

DOSSIÊ

PAUTA GUIADA POR ALGORITMO:

conteúdo e forma no *Estadão Infográficos*



DANIELA BORGES DE OLIVEIRA

Universidade Metodista de Piracicaba, Piracicaba – Brasil
ORCID: 0000-0003-2971-3840

BELARMINO CESAR GUIMARÃES DA COSTA

Universidade Metodista de Piracicaba, Piracicaba – Brasil
ORCID: 0000-0003-0808-8708

DOI: 10.25200/BJR.v16n3.2020.1284

Recebido em: 26/03/2020

Desk Review: 04/05/2020

Editores de Desk Review: David Domingo e Laura Storch

Revisado em: 08/10/2020

Aprovado em: 03/11/2020

RESUMO – O artigo observa como o *Estadão Infográficos* emprega algoritmo para prospecção de pautas e formatação do conteúdo, com o propósito do uso da visualidade tanto no processo de investigação quanto na apresentação das notícias. O objetivo é demonstrar como o webjornalismo está se apropriando das ferramentas do ambiente digital para inovação e diferenciação. Assim, o percurso metodológico consiste em três etapas: entrevista a componentes da equipe responsável pelo menu, seleção de reportagens da seção que utilizam técnicas de inteligência artificial e, por fim, análise qualitativa de investigação de conceitos que explicam os efeitos cognitivos provocados pela técnica nas narrativas jornalísticas, sendo: visualização de dados, imagem-síntese e inconsciente ótico. A investigação conclui que o uso de algoritmo na produção de visualizações permite percepção de novas pautas e potencializa a comunicação com o público na internet, mesmo que tal recurso ainda seja pouco explorado pelo jornal *O Estado de S. Paulo*.

Palavras-chave: Visualização de dados. Inteligência artificial. *Estadão Infográficos*. Algoritmo.

NEWS SCRIPT GUIDED BY ALGORITHM: content and shape in Estadão Infographics

ABSTRACT – This paper looks at how *Estadão Infográficos* use algorithms to examine news scripts and format content; its purpose being to use visuality both in the investigation process and for how the news is presented. We intend to demonstrate how web journalism appropriates the digital tools for innovation and differentiation. Thusly,

the methodological procedure was conducted in three stages: interviewing the team members responsible for the menu, selecting those reports that use artificial intelligence techniques, and lastly, conducting a qualitative analysis upon specific concepts to explain cognitive effects that its news narratives create, which are: data visualization, image-synthesis and optical unconscious. Our investigation concluded that the use of algorithms to produce visualizations provides different perspectives of news scripts and builds communication with the public on the internet, despite the fact that this newspaper (*O Estado de S. Paulo*) does not make much use of this resource.

Key words: Data visualization. Artificial intelligence. *Estadão Infográficos*. Algorithm.

PAUTA GUIADA POR ALGORITMO: contenido y forma em Estadão Infográficos

RESUMEN – El artículo observa cómo *Estadão Infográficos* emplea algoritmo para prospectar pautas y formatear contenido, con propósito de usar visualidad en el proceso de investigación y en la presentación de noticias. El objetivo es demostrar cómo el periodismo web se apropia de las herramientas del entorno digital para innovación y diferenciación. Así, la trayectoria metodológica consta de tres etapas: entrevista con miembros del equipo responsable del menú, selección de informes de la sección que utilizan técnicas de inteligencia artificial y, finalmente, análisis cualitativo de la investigación de conceptos que explican efectos cognitivos provocados por la técnica en las narrativas de periodismo, siendo: visualización de datos, síntesis de imagen e inconsciente óptico. La investigación concluye que el uso del algoritmo en la producción de visualizaciones permite percepción de nuevas pautas y mejora la comunicación con el público en internet, aunque este recurso todavía es poco explorado por *O Estado de S. Paulo*.

Palabras clave: Visualización de datos. Inteligencia artificial. *Estadão Infográficos*. Algoritmo.

1 Introdução

Desde sua origem (anos 1960), a internet é ambiente de interação e simulação que permite à humanidade recuperar e produzir mensagens virtuais de forma descentralizada. Em sua flexibilidade, por meio de código binário (0 e 1) viabiliza tradução em texto, imagem e som, e recombinação para gerar novos significados com aceleração nunca vista antes na sociedade. Outra qualidade é a linguagem hipertextual da navegação em conexões complexas, multissensoriais e marcadas pela lógica da não linearidade. Santaella (2004) menciona a figura do leitor imersivo, caracterizado pela leitura multilinear, que compara, sintetiza e estabelece nexos em permanente transformação. No ciberespaço, observa-se um movimento semelhante ao dos objetos fractais.

Apesar do meio virtual se constituir de memória ampla e circulação de intenso fluxo de informações, os efeitos na sociedade

e nos usuários não garantem real compreensão dos conteúdos, que podem não ser percebidos, ou ser rapidamente esquecidos, comprometendo a noção de experiência (Zuin & Zuin, 2015; Costa, 2002). Na alegoria do ciber-*flâneur* Lemos (2001, pp.50-51) explica a navegação na rede que ocorre à “deriva por espaços descontínuos, fragmentados e efêmeros”. O observador está fisicamente alocado ao mundo real e, simultaneamente, frente a uma abundância de dados e informações virtuais. Ao atuar com os fatos sociais e a estruturação de informações, um dos maiores desafios do jornalismo é garantir que as mensagens cheguem efetivamente ao receptor e provoquem reflexão, antes que este consuma e distribua conteúdos sem verificação e credibilidade.

A incomunicação na internet, ou a falha no processo comunicativo de troca de mensagens, é questão a ser observada pelo profissional da comunicação (Wolton, 2011). Outro ponto é o processo de desinformação, pela falta de contextualização dos fatos ou produção de conteúdos travestidos de verdade, como a pseudociência propagada nas redes sociais e sites que assumem caráter de verdade insofismável, influenciando na agenda política e no comportamento dos usuários. Diante desse contexto, cabe observar como o jornalismo pode contribuir para interpretar o fluxo contínuo de informações circulantes na rede e mediar educativamente para que o público tenha condições de perceber a distorção e o engodo.

Dentre as adaptações da profissão, destaca-se o uso de programas especializados em mineração e tradução de dados para estratégias da narrativa jornalística, como criação de visualizações e infografias, presentes em experiências de veículos tradicionais que se aventuram na busca por soluções gráfico-visuais e de produção de conteúdo para formatos que dependem da computação gráfica. Neste estudo, são analisadas narrativas que utilizam inteligência artificial e visualização de dados pelo *Estadão Infográficos*¹, menu da plataforma digital do jornal *O Estado de S. Paulo*, também conhecido pelo nome *Estadão*², como exemplo de setor que surge em portais de jornais consagrados e como tais estratégias alteram a forma de pensar e conceber os produtos jornalísticos na busca pela diferenciação na internet.

Nesse sentido, é necessária retrospectiva para entender as ferramentas utilizadas na produção de matérias pelo *Estadão Infográficos*, baseadas principalmente no uso de algoritmo, e a razão de a editoria ser nomeada como infográfica, mesmo utilizando mais formatos. A escolha d'*O Estado de S. Paulo* deve-se à sua importância

em número de assinaturas digitais e relevância na implantação de infografias no jornalismo brasileiro, sendo considerado o primeiro jornal a circular com tal formato. Em 2019, o portal recebeu oito premiações por infografias multimídia³, cabendo observar seu movimento de adaptação ao digital como modelo de negócios em comunicação.

Para isso, a metodologia se estrutura nas fases de seleção do *corpus* de análise – três reportagens de agosto a setembro de 2019 do *Estadão Infográficos* –, bem como execução de entrevistas semiestruturadas com o Editor Assistente de Multimídia, Vinicius Sueiro, e o jornalista de dados da seção, Rodrigo Menegat, à época. A tensão entre objeto e fundamentação teórica é conduzida pelos conceitos: visualização de dados (Cairo, 2016), inconsciente ótico (Benjamin, 1987) e imagem-síntese (Quéau, 1993). As categorias explicitam a relação entre visualidade e ampliação da percepção, com o uso de gráficos e imagens que representam conexões entre dados, e são essenciais para esclarecer a relevância do algoritmo para descoberta de novas pautas jornalísticas, que devem prender a atenção do público no meio digital de distração. Assim, também o artigo pretende demonstrar que a internet não é espaço somente para o efêmero, podendo comportar, como costuma ser associado ao meio impresso, profundidade na narrativa.

A fim de compreender o processo de produção do *Estadão Infográficos*, o atendimento do portal foi contatado⁴ pedindo informações da seção. Em resposta, a autora foi convidada via aplicativo de mensagens pelo então Editor Assistente Multimídia, Vinicius Sueiro, a visitar pessoalmente a editoria no prédio do *Estadão*, em 03 de setembro de 2019⁵. O encontro, das 15h às 18h, foi conduzido por entrevista semiestruturada com questões sobre o objetivo do menu, regularidade das publicações, composição da equipe, ferramentas utilizadas, sobre a escolha das pautas e inovações que o *Estadão* propõe para a *web*. Na ocasião, perguntas foram acrescentadas aproveitando a participação do jornalista do *Estadão*, Rodrigo Menegat.

Outro aspecto do artigo é constatar a existência de diferentes aplicações de algoritmos para construção de narrativas jornalísticas, que não sejam dependentes integralmente da automatização: *softwares* reúnem dados e produzem reportagens e artigos completos com velocidade sobre-humana, com moldes nas técnicas do jornalismo, como *StatsMonkey* e *Narrative Science* (Kaufman, 2019). Nesta pesquisa, a linguagem matemática de programação,

o design e as técnicas jornalísticas são apresentados em ação de interdisciplinaridade para construção dos conteúdos, como veremos adiante. Porém, é preciso primeiro entender o conceito de algoritmo e como vem sendo usado pelo jornalismo.

2 Algoritmo: o que é e auxílio no jornalismo

Antes de compreender o jornalismo munido do algoritmo como ferramenta, é necessário esclarecer o que se entende por “algoritmo”. O termo surge frequentemente no contexto das redações interdisciplinares de portais online, apesar de observado que jornalistas normalmente não são capacitados para o uso de algoritmos e trabalho com dados – em entrevista a profissionais de 149 países, incluindo na América Latina, 79% disseram querer treinamento em análise de dados (*International Center for Journalists* [ICFJ], 2019). A equipe do *Estadão Infográficos*, por exemplo, é composta majoritariamente por designers e jornalista com habilidade de programação, aptidão essencial na produção de formatos inovadores para a *web* e no tratamento de dados segundo observações do Editor Assistente, Vinicius Sueiro, e do jornalista Rodrigo Menegat.

O algoritmo remonta à aplicação manual no século IX, com o matemático al-Khwārizmi (Kaufman, 2019), e sua definição depende do contexto social (Dourish, 2016). No ambiente digital: “um algoritmo pode ser definido como uma série de passos realizados para resolver um problema particular ou alcançar determinado resultado” (Diakopoulos, 2015, p.400, tradução nossa⁶). Já Kaufman (2019, p.53) destaca: “algoritmo é um conjunto de instruções matemáticas, uma sequência de tarefas para alcançar um resultado esperado em um tempo limitado”, o que supõe a definição de regras e a capacidade de ação. Considerando o fator humano, o algoritmo pode guiar e influenciar sua perspectiva nos procedimentos de criação de conteúdo:

Decisões de algoritmos podem ser baseadas em heurística e regras, ou cálculos sobre conjuntos massivos de dados. Regras podem ser articuladas diretamente por programadores, ou serem dinâmicas e flexíveis baseadas em *machine learning*. Algumas vezes, um operador humano tem controle na decisão final do processo, mas mesmo neste caso o algoritmo direciona a atenção do operador para um conjunto de informações. (Diakopoulos, 2015, p.400, tradução nossa⁷).

De modo geral, Lindén (2017, p.62, tradução nossa⁸) define algoritmo como conjunto de “operações passo a passo a serem executadas, como cálculo, processamento de dados e raciocínio automatizado – um conjunto de regras que precisamente define uma sequência de instruções precisas que podem ser entendidas por um computador”. Dentre as atividades que realiza, estão a priorização, classificação, associação, filtragem e distribuição de dados (Diakopoulos, 2015; Coddington, 2015).

Para grandes corporações, como os “*Big Five*” Google, Amazon, *Facebook*, Apple e Microsoft, o algoritmo é estratégico na apropriação de dados dos usuários e personalização dos conteúdos apresentados ao público, com base nos perfis de busca. Por isso, o algoritmo adquire condição de estruturador de sistemas da sociedade, o que pode ameaçar a liberdade humana nos processos de tomada de decisões. É no que consiste a “algocracia”, sistema “em que algoritmos são usados para coletar, agrupar e organizar dados a partir dos quais decisões são tipicamente tomadas, e para ajudar no modo como esses dados são processados e comunicados através de um relevante sistema de governança” (Danaher, 2016, p.04, tradução nossa⁹). Ao considerar esta visão, Danaher (2016) pontua preocupações com a ocultação dos dados utilizados pelo sistema e a opacidade desse processo, inacessível para a compreensão pelo ser humano. É nesse sentido que vai na contramão da definição essencial de transparência no jornalismo de dados e computacional (Coddington, 2015).

A investigação aborda o uso do algoritmo e da visualização de dados como estratégias no processo jornalístico de produção de narrativas. Não enfatiza a plena automação que exclui o fator humano na produção de notícias, reportagens, notas e artigos (Lindén, 2017; Kaufman, 2019). Para tanto, o algoritmo seleciona informações de acordo com critérios, o que se torna outro problema visto que, apesar de guiado por deontologia e critérios que sistematizam a produção, o jornalismo não estabelece regras fixas e universais, sendo que a cultura organizacional, regras e códigos estão em constante mudança, apropriados durante o processo produtivo. Ou seja, são inconsistentes para compor regras de um algoritmo (Lindén, 2017). Seria difícil delegar à máquina o controle de toda a produção jornalística, pois depende da experiência e análise humana.

Para Lindén (2017, p.64, tradução nossa¹⁰), “um algoritmo jornalístico precisa refletir o comportamento da profissão que

não é sempre baseado em regras, mas reflexivo e flexível no sentido de que a aplicação de regras explícitas e internalizadas em situações complexas é ambígua”. Porém, o algoritmo pode ser um potencializador do trabalho humano, guiado pelo jornalista, ao auxiliar na prospecção de pautas, confirmação de hipóteses e observação de dados discrepantes.

Tradicionalmente, jornalistas confiaram mais em métodos qualitativos para suas investigações, mas com auxílio de *softwares* fáceis de usar, baratos, ou mesmo gratuitos, para processar grandes bases de dados, novas oportunidades surgem. A capacidade de memória pessoal e recursos de processamento analítico já foram aumentados com auxílio dos computadores. (Lindén, 2017, p.64, tradução nossa¹¹).

É nesse âmbito que plataformas de visualização de dados utilizam o recurso para facilitar a investigação de bases complexas de dados. Se pensado como facilitador ao garimpar dados e unidades informativas na rede, o algoritmo – muitas vezes unido ao uso de gráficos e visualizações – permite a percepção de novas tendências, fugindo da pressão do *hard news*.

Não é mais suficiente a transposição de conteúdos jornalísticos para a internet. O meio digital deve ser pensado em suas ferramentas, formatos e possibilidades, cujo percurso de navegação e interação do público é mais amplo. Sendo a linguagem algorítmica a essência do digital, o uso das bases de dados passa a ser o foco da nova geração do webjornalismo (Dalmonte, 2009), que deve pensar as possibilidades estéticas das plataformas online, como observa Martinez (2007, p.24): “informar bem também passará pelo planejamento e design, pela formatação bem pensada da ‘moldura’ que o *browser* ou navegador apresenta”. Com isso, recursos visuais e multimídia tornam-se chave para novas percepções sobre os fatos e ainda contribuem para melhorar a comunicação com o público, pois vivemos numa sociedade excitada visualmente, que enxerga na ilustração uma linguagem universal e imediata.

Para aplicar coleta, interpretação, análise e tradução de dados com auxílio da máquina, o jornalista necessita conhecer técnicas de programação que permitam obter interrelações entre variáveis quantitativas, extraídas de valores numéricos para a percepção de tendências. Com a escala potencializada de informações na internet, há vasto material disponível – segundo o relatório de 2018 da *Digital Universe Study* (IDC), a expectativa é que em 2025 o volume seja de 175 zettabytes de informação digital – e que, somente com

a capacidade cognitiva humana e suas limitações físicas, seria inimaginável explorar. Não só os dados já existentes em bases digitais, mas características observáveis, como expressões faciais, são reconhecidas, quantificadas e organizadas mais rapidamente por meio do uso da máquina, do algoritmo.

Um exemplo é a reportagem “O que revela uma análise das emoções dos candidatos durante o debate”¹² (13/09/2018), do *Estadão*, resultado da análise das expressões faciais dos presidencialistas brasileiros em entrevista concedida ao jornal. Usando quatro câmeras voltadas a seus rostos, a Interface de Programação de Aplicativos de detecção facial da Microsoft – baseado em inteligência artificial que reconhece atributos na imagem, como as emoções em rostos –, e contando com pesquisas no campo da psicologia e da ciência da computação, a equipe registrou o percentual dos sentimentos de: alegria, desprezo, medo, raiva, surpresa, tristeza e neutralidade. Para isso, cada segundo das filmagens dos seis entrevistados foi transformado em imagem, que foram submetidas ao aplicativo. A reportagem foi produzida por jornalistas em conjunto com designers.

As estratégias utilizadas pelo jornalismo baseado na visualização de dados, no uso de algoritmo e na inteligência artificial serão explicitadas a seguir, com exemplos de reportagens multimídia do *Estadão Infográficos*. Por hora, é preciso observar a tendência desse veículo, que assim como outros tradicionais, como a *Folha de S.Paulo*¹³, define uma editoria específica intitulada como infografia para incubar essas produções, que muitas vezes não são consideradas infográficos. Para isso, é preciso diferenciar infografia do conceito de visualização de dados, próprio do jornalismo guiado por dados.

3 Editoria de formato: *Estadão Infográficos*

As produções multimídia do jornalismo, dentre outras atribuições, exigem conhecimento em *softwares* para criação de gráficos, imagens interativas, ilustrações e sobre linguagem de programação para páginas dinâmicas na web e inserção de visualizações de dados. Visto que muitas vezes o jornalista por si só não tem o conhecimento necessário dessas ferramentas, são necessárias equipes interdisciplinares para produções complexas em jornalismo de dados. Na internet, são diversas as experiências de veículos e menus no ramo.

O mais recente relatório do *International Center for Journalists* (2019) revela interesse crescente de jornalistas no treinamento em análise de dados e a importância do jornalismo de dados, segundo os aspectos: “jornalismo de dados ajuda a engajar a audiência, aprimora a qualidade e aumenta a produtividade” (ICFJ, 2019, p.10, tradução nossa¹⁴). No Brasil, são ilustrativas as experiências do jornal digital *Nexo*¹⁵ e da organização de mídia *Gênero e Número*¹⁶ – nativos digitais –, enquanto que veículos tradicionais separam seções para produções multimídia, como *Infográficos da Folha*¹⁷ e o *Estadão Infográficos*. Este, menu fixo do portal do *Estadão*, possui maior regularidade nas publicações e editoria própria para as produções.

O primeiro infográfico brasileiro “A navegação brasileira” foi registrado no jornal *O Estado de S. Paulo*, em 18 de agosto de 1909 (Teixeira, 2010). Assim, é antigo o uso de infografia no veículo, que se inicia no formato impresso, depois é copiada para a internet para, então, ser pensada no formato digital. Por isso, talvez, seja usado o termo “Infográficos” na maioria dos menus da mídia destinados a formatos multimídia, compostos por ilustrações, dados, gráficos e recursos interativos. Formados por união de texto e imagem, indissociáveis, infográficos apresentam contextualização de um fenômeno específico de forma explicativa e cronológica. São incluídos no gênero informativo, já que apresentam a lógica narrativa de um texto: título, lide, indicação de fontes e autores (Teixeira, 2010). É pressuposto, portanto, que assim como o texto jornalístico o infográfico mostre apenas uma seleção das informações de um fato para demonstrar um ponto de vista específico (Cairo, 2016).

Em entrevista com o então Editor Assistente Multimídia do *Estadão*, Vinicius Sueiro¹⁸, foram traçadas a estrutura, objetivos e práticas da editoria responsável pela maior parte das produções multimídia do *Estadão Infográficos*, que envolvem uso de visualização de dados, algoritmo e inteligência artificial. A Editoria Multimídia divide-se nos setores de Fotografia, Diagramação e Infografia que, por sua vez, desdobra-se na divisão Template, Pura e Digital. É na seção “Digital” que a equipe se responsabiliza por produzir novos formatos para apresentação de conteúdos e dados, que podem tornar-se modelos reproduzíveis em outras pautas – os templates. Até setembro de 2019, a equipe do Digital responsável pela criação de formatos inovadores era composta por seis pessoas, sendo o Editor Assistente Multimídia, Vinicius Souza Sueiro, o jornalista visual e de dados, Rodrigo Menegat, três web designers e um estagiário de jornalismo visual.

Como destaca Sueiro na entrevista, o *Estadão Infográficos* é uma “editoria de formato” mais do que “de conteúdo”, de forma que seu objetivo compreende explorar formas inovadoras de se produzir e apresentar notícias e reportagens. Aqui entram duas questões para que esse trabalho seja possível: o setor não está sujeito à mesma pressão da produção para o *hard news*, ou seja, as pautas muitas vezes fogem da agenda diária; e leva-se mais tempo para as matérias serem finalizadas, podendo seguir alguns dias ou meses de produção.

Por ser página que permite inserir produtos mais pesados e imagens, e por isso mais flexível, o *Estadão Infográficos* abriga matérias multimídia de outras editorias da redação, que utilizam recursos como galerias de fotos e efeitos de reportagem paralaxe, com imagens maiores e plano de fundo. Dessa forma, entende-se que não é dado destaque diferenciado às produções baseadas em jornalismo de dados como estratégia de captação de público que, inclusive, costuma acessar as matérias do menu Infográficos por meio dos links compartilhados nas redes sociais, e não diretamente pelo site.

Para as reportagens do *Estadão Infográficos*, são utilizadas ferramentas como *Adobe Illustrator*, *Data-Driven Documents (D3)*, *Datawrapper*, *Data Viz Project*, *Flourish*, *Python*, *Rawgraphics* e o programa criado pelo próprio *Estadão* para facilitar a produção de visualizações e incorporação no portal, o UVA. Percebe-se o uso de *softwares* para criação de imagens e gráficos, edições estéticas de ilustrações e para a criação de visualizações de dados, com uso de linguagens de programação, a exemplo do *Python* para processar dados e criar páginas dinâmicas na internet, além de catálogo de funções e tipos gráficos (*Data Viz Project*). Com o *Data Viz Project*, são orientadas as escolhas dos formatos e funções de gráficos utilizados nas matérias.

Apesar da estrutura da editoria e do histórico do veículo na produção imagética, há falta de referências em manual ou orientações para a elaboração de infografias e visualizações de dados no portal do *Estadão*, e essa condição reforça o aspecto experimental do *Estadão Infográficos* na exploração estética de formas narrativas. Apesar disso, para o jornalista de dados do veículo, Rodrigo Menegat, o objetivo do menu se assemelha à proposta do projeto *The Pudding*. Este veículo digital reúne histórias interessantes que se orientam pela narrativa essencialmente visual: “Recursos visuais tornam o seu argumento mais acessível e menos complexo do que uma matéria de mil palavras”¹⁹.

4 Jornalismo, visualização de dados e cognição

Na perspectiva de compreender os métodos que orientam o trabalho da equipe multimídia, responsável pela maioria das mais complexas produções do *Estadão Infográficos* – em vista do veículo não expressar referências para o webjornalismo praticado – torna-se fundamental investigar os modelos de jornalismo no ambiente digital, de forma a enquadrar as experiências d’*O Estado de S. Paulo*. Segundo Coddington (2015), as práticas jornalísticas vêm se aproximando da área de programação, marcadas pela cultura do código aberto e o aumento do acesso a dados e compartilhamento de informações. Para compreendê-las, o autor (Coddington, 2015) observa três dimensões do jornalismo na era da mediação digital: Reportagem Auxiliada por Computador (RAC), o jornalismo de dados e o jornalismo computacional.

Mais próximo das técnicas das ciências sociais, métodos da estatística e da reportagem investigativa, a RAC propõe as bases para o jornalismo de dados e o jornalismo computacional, mas difere de ambos no quesito de utilizar dados em amostragem, e por considerar o público como receptor passivo e ser guiada pelos princípios deontológicos do jornalismo (Coddington, 2015). Por se tratar de técnicas mais básicas da computação, como o uso da memória digital ou a troca de e-mails, a dimensão da RAC não será detalhada como as outras, mais próximas das experiências encontradas no menu do *Estadão*.

Marcadamente empregado a partir dos anos 2000, o jornalismo de dados pode ser caracterizado pelo uso de análise de dados e sua apresentação visual, com a utilização de técnicas de tratamento jornalístico e do design. Ao examinar dados, seria possível descobrir histórias e contá-las de forma personalizada. É o que esclarece Howard (2014, p.05, tradução nossa²⁰), quando afirma: “em sentido amplo, o jornalismo de dados está contando histórias com números ou encontrando histórias neles”, e Mancini e Vasconcellos (2016, p.75), ao diferenciar a reportagem produzida “com” dados e a reportagem “de” dados: diferente da primeira, que se utiliza dos dados de forma ilustrativa e complementar, a reportagem de dados é fundamentada pela relação obtida entre os dados, “os dados seriam a própria razão da reportagem”.

São qualidades do jornalismo de dados: a transparência no método de produção, como herança da cultura do código aberto; a exploração de largas bases de dados; e a possibilidade de o próprio

público abstrair reflexões sobre os dados apresentados, de forma ativa. “Ao desenvolver produtos de jornalismo de dados – frequentemente visualização de dados – sua utilidade para a audiência é consideração principal” (Coddington, 2015, p.343, tradução nossa²¹).

É importante levar em conta algumas habilidades que sustentam o trabalho de jornalismo de dados em uma redação. Segundo Gordon (2013), existem poucos profissionais habilitados para tratar das três dimensões necessárias em uma equipe de jornalismo de dados, sendo: o repórter assistido por computador, que entrevista os dados com conhecimento básico de alguns *softwares*; o desenvolvedor de aplicativos de notícias, que conhece programação para produção de interfaces que melhor atendam o usuário; e o especialista em visualização de dados, que sabe como apresentar dados de maneira acessível e interessante, também com conhecimento em programação.

A visualização de dados, tão citada para diferenciar o jornalismo de dados da RAC, é concebida por Cairo (2016) como exibição que permite exploração e descoberta pelo próprio público, uma vez que todos os dados se encontram presentes em uma mesma representação e podem ser explorados conforme a navegação, o que reforça a participação ativa do receptor:

Uma visualização de dados é uma exibição de dados projetada para permitir análise, exploração e descoberta. Visualizações de dados não se propõem principalmente a transmitir mensagens que estão predefinidas por seus designers. Pelo contrário, elas são frequentemente concebidas como ferramentas que permitem às pessoas extrair suas próprias conclusões dos dados. (Cairo, 2016, s/p, tradução nossa²²).

Entendimento semelhante é constatado pelo jornalista Rodrigo Menegat, para quem a visualização de dados com gráficos é forma de demonstrar um fenômeno quantitativo, abstrato, para facilitar a operação mental do leitor²³. É o que Cairo (2016, s/p, tradução nossa²⁴) chama de auxiliares cognitivos: “Gráficos que codificam informações funcionam como auxiliares cognitivos”. Por meio da representação visual, o processo de o indivíduo visualizar mentalmente um fenômeno é delegado à imagem apresentada, facilitando a absorção. A intenção é priorizar antecipadamente os aspectos a que o cérebro se atém (Cairo, 2013).

Em jornalismo computacional, a interatividade do público é outro aspecto marcante: “o objetivo do jornalismo computacional é fornecer a ele [o público] as ferramentas que precisam para executar

sua própria filtragem e abstração” (Coddington, 2015, p.343, tradução nossa²⁵). No caso, é a partir da coleta, análise e estruturação da informação que o jornalismo realiza abstração apoiado em vastas bases de dados por meio da computação. Essa prática só seria possível pela atividade colaborativa e pelo uso de algoritmos. Para Lindén (2017, p.71, tradução nossa²⁶), o pensamento computacional é: “a integração das dimensões lógica, algorítmica, científica e de inovação, da cognição humana, direcionadas à abertura para novas ideias, revolucionando todos os tipos de ocupações, incluindo o jornalismo”.

Dessa forma, verifica-se a proximidade que técnicas do jornalismo de dados e do jornalismo computacional possuem com relação à atuação do *Estadão Infográficos*, tanto nas práticas de estruturação de dados com programas e algoritmo quanto na composição da equipe. O aspecto da visualização de dados fica destacado, uma vez que a própria editoria, formada por profissionais de áreas da computação, design e do jornalismo, dá foco à produção de narrativas visuais no menu.

5 O imperativo da visualidade

Coddington (2015) observa que a abstração no jornalismo computacional é a habilidade cognitiva de desmembrar informações de sua condição material, como os dados, que resultam de percepções. Ao valer-se de recursos visuais, o observador jornalista é capaz de abstrair reflexões a partir de organizações e comparações de bases de dados. Segundo Cairo (2013), a visualização é um tipo de tecnologia na condição de ser extensão das capacidades humanas e auxiliar na resolução de problemas. É uma ferramenta denominada arte funcional, pois, para além da beleza do recurso imagético, propõe objetividade e precisão na transmissão de informações estruturadas, através da forma.

Com o distanciamento proporcionado pela imagem, a técnica possibilita encontrar padrões antes obscuros com a separação das informações que compõem a visualização: “o propósito de toda abstração é o de tomar distância do concreto para poder agarrá-lo melhor” (Flusser, 2008, p.16). Não é mais preciso um texto linear, pois a própria imagem pode ser portadora de informações, e mais especificamente, a imagem técnica (Flusser, 2008).

Em visualização de dados, o formato clássico apresentado em reportagens são os gráficos, que reúnem registros e demonstram

relações entre eles. Atualmente, nas editorias destinadas à produção de visualizações baseadas em dados, catálogos como o *Data Viz* são amplamente utilizados para compreender efeitos e escolher as formas. A reunião dos dados em única superfície de visualização atua – como Flusser (2008, p.24) revela sobre a imagem técnica – nas “tentativas de juntar os elementos pontuais em nosso torno e em nossa consciência de modo a formarem superfícies e destarte taparem os intervalos”.

Relembrando a ideia de infografia, cujo nome intitula o menu do *Estadão*, a imagem infográfica pode ser representada sob conceito de imagem-síntese produzida por simulação no ambiente digital a partir de linguagem matemática e que representa situações abstratas diversas (Quéau, 1993). É na ilustração de um acontecimento, como se dá um fenômeno ou para entender a estrutura de determinado objeto, que o jornalismo costuma recorrer à infografia. Essencialmente, quando fotos ou narrativa textual não são suficientes para explicar a dimensão de um fenômeno qualitativo.

Para compreender a ampliação da percepção pela visualização, é possível atualizar o conceito benjaminiano de “inconsciente ótico”, cunhado diante da realidade do surgimento da câmera fotográfica, mas possível de se apropriar para explicar os efeitos da criação de imagens virtuais, como o recurso da visualização de dados. A fotografia, para Benjamin (1987), fixa um momento real permitindo acuidade visual ao ser humano, que não é capaz de perceber todos os movimentos da natureza, como o exato instante em que todas as patas de um cavalo estão no ar durante o galope.

O gesto de pegar um isqueiro ou uma colher nos é aproximadamente familiar, mas nada sabemos sobre o que se passa verdadeiramente entre a mão e o metal, e muito menos sobre as alterações provocadas nesse gesto pelos nossos vários estados de espírito. Aqui intervém a câmara com seus inúmeros recursos auxiliares, suas imersões e emersões, suas interrupções e seus isolamentos, suas extensões e suas acelerações, suas ampliações e suas miniaturizações. Ela nos abre, pela primeira vez, a experiência do inconsciente ótico, do mesmo modo que a psicanálise nos abre a experiência do inconsciente pulsional. [...] os múltiplos aspectos que o aparelho pode registrar da realidade situam-se em grande parte fora do espectro de uma percepção sensível normal. (Benjamin, 1987, pp.189–190).

Em consideração às produções midiáticas resultantes da lógica da indústria cultural, Costa (2002) formula o juízo de que as mensagens atendem a padrões pré-conformados, medianos, que

interferem decisivamente nas possibilidades de produção estética das notícias. Essa característica, porém, envolve a agenda diária de notícias, a produção de conteúdos efêmeros e a velocidade de produção. No caso do *Estadão Infográficos*, o que se denota é uma quebra nessa lógica, diante da condição possível para o jornalismo investigativo, preocupado com o formato e visualidade do conteúdo, de dedicar-se à produção com prazos ampliados. Justamente na imagem que o jornalismo irá procurar a diferenciação e maneiras de se comunicar melhor com o usuário.

6 Análise: investigação do uso de visualização e algoritmo

Pontuados os caminhos que o jornalismo percorre para adaptar-se à internet, ao recorrer às ferramentas das tecnologias digitais que transformam as formas de investigação, captação de informação, análise, checagem e estruturação das notícias, a pesquisa apresenta exemplos de conteúdos do *Estadão Infográficos* que permeiam o uso de inteligência artificial, técnicas de programação e design. Dessa forma, é possível registrar técnicas e soluções que podem ser apropriadas por outras mídias e na formação do profissional de comunicação. São tratadas três reportagens da Editoria Multimídia, acerca dos conceitos de visualização de dados, imagem-síntese e inconsciente ótico, que guiam a análise.

1) Visualização de dados e adoção no Brasil

Em setembro de 2019, o *Estadão Infográficos* publicou a reportagem “Simulação mostra quais crianças são adotadas (e quais não são) no Brasil” premiada pelo *Kantar Information is Beautiful Awards* do mesmo ano, como melhor projeto na categoria “*Best Non-English*”. O conteúdo foi projetado com auxílio de algoritmo e visualização de dados, que permitiram percepção de tendências na adoção de crianças no Brasil. Porém, é relevante antes de abordar seu uso no direcionamento da pauta é importante entender como ela surgiu.

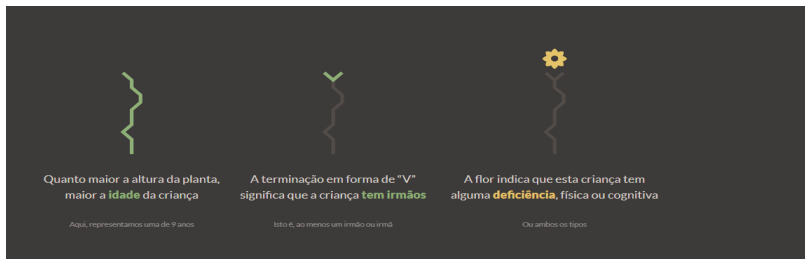
Como explicado, o *Estadão Infográficos* é uma editoria que cria matérias com diferentes perspectivas estéticas, narrativas e interações com o usuário. Para que isso seja possível, a redação permite que a Editoria Multimídia trabalhe tanto com um prazo mais longo de produção quanto com propostas mais criativas e descontraídas. Corrobora com essa proposta, a experiência de escolha da pauta dessa reportagem sobre adoção:

Basicamente, nós sorteamos verbos, substantivos e adjetivos e, em grupos, tínhamos 10 minutos para escrever algumas ideias que alinhavasse essas palavras. Então, nós votamos em nossos favoritos. Depois disso, tivemos uma semana para criar um protótipo bruto com as melhores ideias. Aqueles que parecessem promissores seriam levados adiante. A primeira palavra aleatória que meu grupo produziu foi sobre fertilidade. Nós nos lembramos de um conteúdo do WSJ [Wall Street Journal] que analisou as preferências das famílias no Brasil à procura de doadores de esperma e achamos que seria interessante comparar isso com as preferências em relação à adoção. Como descobrimos mais tarde, nós tínhamos dados suficientes para construir uma história atraente sobre o processo de adoção no Brasil. (Sueiro, 2019, s/p, tradução nossa²⁷).

Por meio da seleção aleatória, a equipe uniu a categoria de “fertilização” – observando os dados disponíveis sobre a temática e as diferentes abordagens possíveis – com o conceito de “adoção”, termos que inicialmente não se encontram interligados. Por isso, destaca-se o aspecto de miscelânea que permeia a editoria no surgimento de pautas, que supõe experimentação e combinações imprevisíveis.

Para ilustrar a fertilização na proposta da pauta, foram utilizadas plantas para identificar as crianças disponíveis para a adoção de acordo com algumas qualidades, como idade, raça, o fato de possuir irmãos e deficiência, física ou cognitiva (Figura 1). A escolha facilita o trabalho de visualização de dados e torna o componente interativo mais atraente pela beleza da representação.

Figura 1 – Legenda do simulador de adoção de crianças no Brasil.

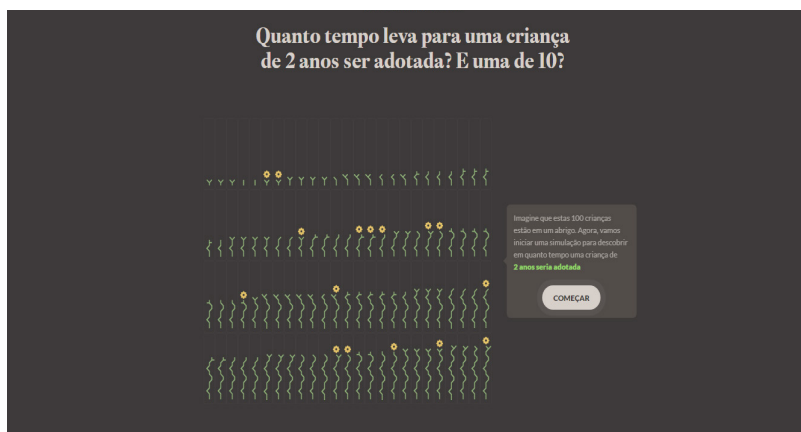


Fonte: *Estadão Infográficos* (2019)²⁸.

O simulador se assemelha à proposta do Sistema Nacional de Adoção (SNA) que seria implementado, em alguns estados brasileiros, a partir de outubro de 2019. Com dados do Instituto Brasileiro de Geografia e Estatística (IBGE) e do Cadastro Nacional de Adoção (CNA), o interativo organiza as plantas que representam as crianças disponíveis para adoção em fila, em ordem crescente de idade (até

18 anos, quando deixam os abrigos) com botões para a reação do usuário que acompanha as etapas do processo: a cada dia passado na simulação, um juiz observa preferências de quatro possíveis pais; no caso de adoção o campo retangular, no original da reportagem, brilha na cor verde e no caso de a criança atingir a maioridade, brilha na cor vermelha.

Figura 2 – Simulador de adoção do *Estadão Infográficos*.



Fonte: Estadão Infográficos (2019)²⁹.

Em entrevista ao *Kantar Information is Beautiful Awards*, Vinicius Sueiro revela as ferramentas utilizadas para a produção do simulador. “O algoritmo de simulação foi o *software* mais louco que eu já escrevi” (Sueiro, 2019, tradução nossa³⁰), enfatizou, ao esclarecer que o algoritmo usado para o interativo foi escrito com linguagem de programação JavaScript. Para melhorar a performance do algoritmo, também recorreu à manipulação do *Document Object Model* (DOM) – interface que representa como o *browser* lê documentos em HTML e XML, que identificam elementos numa página na internet. Assim, Sueiro (2019) conta que foi possível rodar a simulação milhares de vezes para alcançar os padrões apresentados na reportagem. Ainda acrescenta: “A saída [*output*] foi um arquivo JSON que foi analisado em Python, usando Pandas” (Sueiro, 2019, tradução nossa³¹). Dessa forma, foi utilizado o formato *Java Script Object Notation* (JSON) e a linguagem de programação *Python* por meio da biblioteca *Pandas*, com *softwares* em *Python* para análise de dados.

Está evidente a importância do conhecimento em programação para realizar a simulação interativa que fundamenta toda a reportagem multimídia. Esses são instrumentos específicos que dependem da formação do profissional na área e que, muitas vezes, o jornalista não possui. Porém, uma vez que a equipe tem acesso a esses tipos de ferramentas e a expertise necessária, a liberdade no procedimento de projeção da forma e do conteúdo apresentado é maior para criar materiais inovadores no mercado jornalístico.

Eu acredito que o design visual [...] é um fator importante quando transformando números em ideias a que as pessoas se sintam realmente conectadas. O modo como Giorgia Lupi abrange esse tópico em seu manifesto sobre humanismo de dados é incrível – e eu gostaria de citar algumas palavras que ressoam em nossas decisões por trás das histórias: “Quanto mais onipresentes se tornam os dados, mais precisamos experimentar como torná-los únicos, contextualizados e familiares”. (Sueiro, 2019, s/p, tradução nossa³²).

Destaca-se, além da importância dada ao aspecto visual da narrativa, que o veículo disponibiliza os códigos utilizados na produção das reportagens. Nesse caso, o código-fonte é aberto e disponível na plataforma *GitHub*, que abriga ferramentas e projetos que ensinam e compartilham conhecimento de trabalho com *software*. Denota-se, com isso, transparência em todas as produções da equipe.

2) Inconsciente ótico por meio da inteligência artificial

Para celebrar os 200 anos no Museu do Prado (Espanha), a editoria de cultura, em conjunto com a editoria multimídia do *Estadão*, propôs criar uma visualização de dados que contivesse todas as 6.367 pinturas do museu, na reportagem “Uma inteligência artificial vai te guiar pelo Museu do Prado; nossa repórter, pela cena cultural de Madri” (19/08/2019). Por meio da extração de descrições matemáticas de cada obra, os *features*, a equipe realizou a operação de comparar esses valores e calcular a distância entre as pinturas demarcadas por suas semelhanças. Para isso, foi utilizado o algoritmo *t-Distributed Stochastic Neighbor Embedding* (t-SNE) que permitiu representar cada obra em uma única superfície/imagem.

Figura 3 – Visualização das mais de 6 mil obras do Museu do Prado.



Fonte: Estadão infográficos (2019)³³.

Em reforço à transparência do método para o público, a reportagem disponibiliza o código-fonte utilizado para a matéria, bem como a imagem de representação das obras em única visualização, de forma que o leitor possa ampliá-la e explorá-la. Os agrupamentos das pinturas tornam-se improváveis, e mesmo impossibilitados, sem o auxílio do computador, pois o reconhecimento por meio do olhar humano levaria muito mais tempo, estando sujeito a obstáculos físicos, como o distanciamento das salas nas quais as obras estão expostas. Com a visualização, porém, a percepção se amplia ficando mais fácil entrever, por exemplo, que pintores criaram obras com características semelhantes entre si, mesmos dispersos no tempo, podendo ainda observar estilos de cada artista, suas características, propostas estéticas e críticas subjacentes à sua época.

A potencialização do olhar humano corresponde ao efeito que Benjamin (1987) esclarece como o inconsciente ótico, na condição de a imagem fotográfica estender o campo de visão para aspectos não observados diretamente pelo olhar. Com a ilustração construída pelos valores matemáticos que expressam as semelhanças estéticas de cada pintura, fica mais claro abstrair interrelações entre as obras, compreender tendências e mesmo

os limites da IA. Nesse caso, a máquina compreendeu que obras muito semelhantes, como “*La maja vestida*” – Francisco de Goya, 1800 – e “*La maja desnuda*” – Francisco de Goya, 1795 –, eram um mesmo elemento: em razão disso, na visualização essas obras estão sobrepostas uma a outra. É o que explicita a reportagem no trecho: “O computador é incapaz de compreender o contexto histórico ou as características subjetivas de uma obra [...]. Ou seja: cabe a nós, humanos, conferir ao roteiro maiores significados e relações” (Toni et al., 2019, s/p).

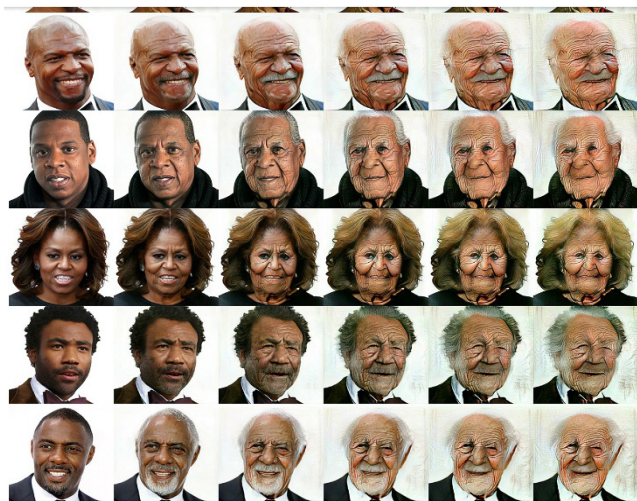
Apesar de algumas limitações da máquina, a figura no formato de nuvem demonstra a inclinação no trabalho de alguns artistas: Carlos de Haes (1829–1898) teve suas obras agrupadas, em razão da semelhança entre as pinturas; enquanto que as pinturas de Francisco de Goya (1746–1828) encontram-se distantes umas das outras, demonstrando variedade de estilo. É interessante notar que a própria escolha de utilizar a ilustração como construção da narrativa da reportagem, além de permitir novas formas de avaliar as pinturas do museu, propõe uma “maneira intrigante de conhecer o acervo da instituição” (Toni et al., 2019, s/p), trecho que explicita a busca pela inovação e curiosidade na construção do conteúdo.

3) Imagem-síntese: simulação e jornalismo

Imagens simuladas são amplamente utilizadas no jornalismo, principalmente quando faltam recursos imagéticos como fotos dos acontecimentos noticiados ou não são suficientes para demonstrar o fato. Um caso diferenciado foi o uso do aplicativo FaceApp no jornalismo, ferramenta que roda filtros de rejuvenescimento e envelhecimento de retratos, na tentativa de encontrar manipulações que reforcem uso de fontes de dados enviesados.

Na reportagem “Além de envelhecer, FaceApp embranquece rostos negros” (02/08/2019), o intuito é demonstrar como bases de dados tendenciosas perpetuam problemática atual, em relação ao padrão de beleza e ao preconceito com a cor negra de pele. Para isso, a equipe multimídia experimenta aplicar cinco vezes o filtro de envelhecimento na mesma foto. O resultado, além de demonstrar embranquecimento dos rostos, surge da exploração de imagens que não possuem referencial no mundo real, ou seja, configuram-se como linguagens abstratas.

Figura 4 – Filtro de envelhecimento do FaceApp aplicado cinco vezes em rostos negros.



Fonte: Estadão Infográficos (2019) ³⁴.

A perspectiva da Editoria Multimídia, que busca apresentar histórias interessantes com narrativas visuais inovadoras, é **resumida no** trecho: “Com exceção dos mais curiosos, ninguém roda o filtro cinco vezes na mesma imagem” (Romani et al., 2019, s/p), para observar o resultado da atuação do algoritmo. A editoria possui, assim, perfil de laboratório, similar à metodologia científica: tem-se uma hipótese, um método utilizado e esclarecido ao público, com experimentação e resultados observáveis por meio das visualizações. Isso corrobora com a concepção de Wolton (2011) sobre a importância dos jornalistas e cientistas como intermediários para a comunicação efetiva com a sociedade.

A imagem simulada está sujeita a distorções, em razão de ser produzida por máquina baseada na base de dados proposta para que realize a operação. Nesse caso, o erro é humano ao apresentar ao computador mais imagens de rostos envelhecidos da cor branca do que negra; a máquina fica treinada a entender que a cor também é critério no envelhecimento. Esse processo caracteriza a técnica de *machine learning*, que consiste em uma máquina realizar operação baseada em dados organizados para ensiná-la. Esse mesmo tipo de tecnologia é utilizado para o reconhecimento facial.

O fator humano se destaca na alimentação dos dados para que o aplicativo rode os filtros, e na reflexão sobre os resultados dessas imagens pela Editoria Multimídia. Não é o caso, portanto, de ser suficiente o uso da máquina para compreender o fenômeno; é preciso que o jornalista esclareça a problemática por meio das técnicas da profissão, checando as informações com fontes e busca de contextualização. Na reportagem, por exemplo, o leitor é informado quanto a erros anteriores envolvendo o FaceApp, que em 2017 com o filtro *hot* responsável por tornar fotos mais atraentes também resultava no embranquecimento dos rostos. Sem esclarecimento, o efeito do filtro de envelhecimento pode transparecer um problema pontual. Sem o uso das ilustrações na reportagem é difícil entender a tendência do embranquecimento, que é justamente o enfoque da matéria. A imagem produzida pela máquina, aqui, é o mote do conteúdo para demonstrar o uso de bases de dados tendenciosas. O procedimento é conhecido pelos usuários, visto que o aplicativo se tornou famoso na época e muitas pessoas o possuíam no celular; assim, a sequência de imagens era parte do repertório do leitor, habituado ao formato do aplicativo. Trata-se de esclarecer, com a imagem, mas também aproximar a experiência do público ao procedimento da reportagem.

7 Considerações finais

Como resultado, entende-se que o *Estadão Infográficos* reforça liberdade frente à agenda diária para escolha de pautas e produção de conteúdos, resultado de conexões inesperadas de conceitos, dados e informações a partir do uso de *softwares* e algoritmos. Isso é potencializado pela ubiquidade de informações disponíveis na rede. Nesse contexto, a curiosidade é aspecto estratégico para propor narrativas, não só na apuração ou definição da pauta, mas, sobretudo na forma de apresentação das histórias e como deixá-las mais interessantes para o público, com novas possibilidades estéticas.

A experimentação é essencial, munida da visualização de dados e das operações de algoritmos que atuam como auxiliares cognitivos para o jornalista abstrair percepções. Surgem novas pautas com a computação unida às técnicas jornalísticas e formas possíveis de apresentação – destaque no *Estadão Infográficos*, que

busca formatos inovadores para compor modelos. O pensamento computacional proporciona novas formas de olhar os acontecimentos sociais, auxiliando na abordagem diferenciada das notícias e reportagens: o jornalista treinado em técnicas tradicionais pode se acomodar à produção diária, que exige velocidade, fragmentação e concisão, o que prejudica a checagem e contextualização dos fatos. A autonomia da editoria de infografia e visualização de dados e a diversificação de técnicas, portanto, ajudam a driblar barreiras das redações tradicionais com criatividade, recorrendo à personalização de contextos complexos que aguçam a percepção do público, para exploração e descoberta.

As equipes interdisciplinares, consolidadas por jornalistas, designers e programadores, revigoram o atual modelo de jornalismo multimídia. Isso demonstra tendência da aplicação do jornalismo de dados para as produções, aspecto decorrente da adaptação às características dos suportes digitais, associados à hibridização de linguagem, multiplicidade de pontos de conexão e de acesso de conteúdos, em estruturas que têm característica de mobilidade e interação. A predisposição é observada nos nativos digitais e em seções de jornais tradicionais, como o *Estadão Infográficos*.

Por último, a pesquisa demonstra que conceitos das ciências sociais como imagem-síntese e o inconsciente ótico, podem ser atualizados para embasar investigações do jornalismo visual como elemento de arte funcional no âmbito da produção na internet. No caso estudado, a elaboração de visualizações de dados e imagens simuladas por meio de algoritmos favorecem processos cognitivos de assimilação de conteúdos e memória, incidindo também pela conformação de novas narrativas, percepções e experiências.

NOTAS

- 1 Aba “Infográficos” localizada entre as opções do menu lateral à esquerda no site do *Estadão*, ou *O Estado de S. Paulo*. Recuperado de www.estadao.com.br/infograficos.
- 2 Jornal impresso fundado em 1875 sob o título *A Província de São Paulo*, com a Proclamação da República brasileira em 1890 é nomeado *O Estado de S. Paulo*, hoje também conhecido como

- Estadão*. Em 2000, surge o domínio www.estadao.com.br. Segundo o Instituto Verificador de Comunicação (IVC), o *Estadão* está entre os veículos com maior circulação impressa e digital no Brasil. Soma mais de 31 milhões de visitas únicas ao site. Recuperado de <http://patrocinados.estadao.com.br/medialab/about-me/>
- 3 Recuperado de <https://brasil.estadao.com.br/noticias/geral,estado-ganha-8-premios-em-concurso-de-design-veja-as-reportagens,70003049883>
 - 4 Contato realizado em 19 de agosto de 2019 via e-mail, para os endereços eletrônicos falecomestado@estadao.com e portal@estadao.com.
 - 5 Entrevista com o Editor Assistente Multimídia do *Estadão*, Vinicius Sueiro, e o jornalista de dados do *Estadão*, Rodrigo Menegat, concedida à jornalista Daniela Borges de Oliveira em 3 de setembro de 2019 na sede d'*O Estado de S. Paulo* (Av. Eng. Caetano Álvares, 55, Limão/SP). O diálogo é resultado do projeto monográfico “Narrativa Digital: Análise Crítica das Estratégias do Estadão Infográficos”. Orientado pelo Prof. Dr. Belarmino Cesar Guimarães da Costa, apresentado ao curso de Jornalismo da Universidade Metodista de Piracicaba (UNIMEP) para obtenção de grau de bacharel.
 - 6 Texto original: “an algorithm can be defined as a series of steps undertaken in order to solve a particular problem or accomplish a defined outcome” (Diakopoulos, 2015, p.400).
 - 7 Texto original: “Algorithm decisions can be based on heuristics and rules, or calculations over massive amounts of data. Rules may be articulated directly by programmers, or be dynamic and flexible based on machine learning of data. Sometimes a human operator maintains agency and makes the final decision in a process, but even in this case the algorithm biases the operator’s attention toward a subset of information” (Diakopoulos, 2015, p.400).
 - 8 Texto original: “a self-contained step-by-step set of operations to be performed, such as calculation, data processing, and automated reasoning – a set of rules that precisely defines a sequence of precise instructions that can be understood by a computer” (Lindén, 2017, p.62).

- 9 Texto original: “in which algorithms are used to collect, collate and organise the data upon which decisions are typically made, and to assist in how that data is processed and communicated through the relevant governance system” (Danaher, 2016, p.04).
- 10 Texto original: “a journalism algorithm needs to reflect a behaviour that is not always rule-based but reflective and flexible in the sense that the application of explicit and internalised rules to complex situation is ambiguous” (Lindén, 2017, p.64).
- 11 Texto original: “Traditionally, journalists have mainly relied on qualitative methods for their research, but the supply of easy-to-use, cheap, or free software for processing large data sets has opened new opportunities. The personal memory capacity and analytical processing capabilities have already been augmented with the aid of computers” (Lindén, 2017, p.64).
- 12 Recuperado de www.estadao.com.br/infograficos/politica,o-que-revela-uma-analise-das-emocoes-dos-candidatos-durante-o-debate,923037
- 13 *Folha, Globo e Estadão* são os veículos brasileiros de maior circulação impressa e digital nacional. O impresso *Folha de S.Paulo* tem origem em 1960 e tem como site: www.folha.uol.com.br
- 14 Texto original: “data journalism helps engage audiences, improve quality and raise productivity” (ICFJ, 2019, p.10).
- 15 *Nexo*: Jornal nativo digital lançado em 2015 com a proposta de produzir conteúdo em profundidade, baseados em dados e estatísticas. Recuperado de www.nexojornal.com.br
- 16 *Gênero e Número*: portal jornalístico que disponibiliza dados sobre equidade de gênero www.generonumero.media
- 17 A seção “Infográficos” da *Folha* agrega produções multimídia no site do veículo: www1.folha.uol.com.br/infograficos/
- 18 Conforme abordado nos procedimentos metodológicos, a autora realizou entrevista semiestruturada ao então Editor Assistente Multimídia Vinicius Souza Sueiro, às 15h do dia 3 de setembro de 2019, na sede d’*O Estado de S. Paulo*. Na oportunidade, também foi possível entrevistar o jornalista de dados, Rodrigo Menegat.

- 19 Texto original: “Visuals make your argument more accessible and less complex than a thousand-word essay”. Recuperado de <https://pudding.cool/about/#pitch>
- 20 Texto original: “in a broad sense, data journalism is telling stories with numbers, or finding stories in them” (Howard, 2014, p.5).
- 21 Texto original: “In developing data journalism products – often data visualization or Web – their usefulness to the audience is a prime consideration” (Coddington, 2015, p.343).
- 22 Texto original: “A data visualization is a display of data designed to enable analysis, exploration, and discovery. Data visualization aren’t intended mainly to convey messages that are predefined by their designers. Instead, they are often conceived as tools that let people extract their own conclusions from the data” (Cairo, 2016, s/p).
- 23 Conceituação expressa no módulo 4 do curso “Introdução ao jornalismo de dados: como entrevistar dados para reportagens investigativas”, oferecido na plataforma *JournalismCourses.org* pelo *Knight Center for Journalism in the Americas* (5 de agosto a 8 de setembro de 2019). Conferir: <https://journalismcourses.org/pt-br/course/dados/>. A concepção foi reforçada durante entrevista em 03 de setembro de 2019.
- 24 Texto original: “Graphics that encode information function as cognitive aids” (Cairo, 2016, s/p).
- 25 Texto original: “the goal of computational journalism is to provide them [the public] with tools they need to perform their own filtering and abstraction with it” (Coddington, 2015, p.343).
- 26 Texto original: “the integration of logical, algorithmic, scientific, and innovative dimensions of human cognition, pointing towards openness to new ideas, revolutionising all kinds of occupations, including journalism” (Lindén, 2017, p.71).
- 27 Texto original: “Basically, we sorted verbs, nouns and adjectives and, in groups, we had 10 minutes to write down a few ideas that stitched those words together. We then voted on our favorites. After that, we had one week to create a raw prototype with the best ideas. The ones that seemed promising would be taken forward. The first random words my group produced had to do

with fertility. We remembered a piece by the WSJ that analyzed the preferences of families in Brazil regarding sperm donors and we thought it would be interesting to compare it to the preferences regarding adoptions. As it later turned out, we had enough data to build a compelling story on the adoption process in Brazil” (Sueiro, 2019, s/p).

28 Recuperado de <https://arte.estadao.com.br/brasil/adocao/criancas/>

29 Idem.

30 Texto original: “The simulation algorithm was the craziest piece of software I have ever wrote” (Sueiro, 2019, s/p).

31 Texto original: “The output was a JSON file that was analysed in *Python*, using *Pandas*” (Sueiro, 2019, s/p).

32 Texto original: “I believe that visual design [...] is an important factor when transforming numbers into ideas that people really feel connected to. The way in which Giorgia Lupi has covered this topic in her manifesto about data humanism is amazing – and I would like to quote a few words that really resonate with our decisions behind this story: “The more ubiquitous data becomes, the more we need to experiment with how to make it unique, contextual, intimate.” – Giorgia Lupi on data humanism” (Sueiro, 2019, s/p).

33 Recuperado de www.estadao.com.br/infograficos/viagem,uma-inteligencia-artificial-vai-te-guiar-pelo-museu-do-prado-nossa-reporter-pela-cena-cultural-de-madri,1025741

34 Recuperado de www.estadao.com.br/infograficos/link,alem-de-envelhecer-faceapp-embranquece-rostos-negros,1018384

REFERÊNCIAS

Benjamin, W. (1987). A obra de arte na era de sua reprodutibilidade técnica. In W. Benjamin (Org.), *Magia e Técnica, Arte e política: Ensaio sobre literatura e história da cultura* (3ª ed.) (pp.165–196). São Paulo: Editora Brasileira.

Cairo, A. (2013). *The Functional Art: An introduction to information graphics and visualization*. Berkeley: New Riders.

Cairo, A. (2016). *The Truthful Art: Data, charts, and maps for communication*. Berkeley: New Riders.

Coddington, M. (2015). Clarifying Journalism's Quantitative Turn: a typology for evaluation data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism*, 3 (3), 331-348. DOI: 10.1080/21670811.2014.976400

Costa, B. C. G. da. (2002). *Estética da violência: Jornalismo e produção de sentidos*. Campinas: Autores Associados, Piracicaba: Editora Unimep.

Dalmonte, E. F. (2009). *Pensar o discurso no webjornalismo: temporalidade, paratexto e comunidades de experiência*. Salvador: EDUFBA.

Danaher, J. (2016). The Threat of Algocracy: Reality, Resistance and Accommodation. *Philosophy & Technology*, 29, 245-268. DOI: 10.1007/s13347-015-0211-1

Diakopoulos, N. (2015). Algorithmic Accountability: Journalistic investigation of computational power structures. *Digital Journalism*, 3 (3), 398-415. DOI: 10.1080/21670811.2014.976411

Dourish, P. (2016). Algorithms and their others: Algorithmic culture in context. *Big Data & Society*, 3 (2), 1- 11. DOI: 10.1177/2053951716665128

Flusser, V. (2008). *O universo das imagens técnicas*. São Paulo: Annablume.

Gordon, R. (2013, 28 de jun.). Want to build a data journalism team? You'll need these three people. *Knight Lab*. Recuperado de knightlab.northwestern.edu/2013/06/28/want-to-build-a-data-journalism-team-youll-need-these-three-people/

Howard, A. B. (2014). *The Art and Science of Data-Driven Journalism*. Tow Center for Digital Journalism: Columbia Journalism School. DOI: 10.7916/D8Q531V1

International Center for Journalists (ICFJ). (2019). *The State of Technology Global Newsrooms*. Recuperado de www.icfj.org/sites/default/files/2019-10/2019%20Final%20Report.pdf

Kaufman, D. (2019). A Inteligência Artificial mediando a comunicação: Impactos da automação. In M. Barbosa (Org.), *Pós-verdade e fake news: Reflexões sobre a guerra de narrativas* (pp.49-58). Rio de Janeiro: Cobogó.

Lemos, A. (2001). Ciber-flânerie. In D. F. Silva, & S. Fragoso (Orgs.), *Comunicação na Cibercultura* (pp.45-60). Leopoldo: Editora Unisinos.

Lindén, C. (2017). Algorithms for journalism: The future of news work. *The Journal of Media Innovations 4.1*, 4(1), 60–76. DOI: 10.5617/jmi.v4i1.2420

Mancini, L., & Vasconcellos, F. (2016). Jornalismo de dados: Conceito e categorias. *Revista Fronteiras*, 18 (1), 69–82. DOI: 10.4013/fem.2016.181.07

Martinez, A. G. (2007) A construção da notícia em tempo real. In P. Ferrari (Org.), *Hipertexto hiperídia: as novas ferramentas da comunicação digital* (pp.13–27). São Paulo: Contexto.

Quéau, P. (1993) O tempo do virtual. In A. Parente (Org.), *Imagem-máquina: a era das tecnologias do virtual* (pp.91–99). Rio de Janeiro: Editora 34.

Romani, B., Ponceano, B., & Sueiro, V. (2019, 19 de mar.). Além de envelhecer, FaceApp embranquece rostos negros. *Estadão*. Recuperado de www.estadao.com.br/infograficos/link,alem-de-envelhecer-faceapp-embranquece-rostos-negros,1018384

Santaella, L. (2004). *Navegar no ciberespaço: O perfil cognitivo do leitor imersivo*. São Paulo: Paulus.

Sueiro, V. (2019). Interview with 2019's Best Non-English-Language Winner. *Kantar Information Is Beautiful Awards*. Recuperado de www.informationisbeautifulawards.com/news/540-interview-with-2019-s-best-non-english-language-winner

Teixeira, T. (2010). *Infografia e Jornalismo: Conceitos, análises e perspectivas*. Salvador: EDUFBA.

Toni, B., Menegat, R., Marin, C., & Ponceano, B. (2019, 19 de ago.). Uma inteligência artificial vai te guiar pelo Museu do Prado. *Estadão*. Recuperado de www.estadao.com.br/infograficos/viagem,uma-inteligencia-artificial-vai-te-guiar-pelo-museu-do-prado-nossa-reporter-pela-cena-cultural-de-madri,1025741

Wolton, D. (2011). *Informar não é comunicar*. Porto Alegre: Sulina.

Zuin, A. A. S., & Zuin, V. G. (2015). A indústria cultural, a semiformação e a revolução microeletrônica. In A. F. Maia, A. A. S. Zuin, & L. A. C. N Lastória (Orgs.), *Teoria Crítica da cultura digital: Aspectos educacionais e psicológicos* (pp.202–213). São Paulo: Nankin.

DANIELA BORGES DE OLIVEIRA. Jornalista pela Universidade Metodista de Piracicaba (UNIMEP), foi bolsista PIBIC/CNPq da Iniciação Científica “*Folha de S.Paulo: Cobertura jornalística, violência estética e sociedade midiática*” (ago.2019 a jul.2020), estudante do GP “Comunicação na Era Digital: Cultura, Estética e Linguagem” (UNIMEP). Áreas de pesquisa: Cultura Digital, Indústria Cultural e Teorias do Jornalismo. E-mail: dani7b.o@gmail.com

BELARMINO CESAR GUIMARÃES DA COSTA. Doutor e Mestre em Educação (Unicamp/UFSCar) e Jornalista (UNIMEP). Docente do Programa de Pós-Graduação em Educação da UNIMEP. Desenvolve pesquisas sobre mídia, linguagem e educação; indústria cultural e ambiente digital; comunicação, tecnologia e estética. Membro do GP “Teoria Crítica e Educação” (UFScar). Ex-coordenador do GP “Comunicação e Educação” da Intercom. E-mail: belarminocesar@uol.com.br

PARECERES

Um dos pareceres utilizados na avaliação deste artigo pode ser acessado em: <https://osf.io/8q4z7/>

Seguindo a política de ciência aberta da BJR, o avaliador autorizou a publicação do parecer e a divulgação do seu nome.