem pautas no cotidiano, mostrando que é possível abordar diversos assuntos de forma quantitativa e construir narrativas e visualizações que informam e atraem o público.

CONSIDERAÇÕES FINAIS

O jornalismo de dados é visto como novidade por ter uma práxis diferente a partir das evoluções tecnológicas. Ainda que não esteja em todas as redações, se apresenta como possibilidade para os atuais e futuros jornalistas e se conecta com a cibercultura e com a convergência, que impactaram a produção jornalística, o consumo de notícias e o acesso ao conhecimento. O objetivo do profissional é o mesmo – informar –, mas as ferramentas disponíveis possibilitam a análise de bases de dados massivas e a criação de visualizações para melhor compreensão. Enquanto o trabalho com dados aumenta a credibilidade, veículos que compartilham suas metodologias e praticam a transparência colaboram com jornalistas interessados nas técnicas e dão aos leitores a possibilidade de chegarem às suas próprias conclusões.

A transparência metodológica, que se relaciona com a cultura *hacker*, é fundamental para estudantes da área. *The Pudding* é um dos veículos que servem de manual. Além de registrar as metodologias nas matérias, a equipe do *Pudding* compartilha seus códigos no *GitHub* e produz vídeos no *YouTube* para explicar fases do processo editorial. O *Pudding* se destaca pelas narrativas de dados com foco na visualização, por tratar de temas fora do ciclo de notícias e pelas práticas éticas. Apesar da deontologia do jornalismo não incluir o uso de dados, o *ethos* profissional permite observar que as boas práticas não se afastam do que já era feito, apenas há uma adaptação aos meios digitais.

As etapas do trabalho com dados também são apenas adaptações da rotina jornalística. Apuração, checagem e redação se tornam coleta, limpeza, análise e contextualização. As visualizações mostram o impacto dos dados e são fundamentais para transmitir a informação. No *Pudding*, diferente da maioria dos veículos, as partes textuais

complementam a visualização. Assim, pela necessidade de total personalização, a equipe de jornalistas-engenheiros utiliza linguagens de programação para construir gráficos e recursos visuais, mas deixa claro em seus tutoriais que esse não é o único caminho.

Aprender jornalismo de dados tem mais relação com aprender a lógica do trabalho com dados do que aprender ferramentas complexas e linguagens de programação. É possível construir matérias de dados utilizando planilhas e *sites* para construir gráficos a partir de modelos prontos. As ferramentas gratuitas disponíveis atualmente oferecem resultados de qualidade profissional – no *Pudding*, por exemplo, todas as matérias são construídas com esses recursos gratuitos.

Por não abordar pautas quentes, o *Pudding* consegue desenvolver reportagens aprofundadas de temas da cultura popular e criar conteúdos impactantes. Além disso, não depende de financiamento externo, possibilitando um posicionamento diante de causas sociais. Com liberdade editorial, é um veículo independente que se sustenta por meio do *Polygraph*, agência da mesma equipe que é contratada para criar visualizações. Além da transparência metodológica, essa independência aumenta a credibilidade do *Pudding* e o torna uma ferramenta de aprendizado, como foi exemplificado nesta pesquisa.

REFERÊNCIAS

ARAÚJO, L. V. D. Apontamentos iniciais sobre a prática do jornalismo de dados à luz da ética profissional. *Âncora*, João Pessoa, v. 3, n. 2, p.179-193, jul./dez. 2016.

BERRET, C.; Phillips, C. **Teaching data and computational journalism**. Nova York: Columbia Journalism School, 2016.

DANIELS, M. State of the Pudding, 2019. **Medium**, 2019. Disponível em: <https://medium.com/@matthew_daniels/state-of-the-pudding-2019-3562f80b503d>. Acesso em: 2 mar. 2021.

DANIELS, M. The Journalist-Engineer. **Medium**, 2015. Disponível em: <https://medium.com/@matthew_daniels/the-journalist-engineer-c9c1a72b993f>. Acesso em: 15 mar. 2021.

DANIELS, M. The State of The Pudding, 2018. **Medium**, 2017. Disponível em: <https://medium.com/@matthew_daniels/the-state-of-the-pudding-2018-9661ab4d299c>. Acesso em: 15 mar. 2021.

GARLAND, E. **The Unsung Power of 'Scrollytelling'**, 2018. Disponível em: <<https://www.arcgis.com/apps/Cascade/index.html?appid=9b821a8e527d4dd28298be79c23fcfa1>>. Acesso em: 28 out. 2020.

GEHRKE, M. Transparency as a key element of data journalism: Perceptions of Brazilian professionals. **Computation + Journalism Syposium 2020**. Boston: [s.n], 2020.

GOLDENBERG, R. Editorial Process. **Github**, 2017. Disponível em: <<https://github.com/the-pudding/FYI/wiki/Editorial-Process>>. Acesso em: 19 mar. 2021.

GRAY, J.; Bounegru, L.; Chambers, L. (org.) **Manual do Jornalismo de Dados**. [S.l.]: O'Reilly Media, 2012.

JENKINS, H. **Cultura da convergência**. 2. ed. São Paulo: Aleph, 2009.

KARAM, F. J. C. **Jornalismo, ética e liberdade**. 1. ed. São Paulo: Summus Editorial, 2014.

MACHADO, E. A Base de Dados como espaço de composição multimídia. In: BARBOSA, S. (org.) **Jornalismo Digital de Terceira Geração**. Covilhã: Livros LABCOM, 2007. p.111-126.

MIELNICZUK, L.; Träsel, M. Jornalismo guiado por dados como inovação profissional e seus desafios para a educação. **Contemporanea: Revista de Comunicação e Cultura**, Salvador, v. 15, n. 2, p.609-629, maio/ago. 2017.

SANTAELLA, L. **Culturas e artes do pós-humano: Da cultura das mídias à cibercultura**. 4. ed. São Paulo: Paulus, 2010.

SER, K. K. K. **Best Practices for Data Journalism**. [S.l.]: MDIF Media Dvisory Services, 2018.

THE PUDDING. About. **The Pudding**, 2017. Disponível em: <<https://pudding.cool/about/>>. Acesso em: 19 mar. 2021.

TRÄSEL, M. **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) - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2014.

VENTURA, M. P.; Paulino, R. D. C. R. Jornalismo de Dados na América Latina: o perfil do ensino. In: MEDITSCH, E.; Kronbauer, J.; Bezerra, J. F. (org.) **Pedagogia do Jornalismo: desafios, experiências e inovações**. Florianópolis: Editora Insular, 2020. p.247-261.