



Núcleo de Estudos em Jornalismo de Dados e Computacional - DataJor
Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IBDEP)

JORNALISMO DE DADOS

das Ciências Sociais à Inteligência Artificial

Taís Seibt (org.)

JORNALISMO DE DADOS **Das Ciências Sociais à Inteligência Artificial**

Organização e edição

Taís Seibt

Redação

Adriano Belisário
Bárbara Libório
Beatriz Farrugia
Francisco Amorim
Jamile Santana
Judite Cypreste
Liliane Ito
Lívia Vieira
Lucas Thaynan
Marcela Canavarro
Marcelo Fontoura
Marcelo Soares
Marília Gehrke
Marta Alencar
Mayane Batista
Polinho Mota
Reinaldo Chaves
Samuel Pantoja Lima
Taís Seibt
Thays Lavor

Revisão

Bárbara Libório
Bruno Borges
Fábio Bispo
Liliane Ito
Lívia Vieira
Marcela Canavarro
Marcelo Fontoura
Nicolás Llano

Capa e ilustrações

Renan Capitão

Diagramação

Marcelo Garcia

Código de catalogação na publicação – CIP

J82 Jornalismo de dados: das ciências sociais à inteligência artificial / Taís Seibt (organizadora); autores Adriano Belisário... [et al]. — Brasília: Instituto Brasileiro Ensino, Desenvolvimento e Pesquisa, 2025.

199 f. : il. color.

Inclui bibliografia.

Vários autores.

ISBN 978-65-87546-44-5

1. Jornalismo digital - Brasil. 2. Jornalismo de investigação - Brasil.
3. Visualização da informação. I. Título II. Seibt, Taís.

CDD 070.40285

Elaborada por Natália Bianca Mascarenhas Puricelli – CRB 1/3439

Agradecimentos

Às autoras e autores desta publicação, que encontraram tempo para colaborar, de forma voluntária, na redação dos capítulos e também na revisão de textos de seus pares. Obrigada por abraçarem a proposta de compartilhar seus conhecimentos como professores e pesquisadores, e suas experiências como jornalistas de dados para inspirar futuros colegas.

Aos integrantes do Núcleo de Estudos em Jornalismo de Dados e Computacional - DataJor (IDP/CNPq) que participaram do processo de revisão dos capítulos, sugerindo ajustes para a qualificação do trabalho final.

Ao Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP) pelo apoio institucional no registro bibliográfico da publicação e subsídio para a editoração gráfica e divulgação do e-book, por meio do Programa de Apoio ao Desenvolvimento de Grupos de Pesquisa 2025/1.

Sumário

Apresentação

Jornalismo de Dados à brasileira <i>Taís Seibt</i>	07
---	----

Parte I: Conceitos e princípios

Jornalismo de Precisão: por uma prática jornalística inspirada nos métodos das ciências sociais <i>Marcelo Soares</i>	13
Feminismo de Dados: contribuições para pensar o jornalismo pela lentes da interseccionalidade <i>Marília Gehrke</i>	25
Precusores do Jornalismo de Dados no Brasil: como os jornalistas brasileiros começaram a entrevistar planilhas <i>Lívia Vieira</i>	33
Dados também são de humanas! Por que perder o medo dos números qualifica o jornalismo <i>Liliane Ito</i>	41

Parte II: Primeiros passos

A força da pergunta: o papel da hipótese em reportagens guiadas por dados <i>Bárbara Libório</i>	49
Lei de Acesso à Informação e dados públicos como matéria-prima <i>Jamile Santana</i>	56
Geração cidadã de dados para criar bases originais para reportagens de impacto social <i>Polinho Mota</i>	64
OSINT no jornalismo: indo mais fundo com fontes abertas <i>Adriano Belisário</i>	71
A ética profissional no Jornalismo de Dados <i>Samuel Pantoja Lima</i>	80

Parte III: Dados têm contexto

Operações básicas com dados e seus sentidos para a narrativa jornalística	
<i>Marcelo Fontoura</i>	89
Da estatística descritiva ao jornalismo de soluções	
<i>Francisco Amorim</i>	96
Dados geográficos muito além dos “mapinhas”	
<i>Thays Lavor</i>	108
Dados em rede: social listening e métodos digitais como recursos para o jornalismo	
<i>Marcela Canavarro</i>	118
Guiado por dados, mas ainda é jornalismo	
<i>Beatriz Farrugia</i>	126

Parte IV: Um passo a mais

Visualização de dados para contar histórias mais envolventes	
<i>Lucas Thaynan</i>	136
Programação ajuda a navegar por dados além da superfície	
<i>Judite Cypreste</i>	152
Fact-checking vai aos dados para comprovar fatos	
<i>Marta Alencar</i>	160
Jornalismo de Dados sob o signo da Inteligência Artificial	
<i>Mayane Batista</i>	171
Perspectivas profissionais: desafios e oportunidades do mercado de trabalho para jornalistas de dados	
<i>Reinaldo Chaves</i>	179

Apresentação

Jornalismo de Dados à brasileira

Por Taís Seibt

O ensino de Jornalismo de Dados cresceu (lentamente) nos cursos de graduação em Jornalismo do Brasil na última década, especialmente pela necessidade de adaptar os currículos às novas diretrizes nacionais, que vigoram desde 2013. Contudo, ainda há lacunas importantes.

Uma delas é a disponibilidade de professores que conciliem conhecimento técnico e prático em Jornalismo de Dados com a titulação acadêmica adequada para dar aulas no ensino superior. Isso faz com que muitos cursos tratem o Jornalismo de Dados apenas como um tópico complementar em alguma disciplina mais genérica, muitas vezes contando com a boa vontade de profissionais convidados para uma palestra, ou então na forma de um curso de extensão ministrado por um especialista externo, atingindo apenas estudantes muito interessados no tema — ou nas horas complementares necessárias para a conclusão do curso.

Alguns professores, por interesse e iniciativa própria, buscam cursos de qualificação oferecidos por entidades como a Escola de Dados e a Associação Brasileira de Jornalismo Investigativo (Abraji) ou mesmo outras instituições de ensino para encontrar referências úteis para suas aulas, mas o incentivo à formação de professores universitários para o ensino de Jornalismo de Dados é inversamente proporcional à exigência dessas habilidades para o profissional de jornalismo formado nos bancos universitários e à velocidade com que a área se desenvolve, diante de inovações tecnológicas constantes que exigem cada vez mais refino técnico — e ético.

Outra lacuna é a ausência de bibliografia formal atualizada e disponível em português para compor as referências básicas de uma disciplina de Jornalismo de Dados compatível com o nível de gradua-

ção. As diretrizes curriculares de muitas universidades determinam que bibliografia básica indicada nas ementas esteja disponível inclusive em livros físicos na biblioteca universitária. Os textos online podem compor a bibliografia complementar, desde que disponíveis em repositórios digitais com a curadoria da instituição. Essa norma, não raro, limita a utilização de ótimas publicações disponíveis gratuitamente na internet, como o Fluxo de Trabalho com Dados, da Escola de Dados, e o Manual de Jornalismo de Dados, traduzido pela Abraji, que são a base didática da maioria dos cursos de formação frequentados por excelentes jornalistas de dados do Brasil.

Sem contar a dificuldade de encontrar obras de referência traduzidas para o português, como é o caso de livros do pioneiro Phillip Meyer, pesquisador estadunidense que cunhou o “jornalismo de precisão”, ou mesmo Paul Bradshaw, professor britânico que é uma referência contemporânea em Jornalismo de Dados com muitos textos online traduzidos para o português por alunos brasileiros que ele mesmo orientou, ou ainda Catherine D'Ignazio, que sustenta o “feminismo de dados” como base para compreender que os dados não são neutros, e sim produtos de relações sociais desiguais. Na pós-graduação, espera-se que estudantes leiam em outros idiomas, mas na graduação, exigir que os alunos leiam textos dessa complexidade em inglês é incompatível com a realidade.

Por fim, também é preciso considerar que ótimas referências internacionais traduzidas para o idioma nacional continuam sendo referências internacionais, ou seja, carregam exemplos e contextos externos, distantes ao cenário brasileiro da profissão. Fica em aberto a adaptação necessária para a realidade local, que felizmente já conta com inúmeros trabalhos de referência em reportagens e projetos especiais com recursos avançados de visualização de dados, goza de alguma maturidade em métodos e práticas para o acesso a dados públicos com uso da Lei de Acesso à Informação — em vigor no Brasil desde 2012 — e oferece boas práticas para criação de bases de dados próprias com finalidade jornalística.

Reunir essa gama de conhecimentos em uma só publicação seria pretensão demais para este e-book, porém não nos impede de oferecer um panorama do Jornalismo de Dados sob a perspectiva brasileira como fio condutor de uma discussão inicial para neófitos no

tema. Este é o propósito desta publicação. Dividida em quatro partes, esta obra é uma produção coletiva realizada pelo Núcleo de Estudos em Jornalismo de Dados e Computacional — DataJor vinculado ao MBA em Jornalismo de Dados e ao Mestrado em Comunicação Digital do Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP) com a colaboração de especialistas e pesquisadores convidados.

A escolha dos conteúdos e a indicação dos autores considerou a expertise de cada um, buscando a melhor composição entre academia e mercado, com diversidade regional e de gênero, para a produção de um conjunto de textos qualificados para utilização como material de formação acadêmica sem necessariamente carregar um linguajar teórico rebuscado e aprofundamento técnico exagerado.

Como obra que busca oferecer uma visão panorâmica da área, partimos de conceitos e princípios, dispostos na Parte I, onde apresentamos não só o conceito de “jornalismo de precisão”, mas também seu criador, Phillip Meyer, que inaugurou uma prática jornalística inspirada nos métodos das ciências sociais nos Estados Unidos e contagiou jornalistas brasileiros, como Marcelo Soares. No capítulo, ele mesmo conta como a descoberta de um livro de Meyer na biblioteca da universidade mudou para sempre sua prática jornalística. Na mesma esteira, Marília Gehrke, uma das fundadoras do DataJor, detalha como o conceito de “feminismo de dados”, trabalhado por Catherine D'Ignazio não só na teoria, mas de forma aplicada, orienta seu trabalho atual de pesquisa, à luz de teorias feministas para destacar que os dados não são neutros. São textos seminais estrangeiros na perspectiva de brasileiros que desbravaram a área, como outros precursores que começaram a entrevistar planilhas, cujo legado Livia Vieira homenageia registrando trabalhos de referência para novos jornalistas de dados. Fechamos a primeira parte com Liliane Ito reiterando que dados também são de humanas e que perder o medo das tabelas qualifica o jornalismo em um cenário de desinformação.

Na Parte II, trabalhamos com os primeiros passos, a começar pelas perguntas, com Bárbara Libório destacando a força de hipóteses bem elaboradas na produção de reportagens guiadas por dados. Jamile Santana aponta os caminhos da Lei de Acesso à Informação, a LAI, como forma de desaprisionar dados públicos que servem como matéria-prima para investigações inéditas. Quando os dados necessários

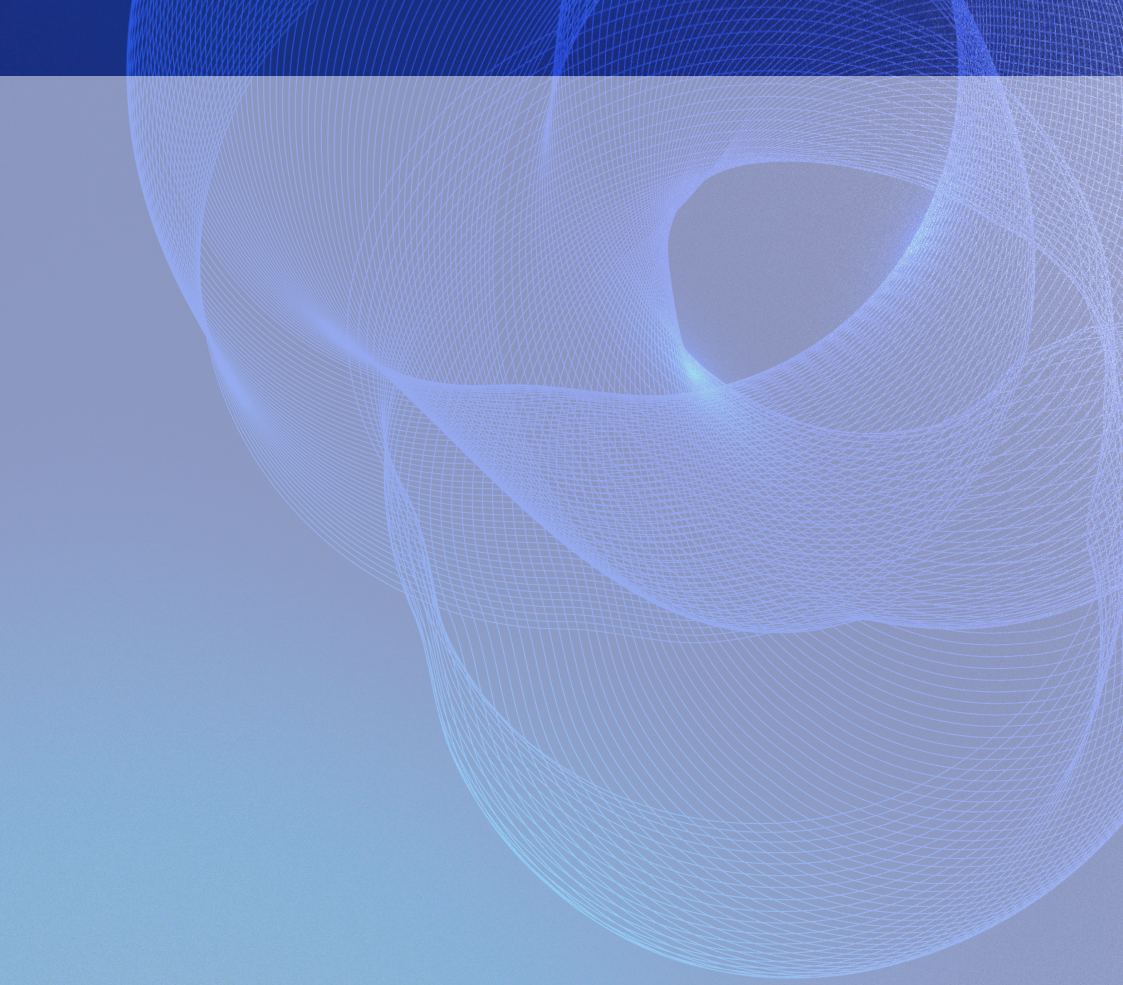
não estão públicos, nem sequer são produzidos pelos entes responsáveis, a geração cidadã de dados é uma alternativa para o jornalismo de alto impacto social, como a experiência de Paulo Motta nos ensina. Também fontes abertas, que podem ser recuperadas com técnicas de investigação na web, permitem ir mais a fundo para explorar histórias ocultas, como explica Adriano Belisário. E, como Jornalismo de Dados ainda é jornalismo, não podemos deixar de lado a ética profissional, problematizada por Samuel Pantoja Lima.

A Parte III desta publicação reforça que os dados têm contexto: de nada adianta dominar todas as técnicas mais acuradas se elas não estiverem a serviço de uma narrativa jornalística qualificada. Marcelo Fontoura explica as operações básicas com dados e o que elas significam em cada história, enquanto Francisco Amorim provoca jornalistas a não se limitarem ao descritivo, mas também explorar oportunidades que os dados oferecem para debater soluções de problemas sociais reveladas pelos números. Thays Lavor apresenta as possibilidades de dados geográficos que ultrapassam a mera produção de mapas para a visualização de informações, assim como Marcela Canavarro explora as informações contidas em redes que conectam pessoas e carregam informações nas plataformas digitais. Beatriz Farrugia fecha a seção com reflexões sobre o diferencial do tratamento jornalístico para todas essas possibilidades de apuração baseadas em dados.

Por fim, a Parte IV convida a pensar nos próximos passos. A visualização de dados é uma especialidade em si no Jornalismo de Dados, com uma gramática própria que mescla competências de comunicação visual com a interpretação dos dados, como mostra Lucas Thaynan. Já aprender programação abre um leque ainda maior de possibilidades de contar histórias e desvendar mistérios que as técnicas convencionais de jornalismo jamais descobririam, como compartilha Judite Cypreste. Habilidades de busca e análise de dados são pré-requisitos também para fact-checkers, jornalistas especializados que se dedicam a verificar informações que circulam nas plataformas digitais, conforme Marta Alencar. Os desafios da Inteligência Artificial, nova fronteira do Jornalismo de Dados, embora permeiem diversos outros capítulos, são discutidos mais a fundo por Mayane Batista. Para concluir, Reinaldo Chaves promove uma reflexão sobre as perspectivas profissionais para jornalistas de dados.



Como se vê, trata-se de uma coletânea com abordagens complementares, ainda que esta não seja uma obra completa. O conhecimento em Jornalismo de Dados é múltiplo, e certamente há mais para se pensar e debater na formação de um profissional. Ainda assim, se os futuros jornalistas de dados brasileiros puderem começar por aqui, sem dúvida encontrarão fontes qualificadas para se inspirar na profissão. E o melhor: terão a oportunidade de encontrar essas fontes no próximo congresso de jornalismo ou em algum fórum de discussão online, poderão tirar dúvidas e trocar experiências diretamente, sem barreiras geográficas ou linguísticas. Esta é, enfim, uma produção de professores de jornalismo e jornalistas de dados do Brasil para professores de jornalismo e jornalistas de dados do Brasil.

An abstract graphic in the top half of the page, featuring a complex, swirling wireframe structure in shades of blue. The structure resembles a stylized, three-dimensional knot or a series of overlapping loops, with a central void. The lines are thin and densely packed, creating a mesh-like appearance. The background is a solid light blue.

Parte I

Conceitos e princípios

Jornalismo de Precisão: por uma prática jornalística inspirada nos métodos das ciências sociais

Por Marcelo Soares

Se você for olhar bem a história da maior parte do que se considera grandes inovações do jornalismo neste século, todas elas foram iniciadas matutamente por Philip Meyer, um sujeito genial, conhecido por sua gravata-borboleta, que morreu, em 2023, com 93 anos bem vividos.

Jornalismo de Dados? Em 1967 ele já montava surveys com a comunidade e perfurava cartões. Ao menos até o final da década de 1990, já como professor, ele deu concorridas oficinas de estatística e análise de dados para jornalistas em que usava piadas de tiozão e saquinhos de M&M para explicar como funcionava a probabilidade.

Produto? No final dos anos 70, ele estava a cargo dos estudos de viabilidade do Viewtron, um pioneiro jornal digital, acessível pela TV, criado pela rede de jornais Knight-Ridder — até hoje, os já sexagenários colegas de escola de suas filhas se referem a ele como “o cara que inventou a internet”, o que obviamente é um simplismo.

Audiência? Muito antes do Google Analytics, Meyer escreveu livros ensinando editores a interpretar dados de pesquisas com o público para entender quem eram os leitores e o que queriam. Nota: Quando fui o primeiro editor de audiência da imprensa brasileira, na Folha de S. Paulo, o “Newspaper Survival Book” ficava na minha gaveta, sempre à mão, e sempre havia algum novo truque de análise a aprender com o velho mestre.

O grande inovador do jornalismo

“Phil é um dos grandes inovadores do jornalismo, e ele provavelmente ficaria vermelho se disséssemos isso na sua cara”, disse Jim Steele, veterano repórter com cinco décadas de experiência em expor

os mecanismos legais pelos quais os Estados Unidos aprofundaram a desigualdade. E tudo começou numa madrugada fria de 1973, quando Steele e seu colega Don Barlett amanheceram perfurando cartões com Meyer para uma das primeiras reportagens de dados já publicadas.

Na noite de 4 de novembro de 2022, quando apresentei o depoimento de Steele numa sessão em homenagem aos 50 anos do "Precision Journalism" no Coda.BR, encontro anual de Jornalismo de Dados organizado pela Escola de Dados, Meyer estava assistindo ao vivo, na casa de uma de suas filhas, na Carolina do Norte. Aos 92 anos, enfrentando o mal de Parkinson em estágio avançado, ele ficou emocionado ao rever seus amigos contando suas histórias. A filha, com quem vivia, nunca me disse se ele ficou vermelho, mas Jim e uma série de outros amigos falaram a verdade na cara dele.

Exatamente meio século antes, Meyer tinha em mãos as provas finais da primeira edição do livro "Precision Journalism", que o consagrou e inspirou gerações de jornalistas que gradualmente criaram o que hoje se chama Jornalismo de Dados. O livro, que teve quatro edições entre maio de 1973 e fevereiro de 2002, foi considerado uma das 35 obras mais significativas sobre jornalismo publicadas no século 20 pela "Journalism & Mass Communication Quarterly". Mas, muito mais importante, inspirou gerações de jornalistas a fazer suas próprias análises de dados para publicação, ganhando graus inéditos de independência em relação às fontes. Por meio da análise de dados, é possível questionar o poder com base nas informações que ele próprio produz.

Ele serve como uma espécie de manual de análise de dados para jornalistas. Em 2025, os trechos de código em Basic parecem divertidos. Em 1991, podiam ser assustadores. Mas Meyer enxergou, lá no tempo do cartão perfurado¹, que, num mundo onde o volume de informação aumenta rapidamente, é urgente que os jornalistas tenham autonomia para coletar dados, processá-los e explicá-los ao leitor. Precisavam ter com os números a mesma liberdade que tinham pra

¹ Antes dos primeiros computadores pessoais, no começo dos anos 70, a interação entre usuário e computador não era feita via teclado. Era preciso perfurar cartões, semelhantes aos de provas de múltipla escolha em concursos, para processar bases de dados, onde cada furo é uma resposta a uma pergunta e cada cartão representa uma linha do que mais tarde seria uma planilha. Na segunda edição de Precision Journalism, publicada em 1979, Meyer incluiu reproduções dos cartões perfurados utilizados em reportagens de que participou.

dizer “corta essa, tá achando que eu sou trouxa?” a fontes humanas que encontrassem em bares de péssima reputação. Spoiler: continuam sem ter essa liberdade, e atualmente sequer as fontes humanas que tentam enrolar os repórteres andam levando muito desaforo pra casa.

Meyer desenvolve, no livro, uma proposta metodológica que deu origem a alguns dos trabalhos de reportagem mais interessantes das décadas de 1970 e 1980, feitas por discípulos seus. Mesmo eleito um dos 20 trabalhos mais significativos sobre jornalismo no século 20, demorou a receber o devido reconhecimento. Meyer guardou para sempre, com humor, cartas raivosas enviadas à editora por professores de jornalismo que acharam o trabalho inútil. Era um livro muito mais técnico do que qualquer outro livro sobre jornalismo e, ao mesmo tempo, muito mais jornalístico do que qualquer outro sobre metodologia de pesquisa nas ciências sociais.

Experiência de quem teve a vida transformada pelo livro de Meyer nos anos 90: até mais ou menos 2010, falar de “jornalismo de precisão” ou “reportagem com o auxílio do computador” no Brasil era recebido mais ou menos como falar de carros voadores. Isso mudou muito desde então, com o nome “Jornalismo de Dados”. Mas a roda começou a girar muito antes.

E sequer esse problema é novo. Brant Houston, diretor da IRE quando criamos a Associação Brasileira de Jornalismo Investigativo (Abraji), lembrou que, nos anos 90, Meyer desconfiava um pouco dos que faziam análises apressadas de dados usando seus PCs, chamando isso de “Reportagem com o Auxílio do Computador”. Por um tempo, chegou a chamar os novos colegas de “caubóis dos dados”. Ao final, fizeram as pazes. E, então, o que temos são três ou quatro gerações de jornalistas que no fundo estão buscando contar verdades da maneira mais precisa possível.

Com a coleta de dados públicos e sua análise em computadores — nas duas primeiras edições, Meyer nos apresentava aos cartões perfurados —, era possível questionar o poder com base na informação produzida na redação, a partir de um olhar jornalístico informado pelas metodologias de pesquisa das ciências sociais. Na última década e meia, a visualização de dados ganhou papel cada vez mais central no que se entende como Jornalismo de Dados, mas a lógica está toda lá.

“Precision Journalism” já foi publicado em espanhol, italiano,

francês, turco, albanês, búlgaro e chinês — português ainda não. Antes de ser publicado pela Indiana University Press, o manuscrito foi rejeitado por cinco editoras. “O que as pessoas às vezes esquecem sobre o Phil é que ele sempre foi, no fundo, um repórter”, disse Sarah Cohen, professora da Arizona State University e presidente do comitê do Philip Meyer Award, oferecido desde 2005 a reportagens inovadoras na metodologia de apuração. Nos anos 90, ela teve a honra de lecionar ao lado de Meyer em “bootcamps” de estatística para jornalistas oferecidos anualmente na Carolina do Norte, onde ele lecionava.

Desde 2006, Meyer se tornou conhecido ao redor do mundo como o professor que previu o fim do jornal impresso no ano de 2043. A referência casual, na página 16 do penúltimo livro que publicou, acabou dominando todo o debate sobre a ambiciosa obra “The Vanishing Newspaper” (publicado no Brasil com o título “Os jornais podem desaparecer?”). As 258 páginas seguintes à suposta previsão, no original, são o relato de um estudo rigoroso, feito ao longo de dois anos, sobre como o investimento na qualidade editorial poderia ajudar a retardar o declínio das empresas jornalísticas.

O repórter da gravata borboleta

Como Meyer ainda é relativamente pouco conhecido no Brasil, vale a pena recapitular sua carreira. Philip Meyer nasceu em 1930, em Dreshler, no Nebraska, numa família de origem modesta. Ainda criança, lendo as primeiras tiras do Superman no jornal local, inspirou-se na história de Clark Kent para decidir o que gostaria de fazer da vida.

“Aquela história teve um efeito profundo nas minhas fantasias. Sendo franzino, eu nunca poderia ser um super-herói. Mas certamente poderia ser um repórter pacato. E, no mundo onde eu vivia, as palavras bem aplicadas poderiam ter poderes próprios. Foi aí que pela primeira vez a ideia de ter uma carreira em jornais passou pela minha cabeça”, escreveu em sua autobiografia, “Paper Route” (2012). Até a gravata que se tornou sua marca registrada foi inspiração de Tommy Bond, primeiro ator que interpretou o jovem fotógrafo Jimmy Olsen nos seriados dos anos 40. Phil contou à sua filha Kathy que começou a usá-la ainda na escola, para dar o que falar aos colegas.

Estudou jornalismo na Kansas State University entre 1948 e 1952,

onde teve suas primeiras experiências em jornais, e em seguida foi convocado para o serviço militar. Por dois anos, foi treinado para a guerra da Coreia, mas ela acabou antes que ele precisasse pegar em armas para valer. A experiência, porém, lhe deu acesso à “GI Bill”, que dava bolsas de estudo para ex-recrutas. Foi com ela que ele pôde cursar pós-graduação em ciência política na Universidade da Carolina do Norte entre 1956 e 1958.

No mestrado, aprendeu como funcionavam pesquisas de opinião. Até tentou aplicá-las quando voltou a trabalhar como repórter, mas a tecnologia disponível – cartões de cartolina coloridos – não ajudou. Levaria uma década para ele conseguir fazer uma boa pesquisa. Em 1958, Meyer começou uma carreira de duas décadas na rede de jornais Knight Newspapers. Inicialmente, como repórter de educação, cobriu os atritos na dessegregação das escolas para o “Miami Herald”. Foi um dos repórteres da sucursal de Washington da rede de jornais, onde teve a chance de entrevistar Martin Luther King e John Kennedy. Quando deixou a empresa para ser professor, era executivo de uma área de inovação.

Foi nesse período que vieram as oportunidades que o tornaram um inovador do jornalismo. No ano acadêmico de 1966/1967, Meyer foi bolsista Nieman em Harvard. Lá, estudou as mais modernas técnicas de pesquisa das ciências sociais e aprendeu a domar um computador IBM 360 usando a linguagem de programação Data-Text, que mais tarde daria origem ao SPSS usado até hoje por cientistas sociais.

Um dos seus colegas em Harvard foi Al Shuster, que mais tarde seria correspondente do *New York Times* em Londres, Saigon e Roma, além de editor de Internacional do *Los Angeles Times*. Em depoimento concedido para a homenagem a Meyer, disse que ninguém entendia o que o colega estava aprontando enquanto escapava de palestras dos

“Meyer enxergou, lá no tempo do cartão perfurado, que, num mundo onde o volume de informação aumenta rapidamente, é urgente que os jornalistas tenham autonomia para coletar dados, processá-los e explicá-los ao leitor.”

maiores professores de Harvard para fuçar no computador. "Eu falei pra ele: Phil, isso não vai dar certo, ninguém vai se interessar por isso. A gente gosta de máquinas de escrever, a gente gosta de ouvir o clic-clac delas na redação", disse Shuster em depoimento para a homenagem que fizemos. "Tá bom, Phil; você venceu."

Na noite de 23 de julho de 1967, na periferia de Detroit, uma festa em homenagem a dois soldados negros que voltavam do Vietnã foi interrompida pela prisão de todos os presentes. Isso levou à explosão violenta dos conflitos sociais latentes na periferia, atizados a cada novo episódio de brutalidade policial contra os negros. A repressão policial aos protestos deixou um saldo de 43 mortos, 1.189 feridos, 7.200 presos e mais de 400 prédios destruídos, muitas vezes pelo fogo. O custo humano daqueles dias está retratado no filme "Detroit em Rebelião" (2017).

De volta à redação, em Washington, após a experiência em Harvard, Meyer atendeu um telefonema de Jack Daniels, diretor do *Detroit Free Press*. A cobertura já durava três dias e a redação estava exausta; ele precisava de dois reforços e foi pedi-los a Washington. "Como fui eu que atendi ao telefone, pude mandar a mim mesmo e ao Sal Friedman", disse Meyer em entrevista a Charles Lewis, fundador do Consórcio Internacional de Jornalistas Investigativos, em 2012. "Na mesma noite, eu estava na área conflagrada, em cima de um caminhão da Guarda Nacional, observando os incêndios, ouvindo tiros e tentando juntar as peças do que estava acontecendo", contou.

Quando a poeira baixou e o jornal buscava um diferencial para avançar na cobertura, Meyer teve a ideia de fazer algo completamente diferente: uma pesquisa de opinião pública, ouvindo os moradores dos bairros conflagrados sobre as causas da revolta. Os editoriais locais especulavam a respeito de maneira bastante preconceituosa e racista, mas uma pesquisa de opinião, com amostragem cientificamente elaborada, ouvindo quem realmente morava na região, poderia ser um diferencial importante.

Com financiamento da Fundação Ford, que tinha sede na mesma cidade, e apoio da Liga Urbana de Detroit, a equipe conseguiu contratar professores negros para entrevistar em casa 437 pessoas que viviam na região sobre o que pensavam a respeito das revoltas. Todas as hipóteses do senso comum foram derrubadas pelo trabalho, que

apontou os frequentes casos de truculência policial como o principal fator. Com essa cereja no topo, a cobertura feita pelo jornal ganhou o Pulitzer em 1968.

Nos anos seguintes, Meyer se tornou um grande mentor de jornalistas interessados em aprofundar a metodologia de suas reportagens. Em 1973, em meio à cobertura do caso Watergate, Meyer orientou Jim Steele e Don Barlett sobre como fazer uma inédita análise computadorizada de dados de processos judiciais, na série "Crime e Injustiça", do *Philadelphia Inquirer*. Eles coletaram dados de mais de 400 processos, classificaram cada um dos casos e, após uma análise que demandou que o trio passasse várias madrugadas processando cartões perfurados, demonstraram que, quando condenados por um mesmo crime da mesma gravidade, réus brancos tinham sentenças mais leves do que réus negros.

No final dos anos 70, devido à sua experiência com análise de dados e pesquisas de opinião, a rede de jornais onde trabalhava, então chamada *Knight-Ridder*, o convocou para um projeto ambicioso: cuidar das pesquisas de mercado que avaliariam a aceitação do primeiro jornal digital do mundo, que chegaria à casa dos leitores via teletexto (um terminal ligado ao aparelho de TV) num serviço chamado Viewtron. A empresa investiu na época US\$ 50 milhões nesse projeto, que no final provou estar ao menos uma década à frente do seu tempo, e naufragou. Antecipando de novo o que viria a ser a internet, o que mais fez sucesso no teletexto sequer foram as notícias — era o chat. Os sexagenários ex-colegas de escola de suas filhas até hoje se referem a elas como "a filha do cara que inventou a internet". Não foi bem assim, mas não é que acaba sendo?

O pensador da inovação no jornalismo

Em 1981, depois de sua experiência como executivo de inovação, decidiu aceitar um convite para se tornar professor da Universidade da Carolina do Norte. Foi lá que trabalhou até 2008, quando se aposentou. No período, orientou dezenas de estudantes e produziu estudos relevantes sobre o negócio do jornalismo e sobre pesquisas de opinião pública.

A partir dos anos 80, também trabalhou como consultor do

jornal *USA Today*, onde recrutou a primeira equipe do mundo focada exclusivamente em Jornalismo de Dados, e da Sociedade Americana de Editores de Jornais (ASNE), para a qual produziu a pesquisa que levou ao livro *"Ethical Journalism"* (1987, publicado no Brasil pela editora Revan como *"A Ética no Jornalismo"*). Esse livro, aliás, foi feito completamente dentro da tradição do jornalismo de precisão: Meyer montou uma pesquisa sobre padrões de conduta e a aplicou a editores de jornais pequenos, médios e grandes do país inteiro, para ver como o que é aceitável varia de acordo com a estrutura da redação.

Essa parceria com a ASNE também levou à publicação do livro *"The Newspaper Survival Handbook"* (1985), que ensinava editores a analisar dados de circulação e outras pesquisas de mercado com um olhar jornalístico. Se *"Precision Journalism"* era voltado aos repórteres, este retrabalhava mais ou menos as mesmas técnicas para refinar o ponto de vista da chefia (gente que sempre julga saber tudo), e portanto nunca foi um livro muito popular.

Durante os anos 80 e 90, em paralelo às atividades como professor de jornalismo e consultor de redações, Meyer teve uma forte atuação no setor de opinião pública, exercendo o cargo de presidente da American Association of Public Opinion Research (AAPOR) entre 1989 e 1990, e da World Association of Public Opinion Research em 1993-1994. Chegou a participar dos conselhos editoriais de duas das mais importantes publicações acadêmicas da área, a *Public Opinion Quarterly* e o *International Journal of Public Opinion Research*.

Foi nessa fase que ele uniu suas duas áreas de interesse acadêmico, o jornalismo e as pesquisas de opinião, numa só proposta: o "jornalismo cívico". Pela proposta, a cobertura política devia usar ainda mais pesquisas de opinião, mas não no modo "corrida de cavalos", tipo qual candidato está na frente e qual vai perder. Deviam ser usadas para entender quais temas da sociedade são importantes para os eleitores, como os eleitores interpretam essas questões e o posicionamento da classe política em relação a elas. A proposta foi mais popular nos meios acadêmicos do que nas redações; algumas redações que anunciaram fazer algo do gênero acabaram usando mais para fins marqueteiros do que realmente de compreensão do público. E, quando virou o século, essa proposta já havia "flopado".

Nota: Quando Donald Trump ganhou sua primeira eleição,

em 2016, e as pesquisas de opinião foram popularmente vistas como falhas, escrevi a ele para perguntar como via a questão. Ele disse para ficar ligado porque a AAPOR estava fazendo uma auditoria das pesquisas e ainda brincou: "as pesquisas não erraram se, como disse o Nate Silver, havia sete chances em dez de Hillary passar de 50% e o resultado caiu numa das outras três; por chances menores de dar errado — uma em seis —, roleta russa não é um esporte muito popular".

Seu último grande projeto como professor era ambicioso: num cenário de encolhimento do mercado de jornais, Meyer decidiu estudar como os fatores que garantem a qualidade dos jornais ajudam a garantir sua sobrevivência. "The Quality Project" foi um projeto de pesquisa de dois anos, que gerou uma série de estudos publicados por Meyer e seus colaboradores nas principais revistas acadêmicas do setor. Nesses estudos, o velho Phil mobilizou todo o arsenal analítico que ensinava em suas aulas, só que em vez de olhar tendências sociais ele procurou relacionar dados do próprio setor pelo qual sempre foi apaixonado.

As conclusões foram consolidadas no livro "The Vanishing Newspaper" (2004), publicado no Brasil pouco tempo depois com o título "Os jornais devem desaparecer?". Nele, o principal alerta de Meyer era: cuidado com a enganosamente óbvia estratégia de lidar com o declínio do negócio por meio do corte na qualidade do jornalismo (principalmente repórteres, edição, cobertura local) visando manter as margens de lucro. Isso, explica, leva à corrosão da credibilidade da publicação, que é o que mantém sua relevância.

Como mencionado anteriormente, porém, quase todo o debate sobre o livro se resumiu a uma observação casual, feita na página 16,

“Conheci a terceira edição do livro clássico de Meyer em 1996, praticamente esquecida num canto da biblioteca da Fabico. Ao lê-lo, aquele garoto tímido que não se via como um grande entrevistador, mas tinha bastante experiência com computadores, encontrou uma vocação para sua vida.”

de que continuada a tendência observada até então a linha chega ao zero em 2043. Na *The Economist*, a descrição disso incluiu um último leitor exausto e um último exemplar amarfanhado. Meyer precisou lançar uma segunda edição para explicar o evidente: que nenhum empresário vai continuar imprimindo jornais até que sobre apenas um assinante para recebê-lo.

Aposentado desde 2008, Meyer acompanhou com interesse a explosão do Jornalismo de Dados e viu os mais inovadores trabalhos de reportagem ganharem o prêmio que leva seu nome, oferecido anualmente pela Investigative Reporters and Editors. Continuou escrevendo sobre pesquisas de opinião, política e jornalismo para o *USA Today* até 2013. No ano de 2016, em sua conta no Facebook, avisou aos amigos que tinha o mal de Parkinson e que sua qualidade de vida tendia a reduzir progressivamente nos anos que viriam. E ainda brincou: o Parkinson é que teria de suportá-lo até o fim.

O legado que inspira gerações de jornalistas de dados

Conheci a terceira edição do livro clássico de Meyer em 1996, praticamente esquecida num canto da biblioteca da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul (UFRGS), a Fabico. Como mencionei anteriormente, o livro era praticamente desconhecido no Brasil até há relativamente poucos anos, e naquela época era um bicho estranho na biblioteca. Ao lê-lo, aquele garoto tímido que não se via como um grande entrevistador, mas tinha bastante experiência com computadores, encontrou uma vocação para sua vida. Em 1998, consegui publicar minha primeira reportagem com análise de dados; em 2006, veio meu primeiro prêmio, com um projeto que começou com uma troca de e-mails com Meyer. Por duas décadas, caçei pistas perdidas até descobrir como foi que aquele exemplar foi parar na Fabico. Cheguei à conclusão de que foi doado a alguém pelo próprio Meyer, quando visitou Porto Alegre em 1991, e essa pessoa resolveu doar à biblioteca para ver se alguém se interessava. Calhou de cair na minha mão.

Embora nunca tenhamos nos visto pessoalmente, passei duas décadas consultando Meyer eventualmente por e-mail. Enquanto pôde, ele sempre respondia; às vezes me apresentava amigos seus, como Jim Steele e Everette Dennis (que batizou o jornalismo de precisão).

Na primeira troca, perguntei a ele como fazer Jornalismo de Dados no Brasil se não tínhamos tantos dados públicos disponíveis — era o fim dos anos 90, enfim. “Crie você mesmo seus bancos de dados desde já”, ele recomendou, e foi o que comecei a fazer.

Mais tarde, fiz em meu tempo livre uma tradução de “Precision Journalism” e tentei publicá-la no Brasil, mas nenhuma editora se interessou — mais ou menos o mesmo problema que ele teve com a primeira edição. Em 2011, Meyer tinha terminado de escrever sua autobiografia e perguntou se eu toparia fazer um pequeno texto para a contracapa da edição em capa dura. Sempre em outubro, eu fazia questão de telefonar para ele para lhe desejar feliz aniversário.

Ele foi o meu professor mais importante de toda a vida, e nunca pude apertar sua mão.

Em 2022, com seis anos de Parkinson diagnosticado, a saúde de Meyer estava tão comprometida que há anos ele não lia mais e-mails. Mantinha o hábito de ler ao sol, bebendo Coca-Cola de canudinho, a edição impressa do *New York Times*, com a qual às vezes cochilava abraçado.

Temendo haver pouco tempo para agradecê-lo como merecia, colaborei com a Open Knowledge Brasil para fazer a curadoria de uma sessão especial no Coda.BR daquele ano em homenagem ao velho professor. Ele não poderia participar, mas uma de suas filhas, Melissa, representaria a família e contaria histórias dele; o Knight Center for Journalism in the Americas patrocinou sua viagem. Expusemos todos os livros que ele publicou e fizemos uma exposição com fotos de todas as fases da sua vida. Recebi fotos dele assistindo à sessão em casa; ele reconheceu rostos e vozes dos amigos em comum que entrevistei para contarem suas memórias. Meyer morreria exatamente um ano depois.

Na sessão, eu e Melissa usávamos gravatas-borboleta, pois é o que Phil usaria. Em 2002, quando ele já havia feito as pazes com os “caubóis”, sua pupila Jennifer LaFleur fez gravatas-borboleta com cartões perfurados e fotografou Meyer e seus amigos usando-as.

Entre eles, estava Steve Doig, que se tornou jornalista de dados após ler a primeira edição de “Precision Journalism” e cuja investigação sobre a relação entre doações eleitorais de construtoras para tornar mais brando o plano diretor de sua cidade e o estrago feito pelo furacão que destruiu sua própria casa, em 1993, foi contada na quarta

edição. Em depoimento à homenagem, Doig disse que gostaria de ser Phil Meyer quando crescer.

É o que todos queremos, enfim. Mas outro assim não aparece tão fácil.

PARA SABER MAIS

- Íntegra da sessão em homenagem a Meyer e entrevistas com seus amigos. Disponível em: <https://lagomdata.com.br/phil/>
- Texto original da terceira edição de Precision Journalism, mesma que encontrei na biblioteca da Fabico. Disponível em: <https://carolinadatadesk.github.io/pmeyer/book/>
- "Paper Route", a autobiografia de Philip Meyer. Disponível em: https://books.google.com.br/books/about/Paper_Route.html?id=hstgDQAAQBAJ&redir_esc=y
- "The Newspaper Survival Book", o manual de Meyer para entender o público. Disponível em: <https://archive.org/details/newspapersurviva0000meyer>
- "A Ética no Jornalismo" - Revan, 1988 (fora de catálogo, mas disponível em qualquer biblioteca de jornalismo que se preze)
- "Os Jornais podem Desaparecer?" - edição brasileira de "The Vanishing Newspaper". Disponível em: <https://www.editoracontexto.com.br/produto/os-jornais-podem-desaparecer-como-salvar-o-jornalismo-na-era-da-informacao/1496581>
- As 35 obras sobre jornalismo mais significativas do século 20. Disponível em: <https://www2.gwu.edu/~media/press-releases/04-11-00-JMCQBooklist.html>

Feminismo de Dados: contribuições para pensar o jornalismo pelas lentes da interseccionalidade

Por Marília Gehrke

Baseado em sete princípios fundamentais (*veja no quadro*), o feminismo de dados se propõe a examinar as estruturas de poder associadas à metodologia, à coleta e à disseminação de dados. Por isso mesmo, toca em um ponto fundamental do jornalismo: os dados, aqui entendidos como fonte para o jornalista, devem ser examinados criticamente antes de sua apropriação e de seu uso.

A abordagem crítica de que “não há nada dado nos dados” e que os números não falam por si vem sendo explorada consistentemente por pesquisadores e jornalistas há pelo menos uma década. Em 2013, por exemplo, Lisa Gitelman editou o livro “Dados brutos são um oxímoro” para alertar que os dados são um produto cultural e pertencem ao contexto em que foram gerados. Em uma perspectiva similar, Yanni Loukissas escreveu “Todos os dados são locais” em 2019, também alertando para a coleta de dados e como esses dados eventualmente são utilizados para a criação de algoritmos e automatização de processos. Em outras palavras, as estruturas de poder associadas ao método da coleta de informações são determinantes para a organização de bases de dados e seu posterior uso. Em 2020, na Escola de Dados, eu e Adriano Belisário, Rodrigo Menegat e Marina Gama Cubas percorremos um caminho intelectual similar para escrever o livro “Fluxo de trabalho com dados: Do zero à prática”.

As abordagens mencionadas acima têm um fio condutor: as relações estruturais de poder que moldam os dados. Tendo essa perspectiva em mente e dando um passo a frente, as professoras estadunidenses Catherine D’Ignazio e Lauren Klein inovam ao empregar uma perspectiva de conhecimento situado e interseccional na obra “Feminismo de dados”, publicada em 2020. Em 2024, colocando essas diretrizes em prática, D’Ignazio publicou seu segundo livro,

chamado “Contando feminicídio”, sobre o qual vou falar mais adiante neste capítulo.

O feminismo de dados nos lembra que nem o jornalismo nem os dados existem fora de um tecido social organizado por hierárquicas estruturas de poder. Assim, D'Ignazio e Klein estruturam sua obra a partir do conceito de interseccionalidade, materializando esse conhecimento nos princípios-base que servem para pensar os dados criticamente. Tal conceito foi introduzido pela acadêmica jurídica Kimberlé Crenshaw no fim dos anos de 1980, nos Estados Unidos, para amparar a discriminação racial e de gênero, demonstrando que características interseccionais – ou seja, que se cruzam – de identidade frequentemente servem de base para a opressão.

Neste capítulo, exploro como a autora Catherine D'Ignazio lançou um novo olhar sobre a ciência de dados alinhada à proposta do feminismo interseccional e, com base em seu contato com ativistas da América Latina, aliou-se aos movimentos de reparação histórica em curso, incluindo uma proposta para revisitar a ciência de dados. Justamente porque explora as relações de poder, o feminismo de dados se torna ainda mais relevante no contexto da Inteligência Artificial generativa e suas implicações para as populações minorizadas.

A interseccionalidade nas formas de opressão

Cientes de que o *feminismo* possui diversas vertentes de acordo com o país e o contexto, D'Ignazio e Klein (2020, p. 6) usam “feminismo” como um termo guarda-chuva que abrange diversos projetos que desafiam o machismo e outras formas de opressão para, em contraste, criar um futuro mais justo e igual. O feminismo de dados desenvolve-se a partir do princípio da interseccionalidade para refletir sobre como uma sociedade sem distribuição igualitária de poder se estrutura e como segmentos se interseccionam para manter privilégios e oprimir certos grupos de pessoas. Machismo, racismo, capacitismo, colonialismo e classismo estão entre as formas de opressão recorrentes.

No Brasil, por exemplo, a opressão de mulheres negras foi estudada por Lélia Gonzalez, para quem as ideologias políticas e formas de dominação coloniais ainda são encontradas nas desigualdades

do cotidiano. Para ela, a falsa ideia de que o Brasil é uma democracia racial é problemática e esconde o fato de que a miscigenação advém do estupro de mulheres negras escravizadas. Historicamente, o papel de subserviência e obediência se materializa na divisão racial do trabalho. Mulheres negras e indígenas são submetidas à condição de empregada doméstica, desta

vez confinadas ao lar de um patrão branco, como se esse fosse o único lugar a elas destinado. Assim, mesmo sem o termo "interseccionalidade" (pelo qual Crenshaw ficou conhecida) no nome, é disso que trata Gonzalez.

A conexão de D'Ignazio com a América Latina, por sua vez, se dá pela Argentina, onde viveu parte de seus anos em diferentes fases da vida. Recentemente, entrou em contato com ativistas de dados da região, principalmente na preparação de seu segundo livro. D'Ignazio se autoidentifica como "mamãe hacker" (*hacker mama*), artista experimental, e professora, além de cientista de dados e feminista de dados. É líder do Data+Feminism Lab no MIT, nos Estados Unidos, país em práticas anteriores ao Jornalismo de Dados, tais como o jornalismo de precisão e a Reportagem Assistida por Computador (RAC), foram desenvolvidas.

“ O feminismo de dados nos faz lembrar de que nem o jornalismo nem os dados existem fora de um tecido social organizado por hierárquicas estruturas de poder”

O feminismo de dados encontra o jornalismo

Vertente moderna do jornalismo de precisão de Philip Meyer (*leia o capítulo sobre o tema*), o Jornalismo de Dados foi calcado em uma ideia científica de método, reprodutibilidade e transparência que segue inspirando acadêmicos e repórteres. Meyer fundou essa perspectiva como uma alternativa ao jornalismo declaratório, ou seja, baseado apenas em fontes humanas.

O problema, particularmente em tempos de Inteligência Artificial generativa, é que essa visão abre precedente para pensar que os

dados numéricos são *sempre* melhores, mais objetivos e mais acurados do que pessoas. No entanto, essas informações são pensadas e coletadas por pessoas que também revelam seus preconceitos e vieses (ainda que não admitam). Em meus tempos de repórter, rapidamente internalizei o que alguns manuais de redação da imprensa de referência recomendavam: fontes documentais (aqui incluídas estatísticas) são mais confiáveis.

Para refletir sobre vieses, veja, por exemplo, a matéria escrita por Rodrigo Menegat e publicada pela *Deutsche Welle* em 2021 sobre os mecanismos de busca do Google e como mulheres de diferentes países são retratadas (*veja a referência no final do capítulo*). Ainda que todas as mulheres sejam, em alguma medida, encaixadas em estereótipos de gênero, o fator raça – mesmo sendo uma construção social – tem peso extra. Mulheres brasileiras e latino-americanas são frequentemente sexualizadas quando buscadas no Google, o que não ocorre, por exemplo, com mulheres alemãs.

O feminismo de dados nos leva a pensar, então, por que informações sobre algumas populações são coletadas e outras são simplesmente ignoradas. Isso fica evidente em reportagem escrita por Vitória Régia da Silva e publicada pelo site *Gênero e Número* em 2021, época da pandemia de Covid-19 e isolamento social (*veja a referência no final do capítulo*). Segundo a matéria, mulheres negras e indígenas eram as principais pessoas expostas a situações de violência em casa ou eram as responsáveis pelo cuidado de pessoas da sua própria família ou de famílias de outros, caso das empregadas domésticas. A vulnerabilidade que vem com a desigualdade social, então, afeta essas grupos historicamente oprimidos. Quando dados sobre gênero e raça não são coletados, perde-se a granularidade desse tipo de conhecimento.

Como o jornalismo guiado por dados trabalha justamente com dados publicamente disponíveis, e nem sempre o jornalista de dados ou redação possui condições para coletar algo novo, torna-se ainda mais essencial examinar criticamente os dados antes de usá-los. Por isso, muitas redações e projetos dentro e fora do jornalismo apostam em coletar informações que não são normalmente contabilizadas por governos e órgãos oficiais.

O feminismo de dados na prática

Em seu segundo livro, “Contando feminicídio”, D’Ignazio traz a história de mulheres que, voluntariamente e devido à ausência ou insuficiência do Estado na contagem de feminicídios – o assassinato de mulheres por causa de seu gênero, ou seja, por conta de ódio e misoginia –, decidiram contabilizar os casos e dar nome às vítimas. No México, país em que não havia uma formulação legal para feminicídio e, por isso, não havia formas de contabilizar esses crimes, a ativista María Salguero reuniu cerca de 5 mil casos de 2016 a 2020 ao acessar notícias e sites governamentais em busca de pistas.

Essa contagem independente de dados perdidos ou não adequadamente mantidos por órgãos oficiais é o que D’Ignazio nomeia de “contradados” (*counterdata*), e não seria de conhecimento público caso uma pessoa não decidisse contabilizar tais números e dar rosto às pessoas. Por isso, trata-se do feminismo de dados em ação: desafia a lógica hegemônica do Estado e de estruturas convencionais de poder. Ao reportar tais iniciativas no livro, D’Ignazio reconhece, pelos mesmos princípios do feminismo de dados, que a atividade laboral das ativistas envolve trabalho criativo, intelectual e emocional.

Depois de criar os princípios do feminismo de dados juntamente com Klein, D’Ignazio diz que seu segundo livro busca oferecer meios de criar uma ciência de dados restaurativa e transformadora. O aspecto restaurativo visa oferecer direitos e dignidade às comunidades historicamente oprimidas. Já o aspecto transformador tem como proposta dismantelar e mudar as condições estruturais que produziram determinado tipo de violência.

Um dos muitos aspectos inspiradores da trajetória de D’Ignazio, além de sua contribuição intelectual, é o retorno prático dado a pessoas como a ativista María Salguero. Em seu laboratório Data+Feminism, D’Ignazio e seu time desenvolveram uma ferramenta que auxilia as

“ Um dos muitos aspectos inspiradores da trajetória de D’Ignazio, além de sua contribuição intelectual, é o retorno prático ”

ativistas de dados a encontrar notícias que exponham potenciais casos de feminicídio não contabilizados pelo Estado. A ferramenta auxilia as mulheres voluntárias a agilizar um trabalho denso que muitas vezes começam com uma simples planilha compartilhada. Seus efeitos em potencial são muitos, incluindo trazer para o jornalismo o debate de que nem sempre os feminicídios são adequadamente reportados pela imprensa. Não raro, esses casos são tratados como excepcionais, e não como resultado de uma violência endêmica calcada no colonialismo e na opressão que cria desigualdades.

A proposta de colocar o feminismo de dados em prática também chegou até mim: em 2024, criei o projeto *Forced to Quit* (Forçada a desistir) após não encontrar informações numéricas sobre o número de mulheres na política, no jornalismo e no ativismo forçadas a deixar seus trabalhos na esfera pública por conta de abusos, campanhas difamatórias e todos os tipos de violência. Essa iniciativa, em minha perspectiva, compreende o que o Jornalismo de Dados e o feminismo de dados têm de melhor: o caráter voluntário e colaborativo guiado por um propósito de dar nome às pessoas e à situação de violência que tira as mulheres da participação na vida pública.

Em alguns casos, como o assassinato da vereadora Marielle Franco, ser 'forçada a desistir' mostra a face mais cruel da violência. *Forced to Quit* é um projeto que une site e mapa, que só pôde ser desenvolvido em colaboração com meus colegas do departamento Geodienst da Universidade de Groningen. Assim como as ativistas de dados e feministas que me antecederam, este é o legado que busco deixar por meio do projeto: informação detalhada, passível de download, para novas análises e descobertas.

O futuro: buscando novos princípios

O recente desenvolvimento da Inteligência Artificial generativa abre espaço para uma extensa análise pelas lentes do feminismo de dados. Tendo em vista que esse aporte teórico-prático é principalmente sobre poder – quem tem e quem não tem o poder – e justiça, faz-se ainda mais urgente para estudar, por exemplo, as implicações da concentração de poder e dinheiro nas mãos de poucos indivíduos, em geral dos donos de grandes empresas de tecnologia.

Quando estive na Universidade de Groningen, em janeiro de 2025, D'Ignazio disse que os sete princípios do feminismo de dados não são o fim ou diretrizes permanentes, mas o começo de um diálogo. A provocação que fica é acompanhar os próximos desdobramentos e construir algo coletivamente, incluindo novos princípios do feminismo de dados para pensar sobre a tecnologia e a exploração colonialista do meio ambiente e de grupos de pessoas. Mais do que isso, como garantir que a ciência de dados inclua o caráter restaurativo e transformador do qual fala D'Ignazio.

7 PRINCÍPIOS DO FEMINISMO DE DADOS

1

Examinar o poder

2

Desafiar o poder

3

Elevar a emoção e a corporalidade

4

Repensar o binarismo e as hierarquias

5

Abraçar o pluralismo

6

Considerar o contexto

7

Tornar o trabalho visível

PARA SABER MAIS

- A ciência de dados tem muito a aprender com as feministas, por Ana Carolina Araújo (AzMina, 2024). Disponível em: <https://azmina.com.br/reportagens/a-ciencia-de-dados-tem-muito-a-aprender-com-as-feministas/>.
- Na pandemia, mulheres ficam mais vulneráveis e são maioria entre desempregados, por Vitória Régia Silva (Gênero e Número, 2021). Disponível em: https://www.generonumero.media/reportagens/mulheres-trabalho/#index_11.

- Yo te nombro: el mapa de los feminicidios en México. Disponível em: <https://mapafeminicidios.blogspot.com/p/inicio.html>.
- Forced to Quit. Disponível em: <https://forcedtoquit.com/>.
- 'Racy' women on Google image search, por Rodrigo Menegat (DW, 2021). Disponível em: <https://www.dw.com/en/google-image-search-cements-national-stereotypes-of-racy-women/a-56767605>.

REFERÊNCIAS

BELISÁRIO, Adriano et al. 2020. **Fluxo do trabalho com dados**: Do zero à prática. São Paulo: Open Knowledge Brasil.

D'IGNAZIO, Catherine. 2024. **Counting femicide**: Data feminism in action. Cambridge: MIT Press.

D'IGNAZIO, Catherine; KLEIN, Lauren. 2020. **Data Feminism**. Cambridge: MIT Press.

GITELMAN, Lisa. 2013. **"Raw Data" Is an Oxymoron**. Cambridge: MIT Press.

LOUKISSAS, Yanni Alexander. 2019. **All Data Are Local**: Thinking Critically in a Data-Driven Society. Cambridge: MIT Press

Precursores do Jornalismo de Dados no Brasil: como os jornalistas brasileiros começaram a entrevistar planilhas

Lívia de Souza Vieira

O Jornalismo de Dados é, hoje, uma prática consolidada em diversas redações do Brasil e do mundo. Mas essa combinação de rigor jornalístico com análise de bases de dados tem raízes que remontam às décadas de 1980 e 1990 no Brasil. Naquele período, antes mesmo da popularização da internet e do acesso facilitado a ferramentas de análise, alguns jornalistas brasileiros começaram a perceber que as informações disponíveis em planilhas, cadastros públicos e documentos administrativos poderiam ser tão valiosas quanto entrevistas e declarações.

Este capítulo apresenta uma linha do tempo dos precursores do Jornalismo de Dados no Brasil, explorando como os profissionais começaram a “entrevistar planilhas” e a produzir reportagens de alto impacto a partir da análise de dados. Também abordamos como essas experiências deram origem a uma comunidade ativa e colaborativa, que fortalece até hoje essa prática no país.

O surgimento

O uso de dados no jornalismo não é, de fato, uma invenção recente. Desde o surgimento do jornalismo investigativo, repórteres sempre trabalharam com informações documentais, estatísticas e registros oficiais. No entanto, o que chamamos hoje de Jornalismo de Dados é algo que se estrutura a partir do uso sistemático de bancos de dados, softwares de análise e visualização de informações, além de metodologias próprias que dialogam tanto com o jornalismo quanto com áreas como estatística e ciência de dados.

Nos anos 1990, alguns veículos jornalísticos brasileiros começaram a experimentar o uso de dados em reportagens investigativas. Fernando Rodrigues e José Roberto de Toledo passaram a adotar técnicas da reportagem assistida por computador (RAC) na *Folha de S. Paulo*. Em 1995, Toledo começou a desenvolver

métodos avançados de pesquisa na internet, analisando pesquisas de opinião pública e adaptando técnicas norte-americanas de uso de planilhas e bancos de dados à realidade brasileira. Para documentar e acompanhar a evolução da RAC no Brasil e no mundo, ele criou o blog Toledol¹. Em 1998, Fernando Rodrigues começou a construção do banco de dados “Políticos do Brasil”, cujo material teve de ser coletado diretamente em todos os 27 Tribunais Regionais Eleitorais e no TSE, pois nada era digital. Em 2006, o trabalho virou livro e, em 2018, foi atualizado e migrou para a base do Poder 360², veículo criado pelo jornalista.

No início dos anos 2000, surgiram reportagens emblemáticas que utilizavam dados de forma central. Em 2004, o jornal *O Globo* publicou a série “Homens de bens da Alerj³”, coordenada pela repórter Angelina Nunes. Com base em declarações de renda entregues ao TRE do Rio de Janeiro, o trabalho minucioso de investigação mostrou o enriquecimento de 113 políticos fluminenses, dos quais 27 aumentaram o patrimônio em mais de 100% entre 1996 e 2001. Em 2009, Daniel Bramatti publicou, em *O Estado de S. Paulo*, a reportagem

“ Desde o surgimento do jornalismo investigativo, repórteres sempre trabalharam com informações documentais, estatísticas e registros oficiais”

1 Disponível apenas no banco de dados da WebArchive: <https://web.archive.org/web/20170201092202/https://toledol.com.br/about/>

2 Disponível em: <https://monitor.poder360.com.br/politicos-do-brasil>

3 Mais informações: <https://memoria.oglobo.globo.com/jornalismo/premios-jornalisticos/os-homens-de-bens-da-alerj-8876006>

"Números derrubam mito do grande palanque peemedebista"⁴, cuja análise de dados das eleições municipais de 2006 relativizou a importância do apoio do PMDB no desempenho de candidatos. Em 2010, o jornal Gazeta do Povo e a RPCTV publicaram a série de reportagens "Diários Secretos"⁵, que revelou casos de corrupção na alta cúpula da Assembleia Legislativa do Paraná. Os repórteres Karlos Kohlbach, Katia Brembatti, James Alberti e Gabriel Tabatcheik descobriram nomeações e exonerações de funcionários comissionados, impressas em edições avulsas do Diário Oficial da Alep, que nunca haviam se tornado públicas até a divulgação das reportagens.

A consolidação

Um dos fatores decisivos para a consolidação do Jornalismo de Dados no Brasil foi a promulgação da Lei de Acesso à Informação (Lei nº 12.527), em 2011⁶. Com ela, qualquer cidadão - especialmente jornalistas - passaram a ter um instrumento legal robusto para solicitar dados públicos de órgãos governamentais. Outra consequência da lei foi a criação dos portais de transparência, que são obrigados a disponibilizar informações como salários de funcionários públicos, gastos de pessoas e órgãos, licitações, entre outros. Isso representou um salto qualitativo na prática jornalística, uma vez que a disponibilidade de dados permitiu não apenas investigações mais profundas, mas também a produção de reportagens baseadas em evidências concretas.

Embora com muitos avanços, a efetiva aplicação da LAI ainda possui entraves, como a decretação de sigilo de documentos que impedem a investigação jornalística. Em que pese ser uma prerrogativa prevista na lei, na prática, acaba se tornando um mecanismo com potencial de prejudicar a transparência das informações públicas.

Ao longo de mais de uma década, são muitas as reportagens

4 Apenas parte da reportagem está disponível no arquivo do Senado Federal: <https://www2.senado.leg.br/bdsf/bitstream/handle/id/341161/noticia.htm?sequence=1&isAllowed=y>

5 Mais informações: <https://www.gazetadopovo.com.br/parana/15-anos-dos-diarios-secretos-relembra-a-necessidade-de-reacao-da-sociedade-aos-abusos-das-autoridades/>

6 Disponível em: <https://www.gov.br/aeb/pt-br/aceso-a-informacao/lei-de-aceso-a-informacao>

feitas com dados e informações obtidos através da Lei de Acesso à Informação⁷. Em 2015, o jornalista Léo Arcoverde criou o site Fiquem Sabendo, que deu origem à agência de dados especializada na LAI em 2018. Em 2016, uma série de reportagens do jornal O Globo⁸ mostrou casos em que centenas de órgãos e tecidos humanos, que seriam usados em transplantes, acabaram sendo perdidos por falta de transporte aéreo imediato. O episódio fez com que o governo federal editasse uma medida para que aeronaves da Força Aérea Brasileira (FAB) ficassem permanentemente de prontidão para atender esse tipo de emergência em qualquer parte do país, em tempo integral.

Em 2020, em meio a um “apagão de dados” que restringiu o acesso a informações sobre a pandemia de covid-19 pelo governo Bolsonaro, foi criado um consórcio inédito entre *O Globo*, *Extra*, *O Estado de S. Paulo*, *Folha de S. Paulo*, *G1* e *UOL*⁹. A iniciativa foi uma resposta à violação da Lei de Acesso à Informação e da própria Constituição do país, e teve grande importância na divulgação de dados sobre a evolução da doença.

A comunidade

Outro sinal de amadurecimento do Jornalismo de Dados no Brasil foi o surgimento de grupos, associações e comunidades. Entre elas, destaca-se a Associação Brasileira de Jornalismo Investigativo¹⁰ (Abraji), criada em 2002, a partir de uma iniciativa do jornalista Marcelo Beraba, para o aprimoramento profissional dos jornalistas; a difusão dos conceitos e técnicas da reportagem investigativa; e a defesa do direito de acesso a informações públicas. Outro exemplo é a Escola

7 Mais informações: <https://www.brasildefato.com.br/2022/05/19/lei-de-acesso-a-informacao-criou-cultura-da-transparencia-mas-ainda-sofre-obstaculos/>

8 Disponível em: <https://oglobo.globo.com/politica/sistema-de-transplante-desperdicou-quase-mil-orgaos-em-cinco-anos-19447655#>

9 Mais informações: <https://www.abraji.org.br/noticias/acao-colaborativa-une-veiculos-para-evitar-apagao-de-dados-sobre-pandemia>

10 Mais informações: <https://www.abraji.org.br/>

de Dados¹¹, iniciativa da Open Knowledge Brasil, que oferece cursos, oficinas e encontros voltados para a capacitação de jornalistas e ativistas no uso de dados. Desde 2016, realizam também a Conferência Brasileira de Jornalismo de Dados e Métodos Digitais (Coda.Br)¹², que se tornou referência na América Latina, reunindo jornalistas, pesquisadores, desenvolvedores e cientistas de dados para discutir metodologias, ferramentas e tendências da área.

Há ainda comunidades informais em aplicativos de mensagem como o WhatsApp, que reúnem jornalistas brasileiros que trabalham com dados, tornando-se locais de trocas de informações, de referências e uma possibilidade de estreitar os laços profissionais no sentido da colaboração.

Para reconhecer a produção em Jornalismo de Dados no país foi criado, em 2019, o Prêmio Cláudio Weber Abramo de Jornalismo de Dados¹³, iniciativa da Escola de Dados, em parceria com a Abraji. O nome do prêmio é uma homenagem ao jornalista Cláudio Weber Abramo, um dos pioneiros no Jornalismo de Dados e transparência pública no Brasil.

A pesquisa

O interesse acadêmico pelo Jornalismo de Dados acompanhou o desenvolvimento da prática profissional. Em 2001, o professor da Universidade Federal de Santa Catarina, Nilson Lage, publicou o livro

“Embora com muitos avanços, a efetiva aplicação da LAI ainda possui entraves, como a decretação de sigilo de documentos que impedem a investigação jornalística”

11 Mais informações: <https://escoladedados.org/>

12 Mais informações: <https://ok.org.br/projetos/coda-br/>

13 Mais informações: <https://premio.jornalismodedados.org/about/>

"Teoria e Técnica de Reportagem, Entrevista e Pesquisa Jornalística"¹⁴, com um capítulo intitulado "Reportagem Assistida por Computador". Lage conta que o livro é resultado de cursos que ministrou em disciplinas de técnicas de jornalismo na UFSC entre 1992 e 1999, e que marca um momento de transformação da reportagem, "com a entrada em cena dos computadores, da internet e dos bancos de dados".

Outro trabalho pioneiro foi o TCC de Marcelo Soares, iniciado em 1998 e finalizado em 2004¹⁵. Com o título "Contribuição ao estudo das condições brasileiras para o uso das técnicas de reportagem auxiliada por computador (Computer-Assisted Reporting)", a monografia, defendida na Universidade Federal do Rio Grande do Sul, marcou o início da carreira de Marcelo Soares, que se tornou um dos grandes especialistas em Jornalismo de Dados do país.

Em 2007, Suzana Barbosa defendeu, na Universidade Federal da Bahia, a tese "Jornalismo digital em base de dados (JDBD): um paradigma para produtos jornalísticos digitais dinâmicos", que identificou os usos e as transformações que a utilização de bases de dados estava fazendo emergir para o jornalismo digital.

Outro destaque é a tese de Marcelo Träsel, que inspirou o título deste capítulo - "Entrevistando planilhas: estudo das crenças e do ethos de um grupo de profissionais de jornalismo guiado por dados no Brasil". Defendida em 2014 na Pontifícia Universidade Católica do Rio Grande do Sul, a pesquisa investigou valores, crenças e condutas dos profissionais de jornalismo guiado por dados no Brasil.

Uma breve busca pelo termo "Jornalismo de Dados" no Google Scholar mostra mais de 1.000 resultados, entre artigos, dissertações, teses e livros. Isso evidencia a relevância do tema como objeto de pesquisa, que desperta interesse dos acadêmicos brasileiros.

Ao longo desses anos, diversas universidades brasileiras passaram a incorporar disciplinas específicas sobre Jornalismo de Dados em seus cursos de graduação, ampliando a formação de novos profissionais com essa competência.

14 Mais informações: <https://nilsonlage.com.br/wp-content/uploads/2017/10/A-reportagem.pdf>

15 Mais informações: <https://portalimprensa.com.br/imprensa%2Bduca/conteudo/79333/tcc%2B-de%2Bmarcelo%2Bsoares%2Bfoi%2Bpioneiro%2Bno%2Bbrasil%2Bpara%2Bo%2Bjornalismo%2B-de%2Bdados>

O presente e o futuro

Atualmente, há equipes especializadas de Jornalismo de Dados em grandes redações, como *Folha de S.Paulo*, *O Globo*, *G1*, *UOL* e *Poder360*. Além disso, veículos nativos digitais, como *Nexo Jornal* e *Agência Pública*, mantêm núcleos de dados que trabalham tanto em investigações profundas quanto na visualização de informações para o público.

A expansão do acesso a ferramentas gratuitas e de código aberto como R e Python também democratizou a prática, permitindo que jornalistas de diferentes regiões do país possam atuar com dados, mesmo fora dos grandes centros.

Por fim, as fronteiras entre Jornalismo de Dados e Inteligência Artificial começam a se estreitar, com redações experimentando o uso de modelos de aprendizado de máquina para análise de grandes volumes de dados e automação de reportagens.

PARA SABER MAIS

Cursos e Tutoriais

- **Escola de Dados** – Cursos online, e-books e tutoriais de Jornalismo de Dados. Disponível em: <https://escoladedados.org/>
- **Knight Center for Journalism in the Americas** – Cursos online de Jornalismo de Dados. Disponível em: <https://journalismcourses.org/>

Audiovisual

- **Spotlight: Segredos Revelados (2016)** - filme sobre a história real de uma investigação jornalística baseada em documentos e entrevistas sobre casos de pedofilia na igreja Católica. Disponível em plataformas de streaming.
- **Privacidade Hackeada (2019)** – documentário sobre o escândalo da Cambridge Analytica. Disponível em plataformas de streaming.

REFERÊNCIAS

BARBOSA, Susana. **Jornalismo digital em base de dados (JDBD): um paradigma para produtos jornalísticos digitais dinâmicos** (tese) - Faculdade de Comunicação, Universidade Federal da Bahia, Salvador, 2007.

LAGE, Nilson. **Teoria e Técnica de Reportagem, Entrevista e Pesquisa Jornalística**. [Versão Eletrônica]. 2001.

RODRIGUES, Fernando. **Políticos do Brasil**. São Paulo: Publifolha, 2006.

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

Dados também são de humanas! Por que perder o medo dos números qualifica o jornalismo

Liliane de Lucena Ito

É fato: a cada semestre lecionando a disciplina de Jornalismo de Dados na graduação, percebo no olhar dos universitários um certo desespero, logo no primeiro dia de aula, quando aviso que operações matemáticas, planilhas e pensamento lógico farão parte do conteúdo do semestre.

O receio, no entanto, é compreensível. Afinal, quantos de nós, jornalistas, escolhemos a profissão por diversos fatores, mas principalmente devido à habilidade comprovada com textos (e talvez uma certa inabilidade com números)? Esse “talento” para a parte textual vem junto a um histórico de elogios de docentes de língua portuguesa; quando não por ótimas pontuações na redação do Enem ou do vestibular. E, se este mesmo histórico passa por desempenho mediano (ou ruim) em disciplinas de Exatas, vai se sedimentando a ideia de que: “não me dou bem com números, meu negócio é Humanas”. Fica fácil entender, numa trajetória assim, a escolha por uma carreira na qual quase nenhuma disciplina de graduação traz conteúdos matemáticos.

Apesar de ter surgido na História como um desenvolvimento natural da cultura, a partir da agropecuária e da agricultura no período Neolítico (10 mil anos a.C.), o conhecimento matemático foi sendo aos poucos descolado de suas utilidades práticas para níveis de abstração que nem sempre eram abertamente democratizados, justamente para promover uma distinção entre “capazes” e “incapazes”, funcionando então como uma ferramenta de poder. Schuré (1986), explica em *Os Grandes Iniciados* que, no século VI a.C., geometria e aritmética passam a ser consideradas ciências e que, à época, a matemática era vista com uma aura religiosa.

Por sua vez, os filósofos que quisessem pertencer ao Instituto de Pitágoras passavam por provas terríveis para fazer parte do gru-

po: para o candidato pertencer ao Instituto de Pitágoras, era obrigado a passar a noite em uma caverna que havia nas proximidades da cidade, onde se lhe fazia crer que existiam monstros e se davam aparições. Aqueles que não tivessem coragem para suportar as impressões fúnebres da solidão e que se recusassem a entrar na caverna, ou que saíssem antes do amanhecer, eram julgados incapazes para a iniciação e despedidos.

A prova moral era mais séria. Bruscamente, sem preparação prévia, o candidato (o discípulo) era trancado em uma cela, onde deveria descobrir o sentido de um dos símbolos pitagóricos, por exemplo: "Que significa o triângulo inscrito em círculo?" ou "Por que é que o dodecaedro compreendido na esfera é a cifra do universo?". O candidato passava doze horas trancado nesta cela tentando decifrar o seu problema, sem outra companhia a mais que um vaso com água e pão seco. Após as doze horas, ele era conduzido a uma sala, com a presença de todos os noviços reunidos, que, nessa circunstância, tinham ordem de zombarem sem piedade do infeliz, o qual, aborrecido e com fome, parecia um criminoso. Irritado pela gozação, humilhado por não ter podido decifrar o enigma incompreensível, deveria fazer um esforço enorme para conter-se (Tato; Scapin, 2004, p. 61).

Algo de residual da Escola Pitagórica ecoa, ainda hoje, em relação ao ensino da matemática, muito pelo fato de que alguns docentes seguem orgulhosos o caminho opressor para com seus alunos. Para Fragoso (2001, p. 98), isso tem até nome: o professor *algebrista*. "Tal denominação é utilizada no sentido pejorativo, para designar todo aquele que vive possuído da preocupação mórbida de complicar [...] o ensino da Matemática".

O fato de haver um descolamento da matéria ao dia a dia do alunado é outra forte razão para que não se compreenda a utilidade da matemática na vida prática. Frente a cálculos que exigem passos longos, uso de fórmulas herméticas e operações indecifráveis, é comum que ouçamos, desde muito crianças, que "a pior matéria é a matemática" ou "matemática é difícil" quando, na verdade, o problema está em *como* essa disciplina é apresentada e ensinada durante os anos de formação escolar.

Vale ressaltar, no entanto, uma outra dificuldade para além da matemática. Trata-se da ausência de habilidades básicas de infor-

mática no caso das gerações mais jovens, como a chamada Geração Z (nascidos entre 1995 e 2010). Conhecer e manejar recursos de softwares de planilhas, que muitas vezes são passos iniciais numa investigação jornalística de dados, é uma dificuldade para os nativos digitais que nasceram em um contexto no qual o consumo de mídia ocorre principalmente por meio do smartphone. Acostumados a utilizar aplicativos e plataformas de redes sociais para quase tudo: conversação, edição de fotos e vídeos, comunicação, dentre outros, os estudantes se deparam com dificuldades básicas, uma vez que “entrevistar” planilhas ou realizar a raspagem de sites funciona muito melhor num computador desktop.

“ No Jornalismo de Dados, a aplicação matemática evidencia como os dados se comportam, mostrando um padrão comum ou uma informação que saia deste padrão – e o olhar jornalístico é que ditará qual enfoque é o mais importante”

Jornalista de dados sem ser expert

A boa notícia é que, mesmo para quem prestou Jornalismo para ficar longe da matemática, é possível produzir boas investigações baseadas em dados. Isso acontece principalmente porque grande parte das perguntas mais importantes realizadas quando se está diante dos dados pode ser resolvida com operações matemáticas ou estatísticas simples – como soma, porcentagem, média, moda e mediana, por exemplo.

No Jornalismo de Dados, a aplicação matemática evidencia como os dados se comportam, mostrando um padrão comum ou uma informação que saia deste padrão – e o olhar jornalístico é que ditará qual enfoque é o mais importante. Assim sendo, a partir das perguntas iniciais de investigação, os cálculos a serem feitos serão os mais simples – para buscar disparidades e elementos comuns. O

importante é seguir realizando as perguntas, ou seja, seguir “entrevistando” a planilha até chegar na informação que se deseja.

Imagine uma investigação que busque mostrar a distribuição dos salários de servidores públicos de uma determinada instituição. Você pode começar fazendo uma simples soma, a fim de descobrir qual é o montante total pago àquele grupo de funcionários públicos. Mas, se dividir esse montante pelo número de funcionários, chegará a uma média – e digamos que seja de R\$ 5000. Já foram duas operações matemáticas simples realizadas. No entanto, você segue fazendo perguntas e começa a perceber informações relevantes – por exemplo, ao calcular a moda, percebe que três funcionários recebem salários cinco vezes maiores do que a média encontrada. Já a grande maioria dos outros servidores, valores menores. A partir destas três operações matemático/estatísticas simples, você já descobriu que há uma disparidade considerável entre o salário dos funcionários da instituição. Se parássemos por aqui, seria possível produzir um infográfico com os detalhes sobre os supersalários, por exemplo.

Obviamente, nem todas as investigações de Jornalismo de Dados funcionam apenas com operações básicas. No entanto, grandes projetos são geralmente realizados por equipes multifacetadas, com profissionais que irão cuidar especificamente de cada parte da reportagem. Em grandes redações jornalísticas profissionais que trabalham com dados há, inclusive, a presença de estatísticos contratados (Ito; Ventura, 2016), essenciais para cálculos mais sofisticados.

É comum que projetos maiores de Jornalismo de Dados, mesmo que independentes, sejam assinados por diversos profissionais, de formações em áreas distintas do conhecimento. Além de jornalistas, frequentemente fazem parte do expediente cientistas de dados ou estatísticos, ilustradores, profissionais de audiovisual, desenvolvedores de sistemas, entre outros. Ou seja, mesmo que existam jornalistas de dados que sabem programar, em alguns projetos estes ficarão com a parte de coordenação editorial ou com o trabalho de campo, não necessariamente produzindo o código para manejar ou apresentar os dados.

Entretanto, ter uma noção das operações matemáticas, de estatística básica, além de saber manejar softwares de planilhas e conhecer um pouco do que é possível produzir com programação

são, sim, diferenciais importantes para quem quer trabalhar como jornalista de dados. Essa “noção” servirá para que, no contexto de um projeto em equipe, o jornalista saiba até onde ir ou o que (e como) solicitar aos outros membros especialistas do seu time de trabalho. E o melhor nisso tudo é que o nível inicial pode ser desenvolvido por meio de recursos gratuitos, como cursos, webinars, tutoriais e

vídeos disponíveis na Internet, assim como investir em congressos ou encontros que têm por objetivo oferecer ferramental prático ao público.

“ É comum que projetos maiores de Jornalismo de Dados, mesmo que independentes, sejam assinados por diversos profissionais, de formações em áreas distintas do conhecimento”

Melhorando seu desempenho com números

Tendo em vista que ignorar totalmente a dimensão numérica incorre em erros fatais para qualquer reportagem de dados (imagine o exemplo dos salários citado anteriormente levando-se em consideração apenas a média), é essencial lembrar, dos tempos da escola, operações matemáticas e, se for o caso, considerá-las tais quais uma técnica a ser aprendida, como a clássica composição do *lead* ou a construção de um bom roteiro de entrevista jornalística.

Assim, no final deste capítulo, estão disponibilizadas algumas opções gratuitas que podem levar você, estudante, a se aprofundar nas habilidades básicas para produzir investigações com base em dados.

Considerações finais

Num país em que 85% dos lares têm acesso à Internet e 60% deste acesso é feito unicamente via smartphone (CeticBR, 2024), é essencial que a formação acadêmica em Jornalismo forneça oportunidades para que os discentes da Geração Z desenvolvam habilidades

básicas de informática. A partir disso, os cursos de graduação que possuem no currículo a disciplina específica de Jornalismo de Dados deveriam incluir, em seu ementário, tais conhecimentos voltados ao manejo de dados em softwares de planilhas, além de conteúdos da área da matemática aplicada ao jornalismo. Acredita-se que este seria um passo essencial a uma formação mais completa dos estudantes de graduação.

Fazer com que os estudantes e profissionais da área percam o medo dos números é um fator decisivo para qualificar o jornalismo, uma vez que toda grande investigação que envolva dados necessita, em algum momento, de uma compreensão matemática elementar para que se realize inferências críticas e reflexões sobre o que se está investigando.

Aos docentes do ensino superior encarregados de ministrar, em cursos de Humanas, disciplinas que requerem uma base matemático/estatística, vale lembrar que, em contraponto à figura do temido (e inútil) professor *algebrista*, há os freireanos, que acreditam que a educação é um instrumento de transformação social, um ato de liberdade e de consciência crítica. No jornalismo, inspiremo-nos neles.

PARA SABER MAIS

- **Curso de Estatística e probabilidade** - Oferecido pela Khan Academy, traz vídeos e exercícios que ensinam cálculos matemáticos e estatísticos de forma acessível. Disponível em: <https://pt.khanacademy.org/math/statistics-probability>
- **Curso de Estatística básica da Universidade Federal do Paraná** - Produzido durante a pandemia, é um curso completo desenvolvido por docentes da UFPR, com aulas gravadas e indicações de leituras. Disponível em: <http://www.leg.ufpr.br/~paulojus/estbas/>
- **Jornalismo de Dados Ambientais** - Neste curso gratuito oferecido em vídeo pela Escola de Dados, com enfoque em temática socioambiental, é possível compreender as etapas básicas de investigação em dados. Disponível em:

<https://escoladedados.org/courses/jornalismo-de-dados-ambientais/>

- **Canal do YouTube da Escola de Dados** - Cursos básicos e avançados sobre Jornalismo de Dados. Disponível em: <https://www.youtube.com/@EscoladeDados>

REFERÊNCIAS

CGI/NIC.br, Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br). **Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros: TIC Domicílios 2023**. São Paulo: Comitê Gestor da Internet no Brasil, 2023.

FRAGOSO, Wagner da Cunha. O medo da matemática. **Revista Educação**, p. 95-110, 2001.

ITO, Liliane de Lucena; VENTURA, Mauro de Souza. Interactive Multimedia Reports: Innovation, Production and Advertising Revenue. **Brazilian Journalism Research**, 12(3), 134–151, 2016. <https://doi.org/10.25200/BJR.v12n3.2016.903>

LABATE, Alice. **Geração digital' tem dificuldade ao usar computadores e enfrenta desafios no mercado de trabalho**. Link – Estadão. Disponível em: <<https://www.estadao.com.br/link/cultura-digital/geracao-digital-tem-dificuldade-ao-usar-computadores-e-enfrenta-desafios-no-mercado-de-trabalho/>>. Acesso em: 10 jun. 2025.

SCHURÉ, Eduardo. **Os grandes iniciados**. Editores Mexicanos Unidos, 1986.

TATTO, Franciele; SCAPIN, Ivone José. Matemática: por que o nível elevado de rejeição? **Revista de Ciências Humanas**, v. 5, n. 5, p. 57-70, 2004.



Parte II

Primeiros passos

A força da pergunta: o papel da hipótese em reportagens guiadas por dados

Bárbara Libório

No jornalismo guiado por dados, tudo começa com uma pergunta. Embora a coleta, limpeza, análise e visualização de bases numéricas ou textuais demande tempo e técnica, é a hipótese, ou a curiosidade transformada em caminho investigativo, que confere sentido e direção ao trabalho. Mais do que um ponto de partida, a pergunta orientadora atua como bússola editorial e cognitiva: ela determina qual fenômeno será investigado, quais métodos serão mobilizados, qual a profundidade da análise e quais os possíveis impactos da apuração para o público.

Essa dimensão é especialmente relevante em tempos de abundância informacional. Como observa Jonathan Stray (2016), o Jornalismo de Dados não começa nos dados em si, mas nas perguntas feitas a eles, que ajudam a revelar o que está sendo contado, como e por quê. Nesse novo contexto, marcado por bancos de dados públicos, ferramentas de scraping e técnicas de análise computacional, o papel da hipótese se intensifica: é ela quem separa uma mera exploração desorientada de um processo jornalístico comprometido com a produção de conhecimento.

A hipótese como ferramenta metodológica

A formulação de hipóteses para o enfrentamento de um problema é uma prática presente de forma tácita em reportagens investigativas ou aprofundadas, mesmo que nem todo jornalista a nomeie conscientemente, e jornalistas experientes em reportagens de profundidade se deparam com a confirmação ou negação de suas hipóteses (Demeneck, 2008). A hipótese é entendida como um

enunciado provisório, cuja veracidade será testada por meio da constituição de uma base de dados confiável e do trabalho de apuração.

No Jornalismo de Dados, a relação com as hipóteses pode se manifestar de duas maneiras distintas, conforme apontado por Parasie (2015):

- **Investigação orientada por hipótese (hypothesis-driven):** nessa abordagem, típica do Jornalismo de Dados tradicional, parte-se de uma suposição previamente formulada e os dados são utilizados para testá-la ou verificá-la, com o objetivo de compreender relações de causalidade. Essa estratégia se alinha a métodos do jornalismo investigativo clássico e é compatível com o modelo Story-Based Inquiry (Hunter et al., 2011), que propõe hipóteses narrativas a serem validadas ao longo da apuração.
- **Hipótese orientada por dados (data-driven hypothesis):** nesse caso, o ponto de partida são os próprios dados, frequentemente em grande volume e diversidade. Não há, inicialmente, uma hipótese clara; espera-se que padrões relevantes ou histórias inesperadas emergjam do cruzamento e análise. Essa abordagem, mais comum no jornalismo computacional, tende a gerar conhecimento baseado em correlações, cuja validade depende fortemente da completude e da qualidade dos dados analisados.

A investigação orientada por hipótese oferece maior controle narrativo e favorece investigações em áreas com forte interesse público, como políticas públicas, direitos humanos e desigualdades estruturais. Já a orientação por dados pode gerar descobertas inusitadas, mas exige maior rigor na verificação e um cuidado extra na interpretação, já que padrões não explicam por si só.

Essas duas abordagens evidenciam que o Jornalismo de Dados não é apenas uma técnica, mas uma forma de conhecimento situada, e isso implica considerar seus fundamentos epistemológicos. Produzir uma reportagem a partir de dados envolve mais do que selecionar variáveis e gerar gráficos: é um processo de interpretação que parte de

decisões humanas, escolhas políticas e recortes editoriais. Cada decisão carrega valores, perspectivas e interesses que moldam o conhecimento produzido.

“ A força de uma hipótese está na clareza da pergunta que a origina e na existência de dados disponíveis para testá-la”

A importância da pergunta

São principalmente as perguntas que guiam o que está sendo contado e como é contado. Essas perguntas são fundamentais para evitar que o repórter aceite os dados como espelhos neutros da realidade e os encare como construções sociais que têm lacunas, limites, pressupostos. A pergunta ajuda a revelar esses contornos.

O dado bruto, por si só, não responde às questões sociais mais complexas: ele precisa ser interrogado, contextualizado e, muitas vezes, cruzado com fontes qualitativas, como entrevistas, relatos e documentos. O jornalista precisa interrogar os dados tanto quanto consulta suas fontes. É nessa interrogação que se estabelece o vínculo entre informação e sentido.

Mancini e Vasconcellos (2016) defendem que o Jornalismo de Dados exige competências que vão além da técnica, como a capacidade de formular boas perguntas investigativas e interpretar os dados com senso crítico. É preciso sensibilidade para identificar quais perguntas merecem ser feitas, quem é afetado pelas dinâmicas que os números revelam (ou escondem) e quais vozes precisam ser ouvidas para além da frieza dos gráficos. Trabalhar com dados é, portanto, uma prática de escuta e interpretação. É preciso enxergar as histórias humanas por trás das tabelas.

Sabemos também que a formulação de hipóteses no jornalismo enfrenta desafios específicos, como a necessidade de traduzir problemas complexos em narrativas acessíveis, a limitação de tempo nas redações e a dependência de bases de dados muitas vezes incompletas, defasadas ou enviesadas. Esses fatores tornam ainda mais relevante o papel da pergunta como filtro e orientação.

Hipóteses fortes e fracas

Uma boa hipótese jornalística é aquela que parte de um problema real e delimitado. Por exemplo: “as escolas públicas em bairros com maior renda têm mais professores efetivos que as dos bairros de menor renda”. Essa hipótese é verificável com dados de educação e censitários. Em contraste, hipóteses vagas como “a educação está piorando” são pouco operacionais e de difícil verificação.

Hipóteses fortes geram investigações possíveis. Hipóteses fracas, por outro lado, produzem pautas difusas ou imprecisas. A força de uma hipótese está na clareza da pergunta que a origina e na existência de dados disponíveis para testá-la.

O Jornalismo de Dados tem processos que envolvem a análise estatística (como forma de testar relações entre variáveis) e a inferência, por exemplo, por meio da qual o jornalista propõe interpretações ou projeções com base nos dados disponíveis. Em ambas as fases, a hipótese cumpre papel estruturante, mas também surgem riscos, como a conclusão precipitada, a omissão de variáveis relevantes e o uso acrítico de conjuntos de dados incompletos.

A formação da base de dados com a qual a hipótese será comparada corresponde à apuração jornalística (Demeneck, 2008). É nessa etapa que se confirma, refuta ou reorienta a hipótese inicial. Às vezes, a investigação não pode ser concluída devido às limitações de tempo, acesso a fontes ou dados incompletos. Mas a hipótese segue sendo um guia, mesmo quando não totalmente testada.

Essas limitações são estruturais: o jornalismo lida com sistemas opacos, dados burocráticos e resistências institucionais. Muitas vezes, o repórter não consegue cruzar bases por incompatibilidades técnicas, ausência de padronização ou falta de transparência ativa. Ainda assim, a simples tentativa de testar uma hipótese pode gerar conhecimento relevante e provocar debates públicos.

Como formular boas perguntas

No Jornalismo de Dados, *insights* relevantes muitas vezes nascem da fricção entre dados públicos e problemas sociais urgentes.

Essa é uma característica clara em projetos como Elas no Congresso, iniciativa do Instituto AzMina que monitora a atuação parlamentar com base em dados legislativos. O ponto de partida não foi uma base de dados, mas uma inquietação: *o que de fato fazem os deputados e senadores quando o assunto são os direitos das mulheres?* Essa pergunta orientou toda a estruturação do projeto: da coleta ao cruzamento de proposições legislativas, posicionamentos, partidos e temas.

Outro exemplo é o MonitorA, também do Instituto AzMina, observatório que acompanha a violência política de gênero nas redes. A hipótese inicial, de que mulheres, especialmente negras, seriam alvos mais frequentes de ataques durante o período eleitoral, serviu como guia para definir que dados coletar, como classificar os ataques e quais categorias de violência seriam monitoradas. Nesse caso, o insight não veio dos números em si, mas da escuta ativa de candidatas e da observação de um fenômeno recorrente ainda pouco quantificado.

Esses dois projetos ilustram que boas perguntas não surgem por acaso: elas derivam da combinação entre conhecimento de contexto, observação crítica da realidade e domínio técnico sobre o que é possível extrair dos dados disponíveis. Mais do que um enunciado metodológico, a pergunta norteadora é uma ferramenta de foco: ajuda a evitar dispersão, delimita escopo e orienta decisões técnicas.

Para formular perguntas eficazes em reportagens guiadas por dados, é importante começar com um problema concreto, algo que afete grupos sociais específicos, e traduzi-lo em termos observáveis. Por exemplo, em vez de perguntar “há desigualdade na política?”, pode-se perguntar “qual o percentual de projetos com impacto positivo sobre direitos reprodutivos que foram barrados por bancadas religiosas nos últimos quatro anos?”. A diferença está na possibilidade de verificação

“ Para formular perguntas eficazes em reportagens guiadas por dados, é importante começar com um problema concreto, algo que afete grupos sociais específicos, e traduzi-lo em termos observáveis”

e na clareza da rota investigativa.

É também fundamental verificar se há dados disponíveis para sustentar a hipótese - e, se não houver, essa ausência pode se tornar um achado relevante. Muitos dos projetos desenvolvidos pela comunidade de jornalistas de dados revelaram justamente as lacunas da transparência pública: dados ausentes, mal estruturados ou invisibilizados pelo modo como são apresentados. E essa percepção, por si só, é um caminho para reportagens com alto valor público.

Em síntese, perguntar bem é mais do que iniciar uma apuração. Significa decidir o que merece ser investigado, para quem essa investigação importa, e como ela poderá produzir conhecimento que amplie o entendimento da realidade, especialmente em contextos marcados por desigualdades, invisibilização e omissões estruturais.

Por fim, vale lembrar que perguntas nem sempre têm respostas definitivas. Às vezes, o papel do jornalista de dados é justamente o de mostrar a complexidade de um tema, iluminando contradições, limites e incertezas. Ao assumir essa postura, o Jornalismo de Dados contribui para o debate público de maneira mais honesta e robusta.

REFERÊNCIAS

DEMENECK, Ben-Hur. **O uso de hipóteses no trabalho jornalístico**: a apuração se valendo de uma ferramenta metodológica. In: CONGRESSO DE CIÊNCIAS DA COMUNICAÇÃO NA REGIÃO SUL, 9., 2008, Guarapuava. Anais [...]. Guarapuava: Intercom, 2008. Disponível em: <https://www.intercom.org.br/papers/regionais/sul2008/resumos/R10-0526-1.pdf>. Acesso em: 26 maio 2025.

HUNTER, Mark Lee; et al. **Story-Based Inquiry: A Manual for Investigative Journalists**. Paris: UNESCO, 2011. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000186996>. Acesso em: 26 maio 2025.

MANCINI, Paulo Henrique; VASCONCELLOS, Fernanda. **Jornalismo de Dados**: conceito e categorias. São Paulo: Escola do Parlamento; Câmara Municipal de São Paulo, 2016.

PARASIE, Sylvain. Data-driven revelation? Epistemological tensions in investigative journalism in the age of big data. *Pre-print*, 2015. Disponível em: <https://hal.science/hal-01134296/document>. Acesso em: 26 maio 2025.

STRAY, Jonathan. **The Curious Journalist's Guide to Data**. 2016. Disponível em: <https://towcenter.gitbooks.io/curious-journalist-s-guide-to-data/content/>. Acesso em: 26 maio 2025.

Lei de Acesso à Informação e dados públicos como matéria-prima

Jamile Santana

Não é exagero dizer que a Lei Federal nº 12.527/11, a Lei de Acesso à Informação (LAI), é uma das ferramentas legais mais importantes para o exercício prático do jornalismo. Em vigor desde 2012 no Brasil, a lei transformou a relação entre o Estado e a sociedade ao estabelecer o direito de acesso a informações públicas como regra, e o sigilo como exceção. Para o Jornalismo de Dados, a LAI é uma fonte de matéria-prima. Pedidos de informação abrem caminhos para reportagens que, com base em dados públicos, expõem desigualdades, revelam gastos, monitoram políticas públicas e promovem a transparência.

Também não é exagero dizer que a LAI é subutilizada. Em 2024, a Controladoria-Geral da União recebeu apenas 53.638 pedidos de acesso à informação. Para uma população de mais de 200 milhões de pessoas, esse número pode indicar, entre outros fatores, uma baixa adesão. Além disso, menos de 1% dos pedidos foram feitos por jornalistas de forma identificável, o que pode revelar uma subutilização da lei ou outros fatores, como a insegurança destes profissionais em fazer a autodeclaração.

Em 2021, a Associação Brasileira de Jornalismo Investigativo (Abraji) publicou uma pesquisa que mostrou que praticamente metade dos jornalistas participantes (48,44%) nunca haviam feito um pedido de acesso à informação para apurar reportagens. Já entre a maioria dos jornalistas que fizeram pedidos via LAI (51,6%), encontraram dificuldades para obter respostas, principalmente do Poder Executivo, em todos os níveis (federal, estadual e municipal).

A LAI prevê dois mecanismos de acesso: a transparência ativa e a passiva. Na transparência ativa os dados devem ser publicados espontaneamente nos Portais da Transparência dos órgãos públicos. Já na transparência passiva, cidadãos ou pessoas jurídicas podem solicitar informações.

PEGA ESSA DICA!

No artigo 7º da LAI é possível consultar a lista de informações básicas de publicização obrigatória nos Portais. O rol de informações mais detalhadas e personalizadas pode ser encontrado nas leis e decretos de regulamentação. Uma outra dica é utilizar a Cartilha do Programa Nacional de Transparência Pública (PNTP) da Atricon para solicitar informações que ainda não estiverem disponíveis ativamente, mas que são recomendadas pelo órgão.

Uso de dados públicos em investigações jornalísticas

Investigar o funcionamento do poder público e suas conexões passa pela consulta a bases de dados públicas, disponíveis em Portais da Transparência ou Portais de Dados Abertos. A LAI impulsionou a divulgação de uma série de bases de dados públicas, disponíveis online, que podem ser cruzadas com outras fontes para revelar padrões e práticas de interesse público - como a nomeação cruzada de parentes de agentes políticos em gabinetes parlamentares, uso de cargos comissionados como moeda política ou indícios de fraudes em contratações, por exemplo.

Em 2019, uma investigação publicada pelo jornal O Globo, revelou como Jair Bolsonaro e seus filhos nomearam ao menos 102 pessoas com laços familiares em seus gabinetes ao longo de quase três décadas. Este é um caso emblemático do uso estratégico da LAI com cruzamento de outros dados públicos. Durante a apuração que durou cerca de três meses, a equipe solicitou informações como nomes, cargos, períodos de vínculo com a Câmara de Vereadores do Rio, da Assembleia Legislativa do Rio e Câmara dos Deputados, além de coletar informações publicadas em diários oficiais, portais da transparência e outros bancos de dados.

Bati um papo com um dos jornalistas que assinam a matéria, o Pedro Capetti. Ele me contou que quando havia indício de irregularidade, outros pedidos de acesso foram abertos solicitando informações sobre a emissão de crachás por parte dos órgãos aos funcionários investigados pela equipe. Veja que, nesse caso, a Lei de Acesso à Informação foi utilizada para complementar informações que já estavam disponíveis,

trazendo mais detalhes ao contexto da história.

O uso da LAI como ponto de partida para investigações também funciona muito bem no contexto local. Em 2013, comecei uma investigação sobre uso indevido de veículos oficiais do legislativo municipal com base nas informações orçamentárias disponíveis no portal da transparência. Na época, a Câmara de Mogi havia devolvido à prefeitura R\$ 35,7 mil de sinistro do seguro² após um acidente de trânsito com perda total fora do município. A informação estava disponível na planilha de execução orçamentária da Câmara e, por meio de um pedido de acesso à informação, solicitei a cópia do processo que deu origem ao sinistro e a devolução do recurso. Era uma linha na planilha de orçamento que rendeu uma investigação com vários desdobramentos ao longo dos anos, como o ranking de infrações e multas por parlamentar, por exemplo.

O uso indevido de veículos oficiais era uma prática comum. Para se ter uma ideia, uma investigação publicada em 2016 mostrou que placas de representação de autoridades foram colocadas na frota da Câmara em desacordo com o que prevê o Código de Trânsito Brasileiro. Isso teria sido feito para diminuir o número de autos de infração, já que o número oficial das placas ficava escondido e os veículos “invisíveis” aos radares. Só foi possível descobrir que não houve adequação às normas, a partir de pedidos de acesso à informação.

No caso da investigação do O Globo, a equipe encontrou muitas dificuldades e irregularidades. “Na Câmara de Vereadores, por exemplo, enfrentamos resistência para obter respostas. Eles não cumpriram os prazos previstos em lei e, além disso, não justificavam nem prorrogavam os pedidos. Chegamos ao ponto de pesquisar manualmente os Diários Oficiais de 2001 a 2018, durante pouco mais de uma semana, para conseguir as informações. Era papel e livro, basicamente. No final, a Câmara enviou os dados com muito atraso, mas conseguimos obtê-los”, contou o jornalista Pedro Capetti.

Transparência passiva: como pedir uma informação e aumentar suas chances de retorno

Preciso te dizer que utilizar a Lei de Acesso à Informação para investigação vai exigir (muita) paciência. Isso porque o dado solicitado

pode vir em um formato inadequado para análise, como as tabelas em PDF (por que gestores públicos adoram isso?), ou em arquivos gigantescos que exigem um conhecimento em programação e banco de dados para extrair informações relevantes. Isso, claro, quando respondem.

Mas não desanime. Fazer pedidos mais objetivos e bem estruturados, podem aumentar consideravelmente a chance de você receber uma resposta mais adequada. Vamos as boas práticas:

1. Pesquise. Antes de fazer uma solicitação de informação, verifique se o que você está pedindo, já não foi respondido. A Controladoria-Geral da União (CGU), responsável por gerenciar os pedidos de informação recebidos no âmbito do governo federal, por exemplo, publica na página “Busca de Pedidos e Respostas” boa parte das solicitações feitas aos órgãos federais, com o resultado de cada processo. Ou seja, você pode consultar como as perguntas foram feitas e o que voltou de resposta, tanto no caso de acessos concedidos, como negados. Muitas vezes, um pedido é negado por ser genérico demais, por exemplo. Muitas cidades também divulgam as perguntas e respostas.

2. Delimite o seu pedido de forma objetiva. Descreva exatamente qual informação você quer, sobre qual recurso ou programa, o órgão executor da política pública (ou vinculado à ele), período de tempo específico, formato ideal da resposta (em CSV, Excel, JSON, etc), locais onde você encontrou parte dessa informação. Evite perguntas subjetivas ou amplas demais.

3. Utilize linguagem simples e itens em tópico. Imagine que o servidor público que vai atender o seu pedido de informação, também tem outras dezenas de atribuições a cumprir no dia. Facilitar a compreensão do que está sendo pedido vai agilizar o atendimento consideravelmente. Evite o “juridiquês”.

4. Entre com recurso. Recebeu uma negativa de acesso com ausência de justificativa ou argumento inadequado? Entre com recurso! É chato, é lento, mas muitas vezes funciona. Já recebi respostas completas só depois de recorrer e citar a lei no recurso. Se a sua cidade não tem

instâncias recursais bem estabelecidas (é raro, mas acontece sempre), reclame na Ouvidoria.

5. Evite o tom combativo. Em contextos locais, muitos servidores tem boa vontade em disponibilizar informações, mas encontram dificuldades como falta de estrutura e equipe para dar conta dos prazos, ausência de processos e fluxos bem definidos na execução dos trabalhos e documentação, influência política da chefia imediata. Mas receber uma solicitação de acesso com tom acusatório pode gerar uma resistência e dificultar a resposta. Foque no dado e na informação, não no conflito.

6. Guarde o número de protocolo ou print da solicitação. Muitas cidades não possuem sistemas adequados para o recebimento de pedidos de acesso à informação. Por isso, ao fazer um pedido online, print a tela ou guarde o número de protocolo numa planilha.

PEGA ESSA DICA!

Se a informação estiver prevista em normas ou cartilhas oficiais, como a do PNTP, mencione isso no seu pedido. Mostra que você sabe o que está pedindo — e por quê.

Check-list de boas práticas para pedidos de acesso à informação

- **Evite:** “Quero saber tudo sobre o orçamento da prefeitura”.
- **Escreva:** “Solicito a planilha de execução orçamentária da Secretaria de Saúde, referente ao exercício de 2023, com detalhamento por elemento de despesa.”
- **Evite:** “Quero todos os contratos da Secretaria de Educação”
- **Escreva:** “Solicito os contratos firmados pela Secretaria de

Educação entre janeiro e dezembro de 2023, no município de Itabaiana (SE)."

- **Evite:** Pedir "tudo o que tiver".
- **Escreva:** "Solicito a lista dos contratos com empresas terceirizadas de limpeza, contendo: número do contrato, valor total, objeto, CNPJ da contratada e vigência."
- **Evite:** "Nos termos do artigo 10 da Lei Federal nº 12.527/11, venho, por meio deste requerimento formal, solicitar o encaminhamento da íntegra dos instrumentos contratuais firmados entre esta municipalidade e empresas terceirizadas, no inter-regno temporal compreendido entre 1º de janeiro de 2023 e 31 de dezembro de 2023, concernentes à prestação de serviços de asseio e conservação predial, para fins de instrução de matéria jornalística de cunho investigativo."
- **Escreva:** "Solicito os contratos firmados entre a Prefeitura e empresas terceirizadas para serviços de limpeza e conservação, entre janeiro e dezembro de 2023. Peço que, se possível, os documentos sejam enviados em formato digital editável (PDF pesquisável ou XLS), contendo o número do contrato, valor, objeto, nome e CNPJ da empresa contratada e vigência."
- **Evite:** "Quero saber por que vocês não divulgam os salários!"
- **Escreva:** "Solicito a tabela atualizada de remuneração de servidores, conforme artigo 8º, inciso III, da LAI."

Transparência ativa: onde encontrar dados que já estão disponíveis

Dados sobre a estrutura organizacional, receitas, despesas, licitações, contratos, remuneração de servidores, cumprimento de metas e programas, são alguns exemplos de informações que já devem estar disponíveis nos Portais da Transparência de forma ativa, ou seja, sem a necessidade do órgão ser questionado ou provocado a publicar.

Além dos Portais da Transparência, os Portais de Dados Aber-

tos são fundamentais para o acesso a dados mais detalhados, com conjuntos e temas que ultrapassam os previstos na LAI. No Portal de Dados Abertos da Prefeitura de São Paulo, por exemplo, há uma base sobre as favelas do município, áreas contaminadas, microdados da rede municipal de educação, entre outros. A WikiLAI, desenvolvida pela Fiquem Sabendo, agência de dados independente e especializada em Lei de Acesso à Informação, também tem dicas essenciais sobre o cumprimento da lei e formas de solicitar informações.

PEGA ESSA DICA!

O Índice de Dados Abertos para cidades, da Open Knowledge Brasil, traz uma gama de possibilidades de abertura de dados, a partir de 111 conjuntos distribuídos em 14 dimensões temáticas. Ele pode ser uma grande inspiração para embasar seus pedidos de acesso à informação.

REFERÊNCIAS

BRASIL. **Lei nº 12.527, de 18 de novembro de 2011.** Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 17 jun. 2025.

BRASIL. Controladoria-Geral da União. **Painel da Lei de Acesso à Informação – Painel LAI.** Disponível em: <https://centralpaineis.cgu.gov.br/visualizar/lai>. Acesso em: 18 maio 2025.

DADOS ABERTOS – Prefeitura de São Paulo. **Educação.** Disponível em: <http://dados.prefeitura.sp.gov.br/group/educacao>. Acesso em: 17 jun. 2025.

FIQUEM SABENDO. **Entrevista com Francisco Leali**. Disponível em: <https://fiquemsabendo.com.br/transparencia/entrevista-francisco-leali>. Acesso em: 17 jun. 2025.

G1. **Carro da Câmara de Mogi tem perda total e sindicância conclui imperícia**. G1, 2013. Disponível em: <https://g1.globo.com/sp/mogi-das-cruzes-suzano/noticia/2013/12/carro-da-camara-de-mogi-tem-perda-total-e-sindican-cia-conclui-impericia.html>. Acesso em: 17 jun. 2025.

GLOBO. **Em 28 anos, clã Bolsonaro nomeou 102 pessoas com laços familiares**. O Globo, 2020. Disponível em: <https://oglobo.globo.com/politica/em-28-anos-cla-bolsonaro-nomeou-102-pessoas-com-lacos-familiares-23837445>. Acesso em: 17 jun. 2025.

PORTAL DA TRANSPARÊNCIA. **Radar da Transparência Pública**. Disponível em: <https://radardatransparencia.atricon.org.br/>. Acesso em: 17 jun. 2025.

PORTAL GOV.BR. **Busca de pedidos e respostas – Lei de Acesso à Informação**. Disponível em: <https://www.gov.br/acessoainformacao/pt-br/assuntos/busca-de-pedidos-e-respostas>. Acesso em: 17 jun. 2025.

ASSOCIAÇÃO BRASILEIRA DE JORNALISMO INVESTIGATIVO (ABRAJI). **Uso da Lei de Acesso a Informações por jornalistas: 4º relatório de desempenho**. Disponível em: <https://www.abraji.org.br/noticias/pesquisa-aponta-alto-indice-de-jornalistas-que-nunca-fizeram-pedidos-de-acesso-a-informacao>. Acesso em: 17 jun. 2025.

OPEN KNOWLEDGE BRASIL. **Índice de Dados Abertos para Cidades**. Disponível em: <https://indicedadosabertos.ok.org.br/>. Acesso em: 17 jun. 2025.

FIQUEM SABENDO (Agência de dados). **WikiLAI – Página principal**. Disponível em: https://wikilai.fiquemsabendo.com.br/wiki/P%C3%A1gina_principal. Acesso em: 17 jun. 2025. (Projeto desenvolvido com financiamento da Embaixada e Consulados dos EUA no Brasil; lançado em nov. 2021)

Geração Cidadã de Dados para criar bases originais para reportagens de impacto social

Polinho Mota

Neste capítulo vou apresentar a Geração Cidadã de Dados (GCD) e como ela pode ser útil ao seu trabalho no jornalismo. Vou caminhar junto com você pela definição do método, as boas práticas para aplicação e como ela tem sido útil no meu trabalho para, entre muitas coisas, gerar informações de qualidade para criar ou apoiar narrativas das periferias, favelas e comunidades urbanas.

Pensa comigo: quantas vezes a gente busca dados oficiais sobre um problema urgente na nossa comunidade e não encontra nada? Ou o que achamos é tão genérico que não reflete a real situação? Esse vazio é muito comum, principalmente, se você trabalha com pautas ou pessoas marginalizadas. Ele é um dos reflexos da nossa sociedade excludente, racista e conservadora que faz com que um grande número de pessoas tenham seus direitos fundamentais ou básicos negados cotidianamente. Trazendo um exemplo prático, eu poderia citar os dados de raça/cor das pessoas internadas por causas psiquiátricas no Rio de Janeiro.

Em 2022, o data_labe produziu uma reportagem intitulada “Cemitério dos vivos” (ROSA G, data_labe, 2022). Nesse trabalho, analisamos 10 anos de dados do Sistema de Informações Hospitalares do DataSus e descobrimos que 39,7% das pessoas internadas por causas psiquiátricas não tinham o campo de raça/cor preenchido. Esse parece ser um detalhe, mas na verdade é um reflexo de como pessoas que sofrem com doenças psíquicas são desumanizadas e vistas como “não gente”, e portanto não importa tanto preencher a raça desses indivíduos. Mais chocante ainda foi descobrir em uma visita de campo à Clínica de Saúde Juliano Moreira, onde também abriga um museu sobre saúde mental, que no início dos trabalhos no século XIX os pacientes eram internados sem preenchimento algum em suas fichas, com indicações

como “morena violenta”, “indivíduo sem nome” e era assim que eram nomeados oficialmente pro restante do seu tempo de internação.

Ou seja, a ausência do dado é muito mais um sintoma do apagamento social, política de encarceramento, ausência de cuidado especializado do que uma simples falta de *input* de uma informação. Pior ainda é pensar que certas temáticas não tem dado algum, como por exemplo informações das abordagens policiais no Brasil. Esses exemplos servem para ilustrar que alguns temas e grupos sociais vão ter poucos dados disponíveis e às vezes nenhum dado coletado. É por isso que a GCD pode ser um caminho bom para trazer luz a questões desse tipo. Foi o aconteceu com a série de reportagens do projeto [#PorQueEu?](#), que gerou dados de forma cidadã sobre racismo em abordagens policiais. Para além de reportagens, o Instituto de Defesa do Direito de Defesa usou os dados desta pesquisa para sustentar no Supremo Tribunal Federal sobre a ilegalidade de processos criminais baseados em abordagens policiais racistas.

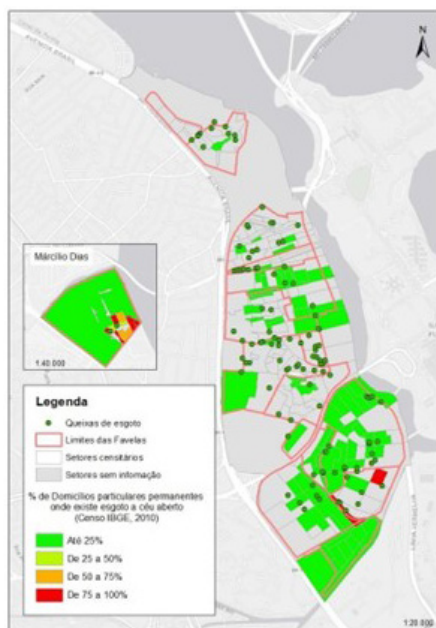
Mas, afinal, o que é essa tal de Geração Cidadã de Dados?

Junto com a equipe do data_labe, desenvolvi uma definição que tem servido aos nossos trabalhos internos e também para produções acadêmicas sobre o tema: **GCD é o conjunto de ações que permite que os próprios cidadãos produzam, colem e usem dados para o bem de suas comunidades.**

Eu quero destacar dois eixos fundamentais para essa definição: os trabalhos não são apenas de gerar o dado, mas sim em todas as atividades que permeiam o engajamento cívico através desta ação; e é sobre o benefício próprio da comunidade. Talvez este segundo eixo seja o mais específico da GCD, que denota um protagonismo e autonomia das comunidades para a tomada de decisão e também execução das ações. Portanto, é uma mudança significativa em processos onde o cidadão apenas é subutilizado para gerar ou coletar dados para outro onde ele pauta, decide e comunica ações sobre os dados que lhe interessa. Ou seja, não são somente os institutos de pesquisa que podem gerar informação relevante. As próprias pessoas que vivem as suas realidades têm o poder de levantar dados importantíssimos para entender e transformar o seu território.

Trazendo a GCD para o nosso mundo: o jornalismo de impacto social

Algumas pessoas acham a ideia e o processo revolucionário, mas na verdade temos alguns exemplos de aplicação desse método, como num estudo que comparou dados de chuva gerados por cidadãos no Nepal com os dados oficiais — e a GCD teve um erro menor (Davids *et.al*, 2019)! Isso mostra o potencial e a precisão que a informação produzida localmente pode ter. Localizando mais ainda os exemplos, posso trazer um outro projeto do data_labe chamado Cocôzap¹, que mapeou irregularidades de saneamento básico na Maré.



Comparação entre os dados gerados pelo Cocôzap entre os disponibilizados pela IBGE sobre esgoto a céu aberto

(Fonte: DataLabe)

¹ <https://datalabe.org/cocozap/>.



Na imagem anterior, podemos ver o mapa de toda a favela da Maré, um conjunto de 16 favelas na região Norte do Rio de Janeiro. No mapa, cada pontinho verde representa uma queixa enviada pelos moradores sobre esgoto a céu aberto, enquanto no fundo do mapa, os polígonos representam os setores censitários coloridos de acordo com a legenda em maior proporção de domicílios que possuem esgoto a céu aberto em seu entorno. A avaliação do mapa informa que em muitos setores onde sequer havia informação é justamente onde tem mais queixas do Cocôzap.

Mas existem outros exemplos pra além do Cocôzap, como as bases de dados sobre tiroteio produzidas pelo Fogo Cruzado², o estudo chamado Território Inventivo³ do Observatório de Favelas e outros que começam a se espalhar pelo país e que, em geral, estão engajados na Rede GCD, um conjunto de organizações que se reconhecem como “fazedoras” dessa metodologia.

Como você pôde ler até aqui, os exemplos de dados gerados pelos cidadãos são em geral de pauta de profundo interesse social, por isso que a GCD possui alto apelo do Jornalismo, pois pode fornecer dados exclusivos, sensíveis e gerados com extremo cuidado sobre situações onde não se possui muitas informações quantitativas. Essa pode ser a matéria-prima para reportagens que impactam, que podem dar plataforma a comunidades excluídas e pressionar por mudanças concretas.

Como o jornalista pode botar a GCD em prática?

Existem duas maneiras que eu indico para você se utilizar da GCD. A primeira e mais óbvia é a de usar os dados gerados para suas reportagens. Em geral, os dados são abertos e disponibilizados na Web, o que garante que você execute suas próprias análises, se quiser, ou apenas extraia as informações disponibilizadas pelas organizações geradoras dos dados. A segunda forma é desenvolvendo seus próprios projetos GCD.

2 <https://fogocruzado.org.br/>.

3 <https://observatoriodefavelas.org.br/territorio-inventivo/>.



Os sete passos para realizar seu próprio projeto GCD

Eu tenho apresentado os passos para realizar a GCD com a imagem de uma cabaça, que é um instrumento muito utilizado pelos povos indígenas e de matriz africana. A cabaça tem um formato que lembra o símbolo do infinito e ilustra a minha intenção de te dizer que, apesar de eu apresentar pedagogicamente a partir da ação de identificar, nada impede que você comece pela ação de estruturar. Porém, mais importante que você saber a ordem é entender o que cada ação significa. Você pode consultar o material completo para isso no site do data_labe (MOTA P, VIEIRA G, 2023), mas para te dar uma visão reduzida eu oriento uma pergunta disparadora para cada ação da cabaça GCD.

Os sete passos da geração cidadã de dados:

1. **Identificar:** Qual problema, assunto, experiência ou realidade precisa ser evidenciada através de dados?
2. **Selecionar:** Quais são as causas que melhor poderiam explicar o problema selecionado por você e seu grupo?

3. **Discutir:** Quais outros subtópicos do problema você não irá abordar e o quanto isso afeta o seu trabalho?
4. **Selecionar:** Existem outras bases de dados que já abordam o mesmo problema que você escolheu?
5. **Engajar:** Quem irá te ajudar e participar do processo da coleta e geração de dados
6. **Estruturar:** Quais processos e sistematizações serão necessárias para gerar dados de forma computacional?
7. **Comunicar:** Quais estratégias você e seu grupo criarão para comunicar os resultados do trabalho?

Indico ainda você ler o manifesto e as iniciativas da Rede GCD para possuir uma visão mais ampliada sobre formas de atuação e princípios norteadores do trabalho.

A tecnologia simples é a que mais funciona

A GCD é um caminho para construir bases de dados originais, com a cara e a voz de quem vive a realidade. No jornalismo, isso significa reportagens mais profundas, mais relevantes e com maior potencial de transformação social. Tudo isso graças a um pensamento simples, de que quem mais sabe sobre o problema é quem está envolvido nele.

Portanto, a GCD luta contra práticas colonizadoras de saberes, onde a população é vista como não sábia, ou incapaz de pensar suas soluções. Além disso, impacta com mais dados reais e sensíveis sobre problemas marginalizados, mas promove um melhor e maior engajamento sobre problemas da sociedade.

REFERÊNCIAS

Davids JC, Devkota N, Pandley A, Prajapati R, Ertis BA et.al. **Soda Bottle Science** - Citizen Science Monsoon Precipitation Monitoring in Nepal. Front. Earth Sci. Volume 7, 2019. Cited March 31. Available from: <https://www.frontiersin.org/journals/earth-science/articles/10.3389/feart.2019.00046/full>

ROSA, G. **Cemitério dos vivos**. Data_labe.2022. Citado em 15 de Abril. Disponível em : <https://datalabe.org/cemiterio-dos-vivos/>

MOTA P, VIEIRA G. **Geração Cidadã de Dados**. Saiba como desenvolver seu projeto de produção de dados com participação social a partir da metodologia utilizada pelo data_labe. 2023. Citado em 15 de Abril. Disponível em: <https://datalabe.org/geracao-cidada-de-dados/>

OSINT no jornalismo: indo mais fundo com fontes abertas

Adriano Belisário

Antiga companheira da prática jornalística, a utilização de fontes abertas em investigações ganhou crescente atenção nas últimas décadas em áreas diversas. De organizações de direitos humanos a forças policiais e militares, passando por empresas de cibersegurança, grandes corporações internacionais e coletivos da sociedade civil, foram muitos os grupos que ampliaram sua capacidade de conduzir investigações graças à digitalização da comunicação e o fácil acesso a conteúdos online gerados por usuários. Ainda que tenham objetivos distintos, os métodos de investigação com fontes abertas empregados por estes grupos muitas vezes são reaproveitados em outros contextos. Nesse sentido, este capítulo lança uma mirada sobre estas abordagens para sintetizar aprendizados práticos para jornalistas e estudantes.

O atalho mais direto para mergulhar no manancial de recursos e ferramentas nesta área encontra-se atrás de cinco letras: OSINT. O acrônimo estrangeiro - *Open Source INTElligence* ou Inteligência com Fontes Abertas - é amplamente utilizado nos bastidores do jornalismo, sendo inclusive mencionado como habilidade desejada em muitas vagas. Contudo, como o nome indica, as práticas de OSINT possuem origem militar. O conceito de “inteligência” na sigla OSINT foi empregado, ao menos inicialmente, no mesmo sentido adotado pelas agências estatais de inteligência, como a ABIN e a CIA.

Por conta deste histórico, há quem prefira utilizar o conceito de “investigações com fontes abertas” (*open source investigations*) para se referir à aplicação destes métodos investigativos para fins jornalísticos ou de defesa dos direitos humanos. É o caso do Protocolo de Berkeley (Ohchir, 2022), que define princípios para o uso de fontes abertas em investigações de direitos humanos. Elaborado em parceria com a ONU, o documento define fontes abertas como informações publicamente disponíveis (não necessariamente gratuitas), ou seja, informações que

qualquer pessoa pode observar, comprar ou solicitar, sem recorrer a privilégios legais específicos ou acessos não autorizados.

Há também quem use o acrônimo OSINT para se referir às investigações com fontes abertas feitas pela sociedade civil. É com este sentido que Fuller e Weizman (2021, p. 5) mencionam a sigla, para referir-se àqueles que “buscam vestígios de políticas que são invisíveis, secretas ou negadas, e trabalham para produzir fatos que contestam declarações ou outras autoridades”. Seguindo esta linha, mesmo ciente das nuances terminológicas, adoto neste artigo a sigla OSINT para me referir ao campo de investigações com fontes abertas no jornalismo.

Apesar do foco instrumental dado por alguns treinamentos na área, dominar as práticas de OSINT não se resume ao aprendizado de ferramentas. Elas são necessárias, porém, muitas vezes, possuem um prazo de validade curto. Ao contrário, a aplicação de técnicas de OSINT no jornalismo demanda antes e sobretudo um certo olhar e uma certa postura, ou seja, um *modus operandi* ou “estado de espírito” (Dutch OSINT Guy, 2018).

OSINT como *modus operandi*

Toda investigação começa com um ou mais elementos. São as primeiras peças de um quebra-cabeça que é preciso montar, por exemplo: nome de pessoas, um endereço, um site ou uma empresa. As técnicas e métodos de OSINT nos permitem derivar novas informações a partir destes elementos iniciais (um endereço a partir de um CNPJ, para usar um exemplo simples) e concatená-los para “ligar os pontos” e produzir conhecimento relevante para um caso específico. Cada novo elemento derivado pode ser um novo ponto de partida para a investigação, de modo que a investigação com OSINT é muitas vezes descrita como um ciclo que se repete.

Abaixo, à esquerda, a Figura 1 mostra estas etapas deste ciclo investigativo e, na Figura 2, há um exemplo de como um elemento (no caso, um nome de usuário) pode levar a outros elementos (redes sociais), cada uma delas representando um novo caminho possível para a investigação. Dentre todas etapas do ciclo, por limitações de espaço, este capítulo se debruça especialmente na primeira etapa, relacionada às buscas online.



De início, para incorporar as práticas de OSINT no jornalismo, é preciso ter um olhar aguçado, detalhista e atento não só ao que é dito explicitamente, mas também a estas informações derivadas, que não são dadas inicialmente. Colocar os “óculos de OSINT” permite ao jornalista ver possibilidades de investigação em quase todo elemento. De fato, o aprendizado de técnicas de busca avançada online é frequentemente o ponto de partida dos currículos de formação em OSINT para jornalistas. Mas não basta “dar um Google” nas informações de interesse. É preciso dominar os operadores de busca avançada da plataforma (*Google dorking*) e de outros buscadores e redes sociais. Mais que isso, é preciso fazer a busca de forma sistemática, minuciosa e quase obsessiva, esgotando cada caminho e cada detalhe possível.

Este *modus operandi* investigativo não se restringe a informações textuais, abrangendo também imagens e outros tipos de dados. Por exemplo, uma foto ou vídeo publicado online apresenta múltiplos caminhos de entrada para derivar novos elementos: cartazes, placas

de carro ou a vegetação do local podem responder quando e onde ele foi registrado. A reportagem 'Anatomy of a killing' (BBC, 2019) é um exemplo icônico, pois a partir de um vídeo de uma execução consegue responder perguntas fundamentais em uma investigação, como onde e quando o crime ocorreu, além da própria autoria da execução - tudo isso usando recursos abertos, como imagens de satélite e redes sociais¹. A investigação combina diferentes técnicas de OSINT e demonstra todo potencial desta abordagem para extrair o máximo de informações a partir de imagens. De fato, há tantas possibilidades nesta área que o New York Times criou em 2017 uma unidade dedicada a "investigações visuais"² para aplicar no jornalismo OSINT e outras metodologias investigativas baseadas em imagens.

Ou seja, os métodos de OSINT são ecléticos e envolvem desde técnicas relativamente simples, como a busca reversa de imagens, a busca de imagens de satélite e a geo e cronolocalização, que são técnicas para derivar o local e o momento em que uma imagem foi registrada a partir de informações visuais (veja o tutorial em português da Escola de Dados nas referências finais para saber mais), até abordagens mais sofisticadas, que podem demandar conhecimentos especializados, como a modelagem 3D e a fotogrametria para reconstituir eventos ou oferecer visualizações imersivas. No projeto 'Territórios de Exceção', por exemplo, a investigação se debruçou sobre o uso de helicópteros policiais como plataforma de tiro no Rio de Janeiro. Para isso, além de coletar imagens registradas por moradores nas redes sociais, foi feita uma parceria com a Redes da Maré, uma organização da sociedade civil que tem um vasto acervo documental³ da ação policial no Complexo da Maré, a fim de geolocalizar as imagens e utilizar modelos 3D para evidenciar como os disparos foram feitos próximo de escolas, por exemplo. Entre outras aplicações, imagens de satélites também já foram utilizadas em uma investigação jornalística sobre mineração ilegal na Amazônia (Andreoni et al, 2022).

1 Esta reportagem em português detalha os passos da investigação: <https://www.bbc.com/portuguese/internacional-45650800>.

2 <https://www.nytimes.com/spotlight/visual-investigations>.

3 <https://documental.xyz/intervencao>.

Para além de textos e imagens, OSINT também envolvem outras fontes e abordagens. A análise de redes entre indivíduos pode ser usada para demonstrar não só relações em redes sociais digitais, mas também interesses económicos. Em 2017, a documentação sistemática das sociedades empresariais dos donos de ônibus no Rio de Janeiro (Belisario, 2017) jogou luz a fatos que explorei em diversas reportagens para o 'Especial Catraca' da Agência Pública, como possíveis táticas para mascarar sobrelucro⁵ e a sociedade entre empresários e o deputado que presidia a Comissão de Transporte da Assembleia local⁶. Outra abordagem é a análise de metadados de arquivos, que já serviu para demonstrar como projetos de lei sobre mineração foram escritos em computadores de advogados de mineadoras (Senra, 2015).

Felizmente, não é preciso decorar todas as ferramentas e abordagens disponíveis, já que existem diversas compilações com os principais recursos para OSINT. Dentre elas, destacam-se coleções colaborativas, como a caixa de ferramentas organizada pelo Bellingcat⁷ e o site OSINT Framework⁸. Ambos organizam tematicamente diferentes recursos online e merecem ser visitados por jornalistas aspirantes nesta área. Outro recurso interessante é o kit 'Exposing the

“ Cada novo elemento derivado pode ser um novo ponto de partida para a investigação, de modo que a investigação com OSINT é muitas vezes descrita como um ciclo que se repete ”

4 <https://apublica.org/especial/especial-catraca/>.

5 <https://apublica.org/2017/08/imobiliarias-podem-mascarar-sobre-lucro-de-empresarios-de-onibus/>.

6 <https://apublica.org/2017/07/presidente-da-comissao-de-transporte-da-assembleia-do-rio-e-socio-de-diretor-da-fetranspor-foragido/>.

7 <https://bellincat.gitbook.io/toolkit>.

8 <https://osintframework.com/>.

Invisible⁹ do grupo Tactical Tech. Tanto o Bellingcat quanto o Tactical Tech possuem também tutoriais em seus sites, com orientações de uso das ferramentas.

Todas estas possibilidades de busca de informação empolgam, mas também podem ser intimidadoras e exaustivas, principalmente porque boa parte destes caminhos investigativos não levam a lugar nenhum. A paciência e persistência são, portanto, fundamentais no *modus operandi* OSINT. Muitas vezes, encontrar uma informação útil é como procurar uma agulha no palheiro. Além disso, é preciso também disciplina e método para documentar e sistematizar os elementos encontrados. Às vezes, uma informação por si só pode não ser útil, somente ganhando relevância quando considerada em conjunto com outras informações. Devido ao grande volume de informações acesadas e revisadas durante uma investigação, sem uma documentação adequada, estas conexões podem não ser visíveis.

Idealmente, a documentação adequada tem um duplo propósito na adoção do *modus operandi* OSINT no jornalismo. Além de apoiar análises complexas, as etapas de coleta e preservação do ciclo (Figura 1) também são necessárias para fins de transparência e reprodutibilidade. Ou seja, investigações abertas apoiam suas conclusões em etapas transparentes e verificáveis, permitindo que virtualmente qualquer pessoa possa refazer o caminho que conecta uma informação a outra. Por isso, o arquivamento é uma prática crucial. Alguém que usa OSINT profissionalmente, inclusive no jornalismo, é um *serial archivist*, um arquivista serial, preservando evidências que podem não estar mais disponíveis quando forem mais necessárias. Neste sentido, é importante manter uma cadeia de custódia sobre as evidências digitais coletadas, buscando garantir sua autenticidade e evitando qualquer ação que possa alterar os arquivos originais.

O *modus operandi* OSINT demanda um certo nível de letramento digital, que varia com os interesses e necessidades de cada um. Não é preciso largar o jornalismo para estudar computação, porém, noções básicas sobre como funcionam computadores e a Internet são fundamentais, até mesmo para conseguir interpretar criticamente as

9 <https://kit.exposingtheinvisible.org/en/>.

fontes abertas online. Além disso, conseguir automatizar a coleta de dados online (raspagem ou *web scraping*) ou aplicar métodos de análise de texto (NLP ou, em português, processamento de linguagem natural) abrem um novo horizonte de possibilidades nas investigações.

Por fim, o uso de OSINT no jornalismo envolve também questões não menos importantes, como a dimensão ética (Van der Woude; Dodds; Torres, 2024), a segurança digital e a proteção à privacidade online. A avaliação de risco é pré-requisito fundamental para o jornalismo investigativo. Um descuido, como acessar o perfil de uma pessoa investigada no LinkedIn usando a própria conta, pode ser suficiente inviabilizar uma investigação em curso ou até mesmo gerar riscos às pessoas envolvidas. Usando ou não técnicas de OSINT, a reflexão sobre a segurança digital no jornalismo precisa ser feita de maneira holística, levando em conta as dimensões de segurança física, digital e psicológica de forma conectada. Para maior aprofundamento no tema, veja o projeto 'Holistic Security'¹⁰ do Tactical Tech.

“Alguém que usa OSINT profissionalmente, inclusive no jornalismo, é um serial archivist, um arquivista serial, preservando evidências que podem não estar mais disponíveis quando forem mais necessárias”

OSINT como paradigma colaborativo

Como apontado inicialmente, o uso de fontes abertas no jornalismo não é nenhuma novidade. Algumas das etapas do ciclo de investigação em OSINT dizem respeito a velhas conhecidas práticas jornalísticas, como a própria análise investigativa e a verificação, que avalia a credibilidade das informações encontradas e faz a checagem cruzada com outras fontes. Ainda assim, há metodologias

¹⁰ <https://holistic-security.tacticaltech.org/>.

investigativas e abordagens na caixa de ferramentas OSINT que são extremamente úteis, mas não são óbvias e nem fazem parte da formação tradicional em jornalismo. Deste modo, o jornalismo vem tirando proveito da aproximação com o universo OSINT para potencializar a busca de informações e diversificar as técnicas investigativas em ambientes online

Curiosamente, a adesão das práticas de OSINT no jornalismo se deu, ao menos parcialmente, graças a grupos que não se enquadram nos formatos tradicionais de "organizações de imprensa". Apesar de não funcionarem como veículos de imprensa, grupos como Bellingcat e o Forensic Architecture¹¹ produzem investigações de inegável valor de notícia, tanto que frequentemente fazem trabalhos conjuntos com veículos de mídia. O fato de serem utilizadas fontes abertas, com uma metodologia transparente e reproduzível, torna as práticas de OSINT como um paradigma propício à colaboração. Ou seja, uma das contribuições mais significativas deste *modus operandi* para as investigações jornalísticas se dá não só por meio da adoção de técnicas e ferramentas, mas também da parceria entre jornalistas, organizações de direitos humanos e especialistas de outras áreas para tratar de temas sensíveis e complexos.

REFERÊNCIAS

ANDREONI, Manuela et al. **The Illegal Airstrips Bringing Toxic Mining to Brazil's Indigenous Land**. The New York Times, 2 ago. 2022. Disponível em: <https://www.nytimes.com/interactive/2022/08/02/world/americas/brazil-airstrips-illegal-mining.html>. Acesso em: 16 maio 2025.

BBC. **Anatomy of a Killing**. BBC World Service TV – Africa Eye, 07 fev. 2019. Disponível em: <https://www.bbc.co.uk/programmes/p0707w39>. Acesso em: 16 maio 2025.

¹¹ <https://forensic-architecture.org/>.

BELISARIO, Adriano. **Checagem de imagens: cronolocalização de fotos.** 2022. Disponível em: <https://escoladedados.org/tutoriais/cheragem-de-imagens-cronolocalizacao-de-fotos/>. Acesso em: 16 maio 2025.

BELISARIO, Adriano. **A teia dos donos do transporte no Rio.** Agência Pública, 8 ago. 2017. Disponível em: <https://apublica.org/2017/08/a-teia-dos-donos-do-transporte-no-rio/>. Acesso em: 16 maio 2025.

DUTCH OSINT GUY. **OSINT as a mindset.** Medium, 14 jan. 2018. Disponível em: <https://medium.com/@Dutchosintguy/osint-as-a-mindset-7d42ad72113d>. Acesso em: 16 maio 2025.

FULLER, Matthew; WEIZMAN, Eyal. **Investigative aesthetics: conflicts and commons in the politics of truth.** London: Verso, 2021.

OHCHR (UNITED NATIONS FOR HUMAN RIGHTS OFFICE OF THE HIGH COMMISSIONER). **Berkeley Protocol on Digital Open Source Investigations: a practical guide on the effective use of digital open source information in investigating violations of international criminal, human rights and humanitarian law.** New York: United Nations, 2022.

SENRA, Ricardo. **Novo código da mineração é escrito em computador de advogado de mineradoras.** BBC News Brasil, 7 dez. 2015. Disponível em: https://www.bbc.com/portuguese/noticias/2015/12/151202_escritorio_mineradoras_codigo_mineracao_rs. Acesso em: 16 maio 2025.

VAN DER WOUDE, Maartje; DODDS, Thomas; TORRES, Gabriela. **The ethics of open source investigations: navigating privacy challenges in a gray zone information landscape.** Journalism, [S. l.], v. 0, n. 0, 2024. Disponível em: <https://doi.org/10.1177/14648849241274104>. Acesso em: 16 maio 2025.

A ética profissional no Jornalismo de Dados

Samuel Pantoja Lima

Vigente há quase duas décadas, posto que aprovado no começo de agosto de 2007, o Código de Ética dos Jornalistas Brasileiros foi concebido no cenário de transição profissional do ecossistema analógico para o digital sob o impacto do advento das novas tecnologias e da *internet*. As técnicas do chamado Jornalismo de Dados (JD) ou Jornalismo Guiado por Dados (JGD) começavam a se popularizar, no país, a partir das iniciativas de cursos de Reportagem Assistida por Computador (RAC), especialmente em função de cursos promovidos pela Associação Brasileira de Jornalismo Investigativo (Abraji), fundada em 2003.

Nesse ínterim, o jornalismo profissional passou a lidar com duas novidades que transformaram a apuração em base de dados digitais para melhor, no caso da Lei de Acesso à Informação Pública (LAI), Lei nº 12.527/2011 - Lei Nº 12.527, de 18 de novembro de 2011 - e a Lei Geral de Proteção de Dados Pessoais (LGPD) - Lei Nº 13.709, de 14 de agosto de 2018. A primeira (LAI), regulamenta o direito de qualquer pessoa solicitar e receber informações dos órgãos e entidades públicas, considerada como instrumento jurídico fundamental visando garantir a transparência e o controle social. Em tese, permite que cidadãos e cidadãs exerçam o direito de solicitar e receber informações dos órgãos públicos, contribuindo para uma sociedade mais democrática e participativa. A segunda (LGPD), conforme expresso em seu Artigo 1º, "dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural".

Na prática do Jornalismo de Dados (JD), as fontes primárias são os próprios dados, em geral acessados em bases públicas disponíveis na internet, seja por determinação legal sustentado pela

legislação da transparência dos entes públicos, seja por boas práticas de entes privados ou do Terceiro Setor, que buscam essa mesma relação de transparência com seus públicos. Portanto, as questões de natureza ética estão igualmente presentes na prática do JD como numa apuração convencional. Na vida real, os dados ou cada linha de informação de uma determinada planilha (Excel, CSV, Google Sheets etc.) identificam determinadas pessoas e/ou organizações. Portanto, as recomendações emanadas do Código de Ética dos jornalistas se aplicam como tal. O cenário se complexifica, a partir da adoção da LGPD, que tem sido usada como subterfúgio para negar o direito à informação amparado pela Lei de Acesso (LAI).

Jornalismo de Dados, LAI e LGPD: separando o “joio” do “trigo”

Tomamos como horizonte maior de garantia à liberdade de imprensa e de manifestação a Constituição Federal (CF, 1988), que em tese afastaria o fantasma da censura das relações sociais. Literalmente, o Artigo 220 prevê que “a manifestação do pensamento, a criação, a expressão e a informação, sob qualquer forma, processo ou veículo não sofrerão qualquer restrição, observado o disposto nesta Constituição”. O legislador constituinte registrou com ênfase, em parágrafo complementar, que “nenhuma lei conterà dispositivo que possa constituir embaraço à plena liberdade de informação jornalística em qualquer veículo de comunicação social, observado o disposto no art. 5º, IV, V, X, XIII e XIV”.

Este é o horizonte ético e político que orienta a prática jornalística e no Jornalismo de Dados tal pressuposto está igualmente presente. O acesso à informação pública nas bases de dados de entes federais, estaduais e municipais não é trivial no labor cotidiano da profissão – especialmente nas 5.569 prefeituras e câmaras municipais, que em sua maioria ainda não cumprem plenamente a Lei nº 12.527/2011. E como o pedido de LAI acaba sendo um procedimento de apuração muito frequente também no Jornalismo de Dados, quando esgotados todos os recursos à obtenção direta dos dados das fontes públicas, a adoção da LGPD, em 2018, trouxe novos desafios à prática jornalística, em geral.

Em outras palavras, quais embaraços mais comuns à prática

do Jornalismo de Dados – e ao labor jornalístico, em geral – que a Lei Geral de Proteção de Dados trouxe? O advogado Marcos Antônio Silveira, especialista no assunto, faz algumas recomendações sobre o suposto impacto da LGPD nas rotinas produtivas de notícias, fotografias, captação de imagens, posts em redes sociais etc. No caso de fotografias e vídeos, por exemplo, mesmo captadas em meios virtuais, Silveira

lembra que quaisquer imagens que identifiquem pessoas devem ser feitas com o consentimento livre e esclarecido do titular, **exceto em situações de interesse público inconteste ou jornalístico**. Num exemplo prático, segundo o advogado, numa cobertura de manifestação pública, você deve priorizar fotos ou imagens da multidão sem focar em rostos específicos: “Se precisar destacar uma pessoa, obtenha seu consentimento ou utilize técnicas de desfoque para preservar a identidade”.

Retomo o grifo para acrescentar que, numa cobertura esportiva da partida final de tênis no Torneio de Roland Garros (em 08 junho de 2025), a quadra central do complexo onde o espanhol Carlos Alcaraz ganhou o título de seu adversário Jannik Sinner, italiano, estava lotada de figuras públicas, que foram fotografadas e filmadas pela imprensa esportiva que cobria o evento. Ora, pedir “autorização” a cada uma delas inviabilizaria a cobertura e a relevância que tais personalidades conferem ao conteúdo jornalístico (atores, atrizes, artistas, esportistas conhecidos etc.).

Retomando a questão da prática do Jornalismo de Dados, o que tem sido frequente é agentes políticos, ao serem flagrados em práticas lesivas ao interesse público (corrupção, prevaricação e malversação de distintas ordens), saírem gritando que seus nomes não podem ser publicados porque supostamente estariam protegidos pela LGPD. Os dados coletados pelo jornalista e pesquisador Fran-

“ A questão mais sensível e complexa que se apresenta ao Jornalismo de Dados é, do ponto de vista ético, equacionar a necessidade de publicação de informações de inconteste interesse público com o direito à privacidade”

cisco Gonçalves, recém doutor pela Faculdade de Comunicação da Universidade de Brasília (UnB), estudando o assunto, revelam que em 10 anos, entre maio de 2012 (ano de regulamentação da LAI) e maio de 2022, os jornalistas brasileiros fizeram 31.826 pedidos pela LAI. Desse total, cerca de dois terços (67%) foram concedidos; o restante foi atendido parcialmente ou simplesmente negado. No caso das negativas, além dos graus de sigilo atribuídos à informação pública (reservada – acesso restrito até 5 anos; secreta – restrição por até 15 anos; e ultrassecreta – acesso por até 25 anos, podendo ainda ser prorrogado por uma vez), parte tem como justificativa a LGPD e outras questões como a inadequação do pedido original.

Com efeito, o advento da chamada sociedade da informação, cujo cenário é marcado pela ubiquidade da tecnologia representado pelo *device* dominante que é o *smartphone*, produziu uma cultura crescente de valorização da transparência (dados abertos, acessíveis ao público). No caso do jornalismo investigativo contemporâneo gerou o chamado Jornalismo de Dados, especialmente a partir do trabalho e da obra de Philip Meyer (*"The New Precision Journalism"*, publicado nos EUA em 1991, ou Jornalismo de Precisão, sem tradução para o português até hoje), que passou a usar a metodologia científica e técnicas de pesquisa das ciências sociais para a produção de reportagens.

A questão mais sensível e complexa que se apresenta ao JD é, do ponto de vista ético, equacionar a necessidade de publicação de informações de incontestado interesse público com o direito à privacidade. Separar o trigo da informação de qualidade e de inequívoco interesse público da justificativa de proteção pela LGPD é um desafio sensível e que precisa ser enfrentado na prática jornalística.

Dados sensíveis, ética jornalística e interesse público

O Código de Ética dos Jornalistas Brasileiros, aprovado em congresso nacional da categoria, em 2007, é claro e preciso no que se refere ao direito à informação. Em seu Artigo 1º, define que "tem como base o direito fundamental do cidadão à informação, que abrange seu o direito de informar, de ser informado e de ter acesso à informação". Além do inequívoco interesse público, o princípio da veracidade dos fatos e o compromisso social devem guiar mãos, olhos e sentidos de

profissionais do jornalismo, de maneira geral.

No caso do JD, os dados sensíveis devem ser observados nos cruzamentos de dados e apuração direta nas fontes de dados públicos. Uma providência sempre à mão é anonimizar as informações, evitando a exposição indevida de fontes primárias, mantendo – quando necessário – o sigilo das fontes, em defesa de sua integridade física, moral e mental. De resto, como recomenda o Artigo 6º do Código de Ética profissional, é dever do jornalista “não colocar em risco a integridade das fontes e dos profissionais com quem trabalha; combater e denunciar todas as formas de corrupção, em especial quando exercidas com o objetivo de controlar a informação”; e, por fim “respeitar o direito à intimidade, à privacidade, à honra e à imagem do cidadão” (Fonte cit.). O código de conduta de jornalistas avança, em seu Artigo 7º para o que não pode ser feito: “Expor pessoas ameaçadas, exploradas ou sob risco de vida, sendo vedada a sua identificação, mesmo que parcial, pela voz, traços físicos, indicação de locais de trabalho ou residência, ou quaisquer outros sinais” (Fonte cit.; grifos meus). Ou seja, no caso do Jornalismo de Dados, cuja apuração é marcada por rigor e precisão, identificar ou expor fontes nessa condição confronta o Código de Ética da profissão.

Na prática, requer um cuidado para não veicular arquivos de dados audiovisuais, imagéticos ou mesmo de fontes documentais que possa expor fontes em situação de vulnerabilidade, especialmente no caso de menores acolhidos pelo Estatuto da Criança e do Adolescente (ECA), que tem regras próprias e rigorosas. A dica aqui é sempre usar técnicas de desfoque de imagens, o uso de programas de edição de áudio para alterar a voz, preservando a identidade da fonte, antes de publicar o conteúdo jornalístico. Do ponto de vista ético, tais procedimentos contribuem para a liberdade de imprensa (proteção da fonte), permitindo que se possa investigar e divulgar informações relevantes, com credibilidade e sem nenhum receio de eventual retaliação ou processo judicial.

Em artigo publicado na mídia jornalística, a pesquisadora Fernanda Campagnucci (Open Knowledge Brasil), em parceria com Bruno Bioni e Rafael Zanatta (Associação Data Privacy Brasil de Pesquisa) alertaram para o uso da LGDP com o objetivo de negar a transparência pública e criar toda sorte de dificuldade à investigação jornalística. Campagnucci et al reconhecem o avanço jurídico da Lei, mas denun-

ciam seu uso "como pretexto conveniente para retroceder nas políticas de transparência. Exemplos não faltam: dos registros dos visitantes ao Palácio do Planalto aos dados de empresas atuadas por trabalho análogo à escravidão, a lei virou mote para ocultar dados públicos". E citando outro exemplo, os pesquisadores citam ainda "remoção de 25 anos de dados educacionais pelo Inep, um dos principais produtores de estatísticas do país, citando uma futura 'adaptação' à LGPD. O novo formato do Censo Escolar de 2021 reduziu dezenas de milhões de registros apenas aos dados de escolas, inviabilizando análises das desigualdades na educação" (Fonte cit.; grifos meus).

Paradoxalmente, o próprio texto da LGPD preconiza que o conteúdo da Lei de Proteção de Dados, em seu Artigo 4º, não se aplicaria ao tratamento de dados pessoais realizado para fins exclusivamente jornalísticos e artísticos. A questão é distinguir o que é jornalístico e de inconteste interesse público, neste cenário dominado pela indústria das fake news, comandada em escala global por um oligopólio digital privado (Alphabet/Google, Amazon, Meta, Microsoft e Nvidia), composto por megacorporações de comunicação e dados. Sim, não se trata de empresas de tecnologia ou "plataforma digitais", cujo faturamento anual alcança a cifra dos trilhões de dólares, e que produz um impacto social, político, cultural e econômico em proporções inéditas na história humana.

Considerações finais

As reflexões sobre a ética jornalística no âmbito das atividades práticas do Jornalismo de Dados (JD), sob a baliza do direito à informação pública pelo princípio da transparência (cuja síntese maior é a LAI) – de um lado – e de outro o direito à privacidade, igualmente acolhida na LGPD, constitui um permanente desafio às/aos profissionais do jornalismo, independentemente da natureza do veículo e da metodologia da reportagem em questão: se tradicional, guiada por dados ou investigativa – ou tudo junto e misturado.

Os riscos de exposição de fontes, em situação de vulnerabilidade social, crianças e adolescentes (menores de idade), pessoas com deficiência (PCDs) e idosos – entre outros grupos tido como "minoritários" na sociedade é o mesmo no Jornalismo de Dados que na apuração baseadas

em fontes primárias pessoais e/ou pesquisa documental. Há formas de mitigarmos possíveis deslizos éticos, recolhidos da experiência empírica e da prática da profissão tais como a anonimização dos dados, desfoque de imagens (uso de pixels e outros recursos), embaralhamento de voz (ou edição com câmeras de eco e outros recursos para impedir o reconhecimento da identidade vocal) etc.

É possível concluir, com base nesta breve reflexão, que o Jornalismo de Dados, embora impulsionado por tecnologias e metodologias inovadoras, demanda um rigor ético proporcional à complexidade que o uso de dados sensíveis impõe. A tensão entre o direito à informação e a proteção da privacidade – intensificada pela aplicação ambígua da LGPD e ação nem sempre republicana de agentes do Estado, nos três níveis – exige do jornalista discernimento apurado, alinhado ao interesse público, à veracidade da informação e à responsabilidade social. A manipulação legislativa para restringir o acesso à informação pública ameaça a própria democracia, ao passo que o Código de Ética dos Jornalistas Brasileiros, não obstante ser um parâmetro não-normativo, coloca-se como instrumento de resistência e orientação, à luz do ordenamento jurídico brasileiro – e é frequente as decisões judiciais que o mencionam (Chaves; Klein *et al* , 2022).

Em última análise, cabe à/ao profissional da informação, especialmente no campo do Jornalismo de Dados, cultivar uma postura crítica e atenta diante de usos distorcidos da legislação, adotando práticas que garantam a integridade das fontes e a relevância social

“ Os riscos de exposição de fontes, em situação de vulnerabilidade social, crianças e adolescentes (menores de idade), pessoas com deficiência (PCDs) e idosos – entre outros grupos tido como “minoritários” na sociedade é o mesmo no Jornalismo de Dados que na apuração baseadas em fontes primárias pessoais e/ou pesquisa documental”

das pautas. Em tempos de opacidade travestida de proteção, ética e transparência seguem sendo trincheiras essenciais da cidadania, na defesa permanente do interesse público e da verdade.

REFERÊNCIAS

BRASIL. LEI Nº 13.709, DE 14 DE AGOSTO DE 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. Disponível em <https://l1nk.dev/CKeM5> [Acesso em: 31/05/2025].

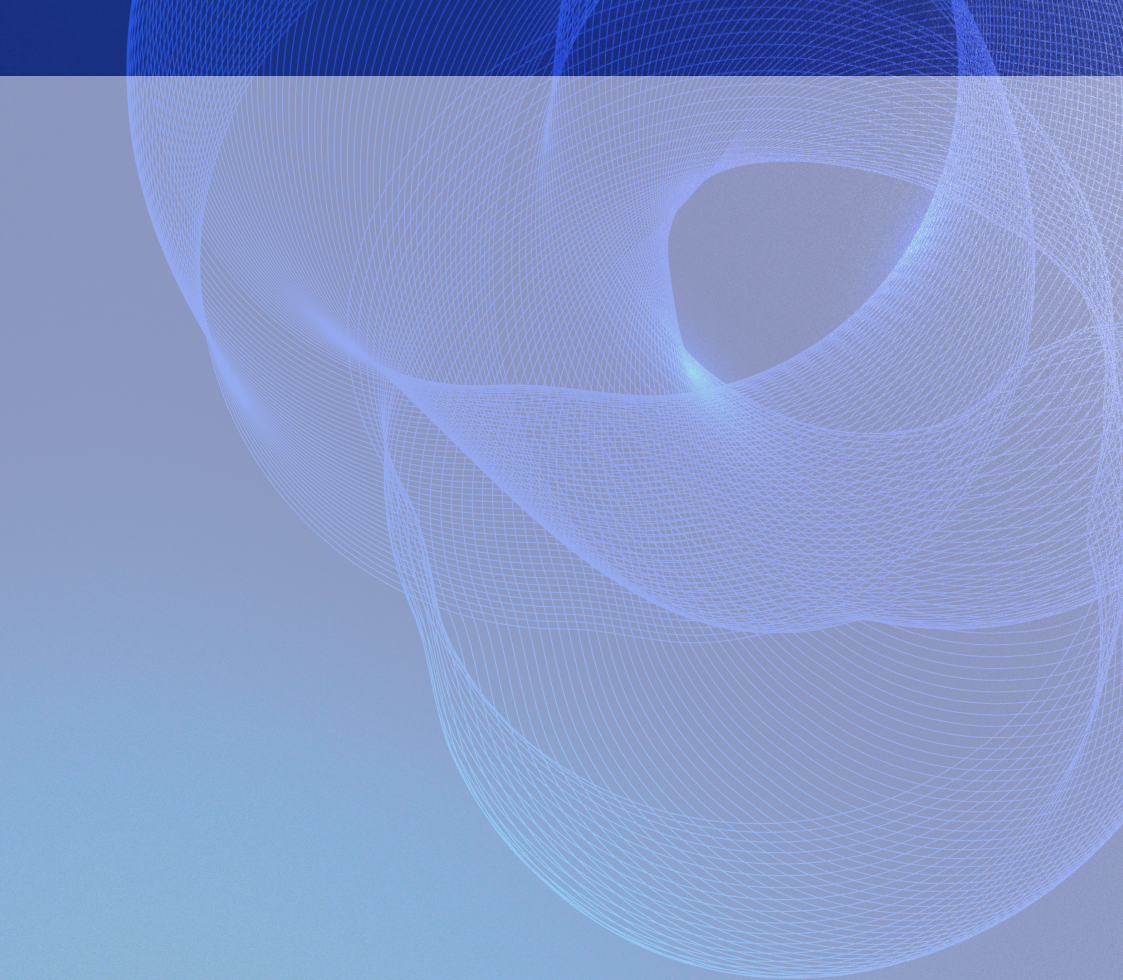
CHAVES, Reinaldo; KLEIM, Letícia; MEIRA, Marina; OLIVEIRA, Lucas; ROSSI, Amanda; SALIBA, Pedro; TOLEDO, Luiz Fernando; VERGILLI, Gabriela; ZAHAR, Cristina; ZANATTA, Rafael. **Jornalismo e proteção de dados pessoais: a liberdade de expressão, informação e comunicação como fundamentos da LGPD**. São Paulo: Abraji, 2022. Disponível em <https://l1nq.com/kYqu6> [Acesso em: maio e junho de 2025].

CAMPAGNUCCI, Fernanda; BIONI, Bruno; ZANATTA, Rafael. **LGPD e transparência: é hora de acertar o passo**. Disponível em: <https://encurtador.com.br/LZTko> [Acesso em: 04/06/2025].

GONÇALVES, Francisco Eduardo. **Limites e potencialidades da Lei de Acesso à Informação**: Dez anos de uso da LAI por jornalistas junto ao governo federal. Tese de Doutorado defendida junto ao Programa de Pós-Graduação em Comunicação da Universidade de Brasília (UnB), em 25 de novembro de 2024. Disponível em: <https://repositorio.unb.br/handle/10482/52240> [Acesso em: 06/06/2025].

FENAJ (Federação Nacional dos Jornalistas). (2007). Código de Ética dos Jornalistas Brasileiros. Disponível em: <https://encurtador.com.br/x8U95>. [Acesso em: 06/06/2025].

SERTSC. Sindicato das Empresas de Rádio e Televisão do Estado de Santa Catarina. Silveira, Marcos Antônio. **Aplicação da LGPD ao conteúdo jornalístico**. Disponível em: <https://l1nq.com/MRkqb> [Acesso em 31/05/2025]



Parte III

Dados têm contexto

Operações básicas com dados e seus sentidos para a narrativa jornalística

Marcelo Crispim da Fontoura

Já sabemos que dados estão em toda parte. Taxas, indicadores, médias, percentuais: o mundo é feito de números. Como disse o jornalista Jonathan Stray, um grande especialista na convergência entre o jornalismo e dados quantitativos, “é impossível entender o mundo moderno sem entender dados” (Stray, 2016, s.p.).

Mas aqui vai um segredo: em que pese a prevalência destes termos e conceitos no noticiário, nas conversas e nas redes, a maior parte das pessoas não compreende o que exatamente eles querem dizer. Até porque existe ainda o outro lado: a frequência com que somos bombardeados com números pode distrair da necessidade de ser crítico com eles. Quando dados são alardeados a todo o momento, e vendidos como a solução para todos os problemas, é fácil esquecer que eles são essencialmente humanos. Como diz a estudiosa norte-americana Danah Boyd, “dados são feitos, não encontrados”.

Qualquer dado é sempre resultado de decisões humanas, normalmente em diferentes níveis, sobre como uma determinada realidade foi interpretada, medida e tratada. A cientista política Deborah Stone (2020) traz ótimas perspectivas sobre essa “humanidade” inerente aos dados. No âmbito jornalístico, isso se torna ainda mais importante: jornalistas precisam entender como o mundo é traduzido em dados a partir de uma visão crítica, assim como devem pensar, eles mesmos, em como aplicar estes números em suas narrativas. É justamente este desafio que vamos abordar aqui. Vamos entender como funcionam os tipos mais básicos de indicadores estatísticos, como usá-los e como pensá-los em contextos jornalísticos.

Medidas de tendência central

Todo mundo conhece a boa e velha média. Está no léxico popular e a população compreende, de modo geral, o que a média quer dizer

– pelo menos a aritmética, embora não tanto suas irmãs, as médias geométrica, harmônica, ponderada. Vamos focar aqui na média aritmética. Ela é o exemplo mais conhecido das medidas de tendência central, que são indicadores que servem para resumir o que está acontecendo com um determinado conjunto, a partir de algum valor central ou típico. Lembre que conjunto é qualquer tipo de grupo de medições – pense nas medições de altura de uma população, ou no registro completo de gols do Brasil em partidas da seleção, por exemplo.

A média tem uma grande vantagem: ela é popular, bem conhecida pela população em geral. Durante a pandemia, uma pequena variação da média também foi muito reportada: a média móvel. Neste caso, se utilizava a média dos últimos sete dias para reportar casos ou mortes de Covid-19. Isso era importante para controlar a sazonalidade dos números: alguns dias da semana tinham mais casos registrados por conta do funcionamento dos órgãos públicos de Saúde. Ao fazer a média dos últimos sete dias, esta variação era controlada e conseguíamos ver melhor o andamento da pandemia. No entanto, a média aritmética, em geral, tem uma grande desvantagem: o fato de que esconde muita informação.

Você viveria em uma cidade cuja temperatura média é $26,5^{\circ}\text{C}$? Parece agradável. Esta é a temperatura média de Riad, na Arábia Saudita. Mas este número esconde uma variação anual de 34° (43° em julho e 9° em janeiro). Já temos aí uma pista sobre o problema da média. Precisamos tomar decisões baseadas também na variação.

Aí entra outra medida de tendência central bem importante, mas de menos apelo popular: a mediana. Ela representa o elemento central de um conjunto, ou seja, aquele que o divide em dois: 50% abaixo dele, 50% acima (este conjunto precisa estar em ordem ascendente ou descendente). E mora justamente aí sua utilidade: a mediana serve para conjuntos que são muito variados e díspares, como renda – muita gente ganha pouco, enquanto pouca gente ganha muito. Por conta disso, o salário médio de um grupo pode acabar não sendo muito representativo. Ou pense no preço médio de um imóvel em uma cidade. A mediana é menos afetada por valores extremos dentro do conjunto em análise.

Uma reportagem da Você S/A é bem didática quanto a isso, mostrando que podemos usar a mediana jornalisticamente, mesmo

que ela seja menos óbvia para a população: “uma das ferramentas usadas por economistas para medir desigualdade, separando ricos e pobres, é a estratificação da população por renda. Primeiro imagine uma fila em que todos os adultos do país são ordenados por renda, da menor para a maior. Depois disso, divida essa população em grupos. Pegue a pessoa no centro, ela tem a renda mediana, que divide a metade pobre da metade teoricamente não-pobre”.

A reportagem da CNN que repercutiu dados do censo também usa bem a medida: “a idade mediana da população brasileira – indicador que divide uma população entre os 50% mais jovens e os 50% mais velhos – atingiu 35 anos, segundo dados do Censo 2022”. Entender a mediana é muito importante para jornalistas, pois ela é uma ferramenta fundamental para descrever conjuntos.

Finalmente, uma terceira medida de tendência central é a moda: o elemento de um conjunto que mais se repete. Ela pode ser útil em algumas situações, mesmo que o jornalista não use a nomenclatura. Por exemplo, em todas as partidas da seleção brasileira de futebol, a quantidade mais frequente de gols marcados pelo time brasileiro é a moda – um dado interessante de ser compartilhado.

E quando sei qual é o momento de usar a mediana ou a média aritmética? Não existe uma regra pronta! A estatística é, sobretudo, humana. O interessante é sempre calcular as duas medidas, e desconfiar quando elas diferem muito entre si. Uma ferramenta útil para isto é o desvio padrão. É um cálculo feito sobre um conjunto e que dá uma característica a mais à média. Um desvio padrão mais alto significa que os dados daquele conjunto estão mais dispersos, variam mais; um desvio padrão mais baixo significa que os dados estão mais concentrados em torno da média.

Os percentis também são úteis para jornalistas

Outra ferramenta útil são os percentis, que nos ajudam a fatiar um determinado conjunto a qualquer ponto, para melhor explicar sua distribuição. Parece complicado, mas é simples: pense na mediana. Ela divide um conjunto no ponto 50%. Seria útil ter outras medidas dessa em outros pontos da distribuição. É exatamente isto que o percentil faz. Imagine que você resolve medir a altura de todas as pessoas numa

sala, e talvez você seja uma das mais baixas. Uma maneira de dizer isso seria, por exemplo: eu tenho 1,60m de altura, e estou no percentil 10 da minha sala. Significa que 90% das pessoas da sala são mais altas, e 10% são mais baixas que você. A lógica é a mesma da mediana, mas em vez de “fatiar” no 50%, nós podemos fatiar em qualquer ponto de 1 a 99.

Novamente, é um daqueles conceitos que usamos jornalisticamente, às vezes sem nem usar este nome. A calculadora do Nexo, que compara um salário frente à realidade brasileira, faz isso. De acordo com ela, com dados de março de 2025, uma pessoa que ganha um salário mensal de R\$ 4 mil ganha mais do que 82% dos brasileiros – ou seja, está no percentil 82. A reportagem da Você S/A, citada acima, menciona que quem recebe R\$ 7 mil está no grupo dos 5% mais endinheirados – ou seja, no percentil 95. É algo que as pessoas conseguem compreender de uma maneira em geral, e ao mesmo tempo dá uma ideia da distribuição de um conjunto.

“Jornalistas precisam entender como o mundo é traduzido em dados a partir de uma visão crítica, assim como devem pensar, eles mesmos, em como aplicar estes números em suas narrativas”

Percentuais: uma das principais ferramentas estatísticas

Para manipular indicadores, jornalistas precisam, também, ter familiaridade com percentuais, uma das grandes ferramentas que temos para representar proporções. Aqui vai uma pergunta simples, mas que embanana muita gente: se, num determinado ano, 40% dos acidentes de trânsito envolveram moto, e, no ano seguinte, 50% dos acidentes envolveram este tipo de transporte, qual o aumento percentual deste tipo de acidente? Se você pensou 10%, errou! O aumento foi de 25%. Se ficou confuso, lembre que % nada mais é que uma unidade, como metro ou segundo. Assim, na verdade, estamos comparando o aumento de 40 para 50. A diferença é de 10; 10 é um quarto de quarenta, e um quarto é o mesmo que 25%. Esta é a diferença entre aumento

percentual e pontos percentuais. Houve um aumento de 25%, ou de 10 pontos percentuais.

Durante anos eleitorais, esta diferença fica bem evidente no noticiário. As pesquisas falam sobre candidatos que subiram dois ou três pontos na pesquisa, mas nunca falam de algum candidato que subiu ou diminuiu 3%. Pois, se alguém tem 7% de intenção de voto e subir 3%, ficará com 7,21%, uma diferença mínima. Usar a terminologia correta é importante. Calcular a variação percentual é importante para comunicar ao público como um determinado número se alterou com o tempo. Para isso, o cálculo é simples:

$$\frac{(\text{número mais recente} - \text{número mais antigo}) * 100}{\text{número mais antigo}}$$

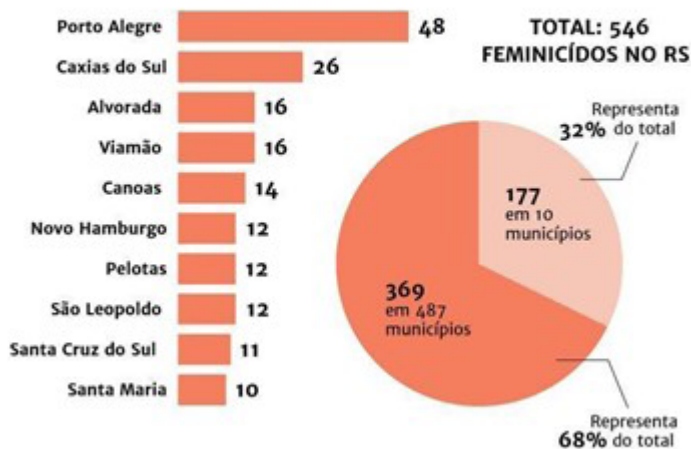
Por outro lado, às vezes, percentuais podem também mascarar a realidade. Com números muito baixos, eles acabam ficando estranhos. Por exemplo, se houve um caso de latrocínio na cidade em um ano e, no ano seguinte, ocorreram três, este é um aumento de 200%, o que soa bem impressionante. No entanto, são números muito baixos, que facilmente causam variações percentuais altas. Nestes casos, é melhor usar os números brutos. Lembre apenas de que, se um número inicial é 0, a variação percentual desse valor para qualquer outro valor é ilegal, ou também descrita como infinito (∞). Isso porque este cálculo envolveria uma divisão por zero, que é ilegal na matemática – e você não quer ser preso!

Levando a população em conta

Outros cálculos muito importantes para jornalistas de dados são aqueles relacionados à população. Considere o gráfico 1, na próxima página, extraído de uma matéria do veículo GaúchaZH sobre o número de feminicídios no Rio Grande do Sul.

Talvez você já tenha percebido o problema. Estas são as cidades mais populosas do estado, então, claro que vão concentrar as maiores quantidades de casos. Onde há mais pessoas, mais coisas acontecem (boas e ruins). Note como o gráfico informa que 10 municípios concentram 32% do total de casos – mas esquece de mencionar que,

Os crimes por cidade



Fonte: dados da Secretaria da Segurança Pública de 2012 a 2017

Gráfico 1 – Feminicídios no RS de 2012 a 2017. Fonte: GaúchaZH

considerando o censo de 2022, estes mesmos municípios concentram 34% da população, ou seja, os casos estão dentro da normalidade estatística para o estado (não que estes números sejam normais em si. Não são! Feminicídio é gravíssimo e deve ser combatido de todas as formas possíveis). No entanto, em nenhum momento, a reportagem usa a população para melhor contextualizar estes números. Há, ainda, um segundo problema, este de visualização de dados: se prestar atenção, verá que as cores no gráfico de pizza deveriam estar invertidas, para representar as mesmas cidades que estão no gráfico de barras.

O correto, neste caso, seria calcular a taxa de casos a cada 100 mil habitantes (para casos de violência de gênero, também se utiliza a cada 100 mil mulheres, bastando dividir a população geral por dois). É bem simples: divide-se o número de ocorrências pela população, multiplicando depois por 100.000. Esta taxa controla o resultado pela população, pois comunica “quantos casos esta cidade teria se ela tivesse 100 mil habitantes”, colocando todas na mesma régua.

No exemplo acima, usando a população do censo de 2022, Alvorada e Santa Cruz do Sul apresentam uma taxa acima de 8 feminicídios a cada 100 mil pessoas, enquanto Porto Alegre, com número bruto de casos bem maior, tem 3,6.

Este tipo de situação é tão comum que existe até um fórum no Reddit chamado, em inglês, “Pessoas vivem em cidades”, para desmistificar aqueles mapas e cálculos que querem representar algo, mas cujo autor não se deu conta de que está apenas apresentando as regiões mais populosas.

Para aprofundar ainda mais

Estas são apenas algumas recomendações sobre como incorporar operações básicas na sua produção jornalística. Nossa principal lição aqui é desconfiar dos dados e abordá-los de um ponto de vista crítico. Caso você queira continuar estudando esta área, Wheelan (2016) é uma ótima referência de livro que explica a estatística em termos leigos e acessíveis, mas construindo uma boa base de raciocínio. Já Bussab e Morettin (2024) escreveram o principal livro-texto da área, fundamental para quem quer entender a estatística mais a fundo.

Acima de tudo: pratique e leia muitos exemplos de jornalismo envolvendo números. As melhores narrativas jornalísticas baseadas em dados são aquelas que se apropriam dos números em uma perspectiva jornalística, nunca perdendo de vista o que querem informar à audiência.

REFERÊNCIAS

BUSSAB, Wilton de O.; MORETTIN, Pedro A. **Estatística básica**. 10ª edição. São Paulo: Saraiva, 2024.

STONE, Deborah. **Counting**: How we use numbers to decide what matters. New York: Liveright Publishing, 2020.

STRAY, Jonathan. **The Curious Journalist's Guide to Data**. New York: Columbia University, 2016

WHEELAN, Charles. **Estatística**: o que é, para que serve, como funciona. São Paulo: Zahar, 2016.

Da estatística descritiva ao jornalismo de soluções

Francisco Amorim

Nos últimos anos, o jornalismo passou por uma profunda transformação em suas formas de produzir conhecimento. A cultura digital, impulsionada pelo avanço tecnológico, ampliou as possibilidades de acesso à informação, mas também contribuiu para o enfraquecimento da credibilidade da imprensa. Nesse cenário, marcado pela desinformação e pela polarização dos debates públicos, propomos neste capítulo uma trilha conceitual que conecta a estatística ao jornalismo de soluções, uma nova racionalidade profissional que vem ganhando força nos últimos anos, tendo o Jornalismo de Dados como elo entre método e narrativa.

A ideia central é mostrar como a análise matemática pode contribuir para uma forma emergente de pensar e praticar o jornalismo, que combina rigor técnico e compromisso público. Com esse objetivo, a discussão neste artigo está estruturada em três partes. Primeiro, analisamos de forma breve as possibilidades da estatística em suas duas dimensões fundamentais: descritiva e inferencial. Segundo, abordamos o Jornalismo de Dados como uma aplicação direta da estatística em nosso campo. Terceiro, apresentamos as principais características do jornalismo construtivo europeu e do jornalismo de soluções desenvolvido nos Estados Unidos, examinando a conexão dessas duas vertentes com o uso de dados, tanto da estatística descritiva quanto dos modelos inferenciais aplicados em abordagens mais avançadas.

Análise estatística como apreensão da realidade

Tradicionalmente associada às ciências exatas, a estatística é uma ferramenta indispensável para um jornalismo que se propõe a lidar com dados. Por meio dela, o repórter consegue mensu-

rar e interpretar fenômenos complexos, baseando-se em evidências matemáticas e não apenas em declarações de suas fontes.

A estatística, em suas duas dimensões descritiva e inferencial, estrutura a passagem do dado bruto à explicação, do número à narrativa. Ela é o método pelo qual o jornalista aprende a observar padrões, identificar desvios e contextualizar histórias. É por meio dela que podemos descobrir se algo é raro ou virou moda, se passa por uma tendência de alta ou de baixa, ou se aquela boa história que ouvimos encontra respaldo nos números.

Em uma frase, é possível dizer que a estatística descritiva ajuda o jornalista a compreender a realidade, enquanto a estatística inferencial permite antecipar tendências e avaliar incertezas. Juntas, tornam-se instrumentos de conhecimento sobre a realidade social (Spiegelhalter, 2022; Tufté, 2009) e estruturam uma prática que, como definem Mielniczuk e Träsel (2017), corresponde a um “jornalismo guiado por dados”, fundamentado na análise e na explicação de evidências quantificáveis.

A estatística cumpre exatamente essa função: traduzir fenômenos complexos em padrões legíveis, permitindo que o repórter reconheça tendências, relações e desvios. Enquanto a estatística descritiva organiza e resume os dados observados, a estatística inferencial utiliza essas observações para fazer previsões ou generalizações a partir de modelos analíticos mais sofisticados.

A estatística descritiva é, portanto, o primeiro passo da análise. Ela serve para apresentar os dados observados de maneira compreensível. Em termos práticos, é o campo dos totais, das médias, das proporções e das taxas, que o jornalista Marcelo Fontoura explica em outro capítulo deste e-book. No Jornalismo de Dados, manifesta-se, por exemplo, nos gráficos de evolução da população, nos percentuais de aprovação de um governo, nos índices de violência ou nas séries históricas de temperatura.

Mas seu papel vai além da simples contagem. A estatística descritiva ajuda a identificar tanto padrões quanto anomalias, permitindo que o repórter formule perguntas: por que um município tem uma taxa de homicídios muito superior à média nacional? Que regiões do país concentram maior número de queimadas? Essa etapa transforma a informação bruta em dado para o teste de uma

hipótese jornalística, que, como afirma Hunter (2013), é o ponto de partida de uma investigação para se contar uma história. No capítulo deste e-book sobre a força da pergunta, escrito pela jornalista Bárbara Libório, você encontra mais detalhes sobre hipóteses no Jornalismo de Dados.

Correspondendo à maior parte dos gráficos e tabelas vistos em conteúdos jornalísticos, a estatística descritiva ajuda a contar visualmente uma história baseada em muitos dados. Em tempos de abundância de números, a estatística descritiva é a ferramenta que recupera o sentido do dado, devolvendo-lhe significado dentro de contextos sociais concretos.

Representar visualmente bem o que se observa é, portanto, um ato de responsabilidade. Como destaca Tufte (2009), a integridade de gráfica começa pela exigência de contar a verdade e se concretiza quando a representação visual dos dados é consistente com a representação numérica.

Saber a diferença entre média, mediana e moda, compreender a função do desvio-padrão ou calcular uma taxa de incidência, como a conhecida taxa por 100 mil habitantes, por exemplo, felizmente, está longe de ser uma tarefa hercúlea. Há publicações que tratam especificamente da estatística descritiva para não matemáticos, conforme indicamos no final do capítulo.

Se a estatística descritiva retrata como as coisas estão, a estatística inferencial busca compreender por que elas mudam e para onde podem estar indo. Baseada em amostras e modelos matemáticos, ela permite estimar a probabilidade de comportamentos futuros a partir de dados já observados. No campo jornalístico, isso significa usar a informação para estimar cenários, identificar tendências e dimensionar incertezas. Está presente nas pesquisas eleitorais, nas

“ É por meio da estatística que podemos descobrir se algo é raro ou virou moda, se passa por uma tendência de alta ou de baixa, ou se aquela boa história que ouvimos encontra respaldo nos números.”

previsões climáticas e nas análises esportivas mais elaboradas.

Em uma redação, elas geralmente aparecem em modelos preditivos elaborados por fontes oficiais ou por especialistas da área. O emprego de estratégias preditivas pelos próprios jornalistas ainda é pouco explorado e esbarra tanto na necessidade de um conhecimento matemático mais profundo quanto na resistência existente dentro do próprio campo. Mesmo assim, nenhum jornalista hoje deveria abrir mão de conhecimentos básicos sobre modelagem estatística e suas aplicações.

A compreensão probabilística da realidade implica uma nova competência profissional, uma espécie de letramento sobre a incerteza. O jornalista contemporâneo precisa saber interpretar percentuais, níveis de confiança e margens de erro, por exemplo, para reconhecer as limitações de modelos e projeções. Isso porque, como afirma Gigerenzer (2014), grande parte do público interpreta mal as probabilidades por tender a pensar em categorias absolutas, como “vai acontecer” versus “não vai acontecer”.

Sem entender a lógica por trás de um modelo, por exemplo, é difícil entender que inferência não é o mesmo que previsão em sentido absoluto, mas uma forma de raciocínio probabilístico que indica o que é mais ou menos provável, dadas as evidências disponíveis. Como lembra Spiegelhalter (2022), adotar um raciocínio probabilístico é aceitar a incerteza. Para o jornalismo, isso é central: o uso responsável de modelos e estimativas permite que o repórter trate o futuro como hipótese, não como certeza, comunicando a incerteza de forma clara e honesta.

Jornalismo de Dados como narrativa estatística

A estatística permite ao repórter compreender o mundo a partir de evidências matemáticas. Mas conhecer não basta. O próximo passo é a tradução desse conhecimento em linguagem pública, realizada pelo jornalismo de dados. Mais do que uma especialidade técnica, ele representa um modo de investigar o mundo por meio da evidência quantificável (Anderson; Revers, 2018; Diakopoulos, 2016; 2019).

Em uma publicação dedicada integralmente ao Jornalismo de Dados, não é necessário, neste texto, defini-lo de forma mais robus-

ta. O que é significativo aqui é compreender seu papel de transformar o raciocínio estatístico em comunicação socialmente relevante. A ideia por trás disso é de que o dado deixe de ser apenas uma fonte e passe a ser um instrumento de raciocínio jornalístico (Appelgren; Nygren, 2014).

O Jornalismo de Dados é a ponte entre a estatística e o jornalismo de soluções, que podemos definir, de forma provisória neste ponto do texto, como um jornalismo que busca compreender o que funciona e por quais razões.

Nesse sentido, a apuração guiada por dados fornece a matéria-prima com a qual o jornalismo de soluções pode apontar, de forma objetiva, caminhos para enfrentar problemas sociais. Em outras palavras, propomos aqui que um jornalismo orientado à solução parta de uma base de dados sobre os problemas que busca compreender e resolver.

A vantagem dessa estratégia é clara: ao conhecer os dados relacionados a um problema, o jornalista amplia sua capacidade crítica, investiga com mais profundidade e contextualiza melhor o que dizem as fontes. Não se trata de assumir uma posição de especialista na questão, mas de estabelecer bases matematicamente sólidas para a compreensão e o questionamento dos problemas.

Uma apuração baseada em dados analisados estatisticamente reduz as chances de análises enviesadas e de adesão a discursos baseados em opiniões vagas. É fundamental para o jornalista compreender, por exemplo, que a média pode representar muito pouco quando o desvio-padrão é alto. Basta pensar que, se dois alunos tiram, respectivamente, notas 2 e 10, a média será 6, um valor que não representa bem o desempenho real dos indivíduos.

Em uma investigação guiada por dados estatísticos também é fundamental desconfiar de inferências feitas sobre bases muito pequenas, pois variações mínimas podem alterar de forma signifi-

“ Se a estatística descritiva retrata como as coisas estão, a estatística inferencial busca compreender por que elas mudam e para onde podem estar indo.”

cativa os percentuais. Em outras palavras, percentuais calculados sobre amostras reduzidas tendem a distorcer as análises sobre a realidade e precisam ser sempre contextualizados.

Se uma cidade registra, por exemplo, um homicídio em um ano e dois no ano seguinte, o aumento é de 100%, mas, quanto maior a população, menos representativo esse percentual se torna. Nesse caso, a avaliação a partir de uma taxa do crime por habitantes poderia oferecer uma interpretação mais adequada do problema da violência neste município. Mesmo assim, a taxa por 100 mil habitantes, costumeiramente usada por demógrafos, pode levar a interpretações distorcidas quando aplicada a municípios muito pequenos. Aqui a solução seria usar uma taxa por mil habitantes. Em outras palavras, a vigilância epistemológica é regra básica do jornalismo de dados.

Jornalismos construtivo e de soluções guiados por dados

Com os cuidados necessários, uma investigação guiada por dados pode sustentar projetos jornalísticos voltados à apresentação de soluções ou à conscientização do público sobre alternativas para enfrentar problemas sociais relevantes. Ou seja, o jornalismo de dados estabelece uma espécie de infraestrutura de informações matemáticas sobre a qual se erguem as práticas mais recentes do campo, como o jornalismo construtivo e o jornalismo de soluções.

Essas duas perspectivas têm ganhado espaço tanto nas discussões acadêmicas quanto nas redações desde o início da década passada. Embora ambas tenham como eixo central a busca por respostas a problemas sociais relevantes, elas propõem caminhos ligeiramente distintos. Em linhas gerais, o jornalismo de soluções se orienta sobretudo pela identificação de respostas concretas, enquanto o jornalismo construtivo busca, primordialmente, estimular o engajamento social e ampliar a compreensão coletiva sobre questões de interesse público. As fronteiras entre as duas vertentes, entretanto, aparecem borradas, e ambos os conceitos seguem em construção. Há autores, como Lough e McIntyre (2018), que chegam a classificar o jornalismo de soluções como um tipo específico de jornalismo construtivo.

Embora o objetivo central deste texto seja examinar as relações entre a estatística e o jornalismo de soluções, parece oportuno, antes, ainda que de forma breve, explorar as possibilidades de uso da matemática na produção de conteúdos orientados por uma racionalidade construtiva. Podemos entender o jornalismo construtivo, segundo McIntyre e Gyldensted (2017), como uma resposta

à percepção crescente de que a ênfase contínua na negatividade afasta a audiência, produzindo uma espécie de fadiga noticiosa que compromete a função pública da imprensa. Outra consequência de um grande volume de notícias negativas seria a distorção da percepção da realidade pelo público, que passaria a ver o mundo como mais perigoso do que ele realmente é.

A partir de uma discussão originada na psicologia, o jornalismo construtivo propõe reequilibrar a cobertura midiática ao incluir contexto, aprendizado e perspectiva de mudança nas narrativas. Seu objetivo é restaurar a relevância social do jornalismo por meio de histórias que iluminam tanto os problemas quanto as respostas possíveis. Conforme McIntyre (2015, *apud* Lough; McIntyre, 2018), a prática envolve o emprego de técnicas de psicologia positiva no trabalho jornalístico para produzir narrativas mais produtivas e envolventes.

Para McIntyre e Gyldensted (2017), o jornalismo construtivo não nega o conflito nem suaviza a crítica, ele os contextualiza, acrescentando informações que apontam soluções possíveis, esperança e resiliência. O foco desloca-se da simples exposição do problema para a compreensão integral da questão. Em vez de perguntar apenas “o que deu errado?”, pergunta também “o que pode ser feito para corrigir?”. Essa mudança discursiva altera o enquadramento emo-

“ O Jornalismo de Dados é a ponte entre a estatística e o jornalismo de soluções, que podemos definir, de forma provisória, como um jornalismo que busca compreender o que funciona e por quais razões.”

cional da notícia e amplia o seu potencial de engajamento.

Nessa perspectiva jornalística mais abrangente, a estatística se apresenta como uma aliada importante ao alicerçar a discussão, evidenciando matematicamente não apenas a dimensão de determinado problema, mas também a viabilidade das propostas para enfrentá-lo. Ela desloca o debate público do campo da especulação para o da evidência, permitindo compreender de forma mais completa as causas e as consequências de questões de interesse público.

A nova prática jornalística tem se difundido, principalmente, no continente europeu por meio de instituições e redes de comunicadores. Um desses projetos é desenvolvido pelo The Constructive Institute, na Dinamarca, que produziu um manual para jornalistas interessados na abordagem. Outro projeto relevante é desenvolvido pelo Bonn Institute, na Alemanha, que promove uma série de ações, entre elas um programa de bolsas. Há ainda o Hellenic Institute for Constructive Journalism, na Grécia, que realiza treinamentos.

Algumas dessas iniciativas europeias mantêm vínculos com o principal projeto de jornalismo de soluções, criado nos Estados Unidos, a Solutions Journalism Network (SJN), que apresentaremos a seguir. Essa integração evidencia que, apesar de diferenças conceituais, ambas as abordagens jurídicas se orientam pela intenção comum de identificar caminhos possíveis para problemas sociais.

Comparado à perspectiva construtiva mais abrangente desenvolvida na Europa, o jornalismo de soluções formulado nos Estados Unidos se apresenta, de certa forma, como mais objetivo, possuindo um método sistematizado e replicável de apuração. Essa vertente parte da constatação de que o jornalismo precisa ir além do diagnóstico: é necessário investigar também as respostas, como afirmam Lough e McIntyre (2018). Conforme os autores, essa estratégia também aumenta o engajamento cognitivo e emocional do público. Ou seja, ao mostrar o que funciona, o jornalismo de soluções reforça a crença de que os problemas são, de fato, solucionáveis.

O jornalismo de soluções tem ganho adeptos nos Estados Unidos nos últimos anos a partir do trabalho da Solutions Journalism Network (SJN). Criada em 2013 por David Bornstein, Tina Rosenberg e Courtney Martin, a organização mantém um laboratório de aprendizagem com treinamentos, kits de ferramentas e guias

voltados ao desenvolvimento profissional de jornalistas. Segundo o relatório anual da SJN de 2024, cerca de 68 mil jornalistas já foram treinados e mais de 16 mil histórias foram catalogadas no Solutions Story Tracker, um banco de reportagens que serve de inspiração para novas investigações.

Parte desse sucesso resulta do fato de que a rede definiu um protocolo de apuração que orienta repórteres em campo, estruturado em quatro pilares:

1. **Foco em uma resposta** a um problema social relevante.
2. **Evidência de eficácia**, sustentada por dados, avaliações ou resultados verificáveis.
3. **Discussão dos limites e desafios** da própria resposta.
4. **Inspiração sem prescrição**, isto é, a reportagem deve informar, não doutrinar.

O protocolo expressa uma racionalidade que aproxima o jornalismo de soluções do jornalismo de dados. Ambos se sustentam na comprovação empírica, compartilhando a mesma episteme: a ideia de que é possível conhecer o mundo por meio de evidências.

Nesse sentido, o método proposto pela SJN evidencia a função da estatística tanto na compreensão do problema quanto na demonstração da eficácia das soluções propostas. A apuração baseia-se na verificação dos dados disponíveis, o que só pode ser realizado, especialmente em fenômenos de maior escala, por meio de informações agregadas. Em outras palavras, na maioria dos casos, trata-se de analisar populações afetadas e não indivíduos isoladamente. Embora as histórias pessoais sejam essenciais para a narrativa, a análise é sustentada por estratégias estatísticas capazes de revelar padrões, tendências e efeitos coletivos.

Essa lógica de apuração começa a se difundir lentamente no Brasil, onde o conceito de jornalismo de soluções ainda passa por uma apropriação gradual por profissionais e acadêmicos. O termo é usado, por vezes, de forma fluida, associado a iniciativas de inovação ou responsabilidade social, indicando um estágio inicial, em que sua adoção prática antecede a sistematização conceitual e metodológica no campo científico brasileiro.

Nesse cenário, o emprego do conceito de jornalismo de soluções pelo Grupo RBS é um exemplo. Nos últimos anos, a empresa tem associado o termo a séries especiais sobre temas sociais importantes, buscando identificar boas práticas e resultados verificáveis. Uma análise de Oliveira, Pastl e Garcia (2023) sobre a cobertura do grupo empresarial, entretanto, revela que o jornalismo de soluções tem funcionado mais como dispositivo discursivo e institucional do que como método de apuração sistemático.

Consequências no ensino do jornalismo

Na formação universitária, estudar essas correntes significa compreender que o jornalismo é, ao mesmo tempo, crítica e construção. Apoiadas em uma investigação guiada por dados, essas abordagens construtivas podem estimar o alcance social de determinados fenômenos, identificar populações mais afetadas e estimar alternativas viáveis de intervenção.

Ao incorporar o raciocínio estatístico à apuração, portanto, o jornalista deixa de ser mero registrador de eventos e passa a ser intérprete de processos sociais, alguém que traduz realidades complexas em conhecimento público útil. Do ponto de vista prático, o dado é transformado em argumento e o número em narrativa pública.

PARA SABER MAIS:

- Análise com estatística descritiva para leigos - tutorial elaborado pela Escola de Dados, programa de alfabetização em dados da Open Knowledge Brasil.
- Estatística Descritiva - artigo do projeto Aprender Fazendo Estatística, de Terezinha Aparecida Guedes, Ana Beatriz Tozzo Martins, Clédina Regina Lonardan Acorsi e Vanderly Janeiro.
- The Constructive Institute - projeto dinamarquês dedicado ao jornalismo construtivo, que disponibiliza online um manual gratuito para jornalistas.

- Bonn Institute - iniciativa alemã que desenvolve pesquisas e formações na área, incluindo um programa de bolsas.
- Hellenic Institute for Constructive Journalism - instituto grego que promove treinamentos.
- Solutions Journalism Network (SJN) - organização que oferece materiais de apoio e programas de capacitação para jornalistas.
- Solutions Story Tracker® - banco de reportagens criado pela Solutions Journalism Network.
- Descobrimo a Estatística Utilizando o SPSS (Andy Field) - apesar de ser um livro voltado para a utilização de um software, a obra explica os principais conceitos estatísticos de forma muito didática e acessível, com muitos exemplos.
- A arte da estatística: Como aprender a partir de dados (David Spiegelhalter) - outra opção interessante Para quem deseja compreender as premissas básicas da estatística inferencial.

REFERÊNCIAS

ANDERSON, C. W.; REVERS, M. **From Counter-Power to Counter-Pepe: The Vagaries of Participatory Epistemology in a Digital Age**. Media and Communication, v. 6, n. 4, p. 24–35, 2018. DOI: 10.17645/mac.v6i4.1770.

APPELGREN, E.; NYGREN, G. **Data Journalism in Sweden: Introducing New Methods and Genres into “Old” Organizations**. Digital Journalism, v. 2, n. 3, p. 394–405, 2014. DOI: 10.1080/21670811.2014.884344.

DIAKOPOULOS, N. **Computational Journalism and the Emergence of News Platforms**. In: FRANKLIN, B.; ELDRIDGE II, S. A. (org.). The Routledge Companion to Digital Journalism Studies. 1. ed. London: Routledge, 2016. p. 176–184.

_____. **Automating the News: How Algorithms Are Rewriting the Media.** Cambridge, MA: Harvard University Press, 2019.

FIELD, A. **Descobrimos a estatística usando o SPSS.** Tradução de Lorí Viali. 2. ed. Porto Alegre: Artmed, 2009.

GIGERENZER, G. **Risk Savvy. How to Make Good Decisions.** New York: Viking, 2014.

GUEDES, T. A.; MARTINS, A. B. T.; ACORSI, C. R. L.; JANEIRO, V. **Estatística descritiva.** Projeto de ensino Aprender Fazendo Estatística, 2005.

HUNTER, M. L.; HANSON, N.; SABBAGH, R.; SENGERS, L.; SULLIVAN, D.; SVITH, F. T.; THORSEN, P. **A investigação a partir de histórias: um manual para jornalistas investigativos.** Paris: UNESCO, 2013. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000226456>. Acesso em: 15 nov. 2025.

LOUGH, K.; MCINTYRE, K. **Journalists' perceptions of solutions journalism and its place in the field.** #ISOJ Journal, v. 8, n. 1, p. 33–52, 2018.

MCINTYRE, K.; GYLDENSTED, C. **Constructive Journalism: Applying Positive Psychology Techniques to News Production.** The Journal of Media Innovations, v. 4, n. 2, p. 20–34, 2017. Disponível em: <https://journals.uio.no/TJMI/article/view/2403/5102>. Acesso em: 15 nov. 2025.

MIELNICZUK, L. P.; TRÄSEL, M. **Jornalismo guiado por dados como inovação profissional e seus desafios para a educação.** Contemporânea: comunicação e cultura, Salvador, v. 15, n. 2, p. 609–629, maio/ago. 2017. Disponível em: <https://lume.ufrgs.br/handle/10183/182425>. Acesso em: 15 nov. 2025.

SPIEGELHALTER, D. **A arte da estatística: como aprender a partir de dados.** Tradução de George Schlesinger. Rio de Janeiro: Zahar, 2022.

TUFTE, E. R. **The Visual Display of Quantitative Information.** 2. ed. Cheshire: Graphics Press, 2009.

Dados geográficos muito além dos “mapinhas”

Thays Lavor

Os dados geoespaciais (ou geodados) contêm informações sobre localização na superfície da terra e permitem relacionar eventos e fenômenos a contextos territoriais. Praticamente toda notícia tem um elemento espacial, pois registrar onde algo aconteceu é parte fundamental da reportagem. Ao combinar notícia com mapas interativos, é possível revelar padrões invisíveis ao método tradicional.

Por exemplo, a plataforma Fogo Cruzado mapeia os tiroteios e violência armada nas regiões metropolitanas do Rio de Janeiro, Recife, Salvador e Belém, permitindo à população e à imprensa compreender os territórios mais afetados e os horários de maior risco. A cobertura se baseia em dados colaborativos, sistemas de georreferenciamento e boletins interativos que fortalecem a vigilância cidadã.

Já a InfoAmazonia, em uma de suas reportagens, revelou o caso emblemático do presidente da Câmara de Recife, Romerinho Jatobá, que acumula R\$ 1,7 milhão em multas ambientais por infrações em uma área protegida da Amazônia. A investigação cruzou dados de satélite com informações sobre políticos multados por crimes ambientais que disputaram as eleições de 2022, mapeando sobreposições entre regiões desmatadas e a atuação de candidatos que votaram contra a legislação ambiental. A investigação expõe, por exemplo, como Jatobá, figura política em âmbito municipal, mantém vínculos com danos em territórios distantes de seu cargo, como a Amazônia, evidenciando redes de influência que transcendem fronteiras geográficas.

O mapa produzido pela plataforma não apenas destaca correlações entre poder político e degradação ambiental, mas também ilustra como os dados geoespaciais conectam escalas distintas de um problema público, de decisões locais a impactos em biomas sensíveis, de atuações em áreas remotas a cadeiras no Congresso Nacional. Essa abordagem inovadora amplia o jornalismo investigativo ao transformar números e coordenadas em narrativas que desafiam

a impunidade e expõem conflitos entre interesses econômicos e proteção socioambiental.

Essa ampliação do papel dos mapas no jornalismo é acompanhada de reflexões teóricas sobre sua função como linguagem. Como destacam os pesquisadores Laranjeira e Aguiar, os mapas vêm se consolidando como uma das linguagens midiáticas mais relevantes para o jornalismo digital. Para os autores, compreender os fundamentos da cartografia é essencial para repórteres e editores que desejam acompanhar as transformações tecnológicas da comunicação: “jornalistas que pretendem trabalhar em organizações digitais nos próximos anos precisam, ao menos, compreender a cartografia como uma das linguagens midiáticas contemporâneas” (Laranjeira e Aguiar, 2024).

Nesse contexto, os mapas não apenas localizam fatos, mas também os narram e interpretam, reunindo dados, território e linguagem visual em uma única interface de compreensão. Essa forma de representação, segundo os autores, envolve um intercampo de estudos entre Comunicação, Geografia e Computação, no qual o jornalismo passa a operar também como mediador de territorialidades.

Longe de serem neutros, os mapas expressam escolhas ideológicas, culturais e políticas. A construção cartográfica exige, portanto, responsabilidade editorial e leitura crítica. O uso de geodados na mídia precisa ir além da simples sobreposição de planilhas e pontos: envolve decisões visuais, curadoria de fontes e verificação técnica dos dados. Trata-se de novas formas de investigar e comunicar com mapas.

Conhecimentos e ferramentas necessárias

Para trabalhar com geodados, o jornalista precisa entender conceitos básicos de SIG (Sistemas de Informação Geográfica): projeções cartográficas, coordenadas (latitude/longitude), dados vetoriais (pontos, linhas, polígonos) e matriciais (imagens raster). Felizmente existem várias ferramentas, gratuitas e comerciais, que viabilizam essa prática. Por exemplo, o QGIS é um software de código aberto amplamente recomendado. Ele permite editar praticamente todos os formatos de dados espaciais (vetoriais e raster) e fazer mapas detalhados.

Em tutoriais didáticos, QGIS é apontado como um ótimo ponto de partida para jornalistas que queiram “explorar dados com mapas”.

Já o Google Earth Engine (GEE) é uma plataforma online com petabytes de imagens de satélite e bases científicas, capaz de detectar, medir e comparar mudanças no território em diversas escalas. Embora tenha curva de aprendizado mais acentuada, o GEE tem sido usado em investigações sobre mudanças ambientais e em atualizações em tempo quase real.

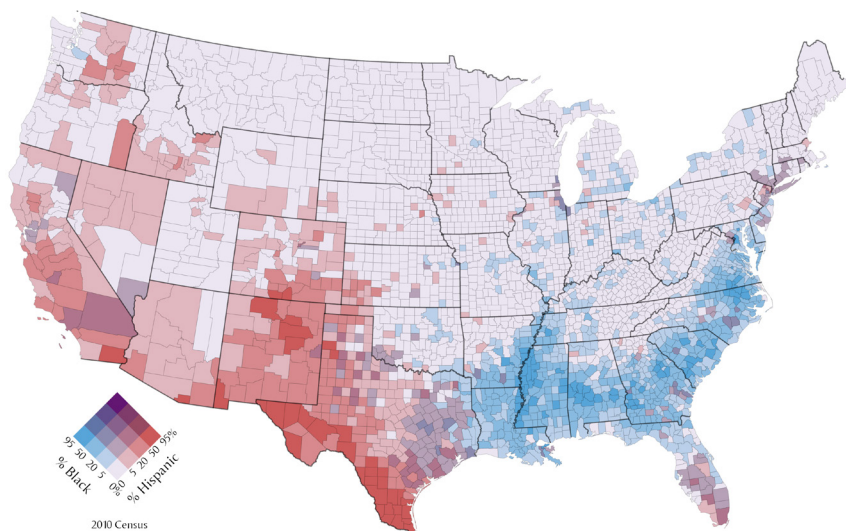
Outras ferramentas complementares incluem provedores de mapas web e bibliotecas de visualização. O Mapbox, por exemplo, permite criar mapas personalizados para a Web com alta qualidade visual. Ferramentas mais simples, como o Google Earth Pro bastam para iniciantes fazerem mapas de pontos a partir de planilhas de endereços. Existem também aplicativos de narrativa geográfica prontos, como o StoryMapJS (Knight Lab), Flourish e o Datawrapper, úteis para inserir mapas em reportagens interativas. Em resumo, a cobertura com geodados demanda familiaridade com SIG básicos, algum treinamento em softwares como QGIS ou plugins GIS, e conhecimento de fontes de dados espaciais (OpenStreetMap, dados oficiais de satélite, APIs de localização).

Principais tipos de Mapas Temáticos

Mapas temáticos são recursos essenciais para representar visualmente a distribuição geográfica de fenômenos sociais, ambientais ou econômicos. Diferenciam-se dos mapas de navegação por se concentrarem em temas específicos, como acesso à saúde, qualidade da educação ou níveis de poluição. Podem apresentar dados categóricos ou quantitativos, como taxas, totais ou índices, e são escolhidos conforme o tipo de dado e a narrativa desejada. Não há um modelo melhor, a decisão depende do objetivo analítico.

Mapa Coroplético

Nesse tipo de visualização, as áreas do mapa são preenchidas com diferentes intensidades de cor para refletir variações numéricas. É ideal para representar indicadores proporcionais, como taxa de alfabetização, cobertura vacinal ou incidência de crimes por região. Sua principal força está na comunicação rápida de padrões geográficos.



Mapa coroplético bivariado comparando as populações negra e hispânica nos Estados Unidos, Censo de 2010.

Fonte: Wikimedia Commons

Mapa de Densidade de Pontos

Utiliza pontos dispersos aleatoriamente dentro das áreas delimitadas, com cada ponto representando uma quantidade específica de ocorrência – por exemplo, um ponto para cada 500 atendimentos em unidades de saúde. É especialmente útil para destacar regiões com maior concentração de um fenômeno, como focos de dengue ou matrículas escolares em áreas urbanas.

Distribuição da população

2022

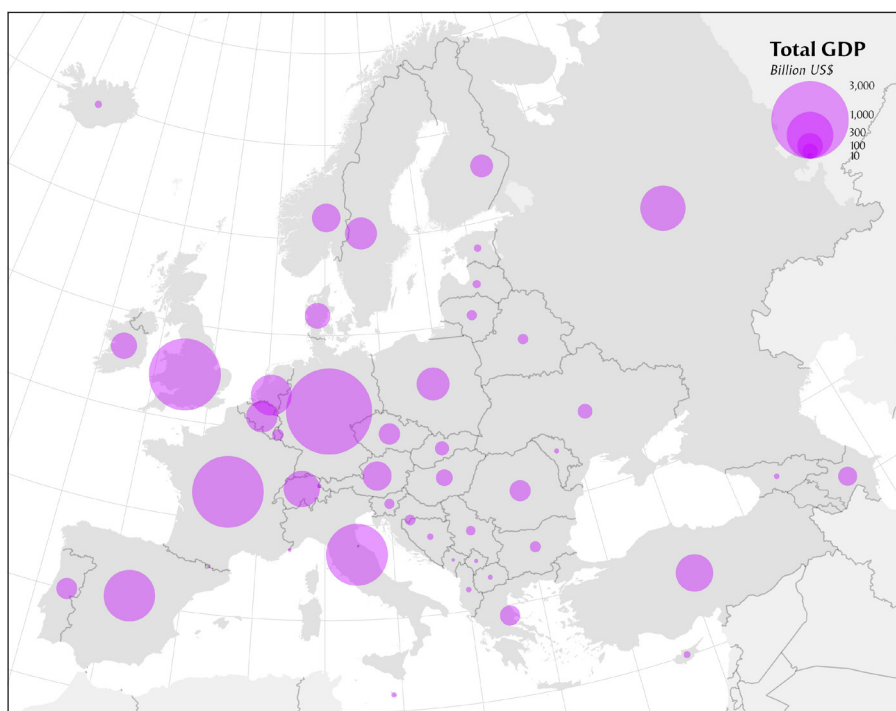


Distribuição da população residente e densidade demográfica (Censo 2022).

Fonte: IBGE. Censo Demográfico 2022: população residente, área territorial e densidade demográfica. In: IBGE. Sidra: sistema IBGE de recuperação automática. Rio de Janeiro, [2023]. tab. 4714. Disponível em: <https://sidra.ibge.gov.br/tabela/4714>.

Mapa de Símbolos Proporcionais

Este tipo de mapa usa símbolos (como círculos ou quadradinhos) com tamanhos proporcionais ao valor que representam. É eficaz para demonstrar volumes absolutos, como número de leitos hospitalares por cidade, quantidade de escolas públicas em determinado estado ou casos de violência em determinadas regiões, entre outros casos de uso.

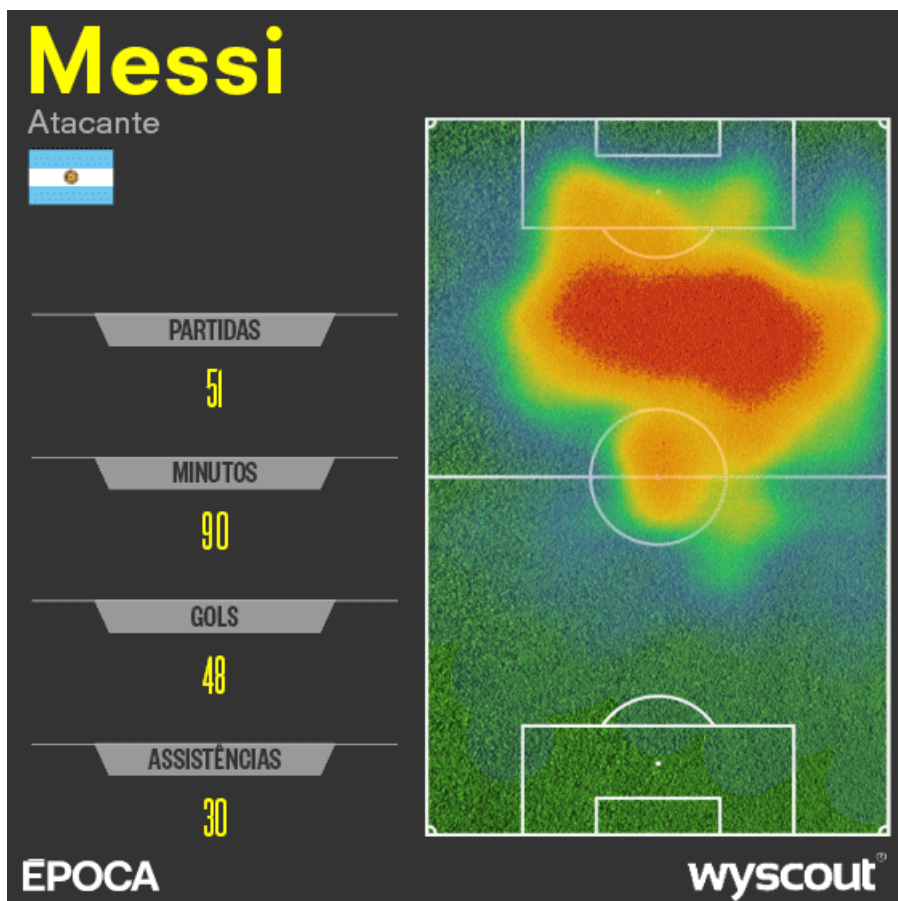


Mapa de símbolos proporcionais do PIB dos países europeus, 2018.

Fonte: Wikimedia Commons

Mapa de Calor (Heatmap)

Trabalha com gradientes de cor sobre uma malha contínua, sem respeitar fronteiras administrativas, representando a intensidade de um fenômeno em uma superfície. É amplamente utilizado para indicar zonas de maior ocorrência de crimes, áreas com alta concentração de poluentes ou regiões com maior mobilidade urbana, como estações de transporte público com grande fluxo. Ou mesmo no futebol, para mostrar áreas dominantes, como no exemplo abaixo.



Exemplos práticos de reportagens com uso de geodados

Vários veículos tradicionais, digitais e projetos independentes têm explorado geodados em escala global. No Brasil, o portal ((o))eco e a InfoAmazonia criaram mapas interativos de desmatamento amazônico. Uma matéria recente do ((o))eco mostrou dados de desmatamento em toda a Pan-Amazônia, ampliando o escopo original do projeto. Internacionalmente, temos casos emblemáticos:

The Guardian (Reino Unido): a série “The Killing Times: a massacre map of Australia’s frontier wars” combina mapa interativo e pesquisa histórica. Ela mostra centenas de locais de massacres de indígenas entre 1788 e 1928, expondo um “processo sustentado e sistemático de conflito” até então pouco divulgado. Esse projeto usou dados arqueológicos, jornalísticos e orais para alimentar um mapa densamente pontuado, proporcionando forte impacto visual e informativo.

Financial Times (Reino Unido): em “Ukraine’s battle against Russia in maps and charts: latest updates”, o FT publica regularmente mapas temáticos sobre a guerra na Ucrânia. São diversos mapas, infográficos e visualizações interativas, integrando dados geográficos (linhas de frente, posições de tropas) a dados sociais (como a distribuição de refugiados ucranianos na Europa). Um mapa de refugiados, por exemplo, usa círculos proporcionais para traduzir a escala do deslocamento humano.

Vantagens do uso de geodados na prática jornalística

O uso de mapas e geodados traz vários benefícios à reportagem:

Precisão espacial: mapas permitem localizar eventos com exatidão. Uma investigação pode, por exemplo, mostrar exatamente onde ocorrem queimadas ou crimes, o que dá mais credibilidade à informação. O exemplo do FT combina dados sociais com mapas, mostrando que o tamanho dos símbolos no mapa traduz a escala real dos deslocamentos de refugiados. Isso evita ambiguidades em texto, indicando onde e quanto.

Impacto visual e clareza: mapas são recursos gráficos poderosos. Um mapa denso e interativo pode “dar uma pausa” na compreensão do leitor, gerando reflexão. Em investigações, mapas atraem atenção e tornam os dados acessíveis, por exemplo, comparações “antes e depois” de imagens de satélite tornam evidente o avanço de incêndios ou enchentes.

Inovação narrativa e proximidade: integrar mapas amplia as formas de contar história. A combinação de dados geográficos com estatísticas sociais, por exemplo, destacando bairros afetados por epidemias ou zonas de influência em eleições, cria narrativas mais ricas e contextualizadas. Esses recursos visuais inovadores ajudam não só na compreensão mas também no engajamento do público, aproximando o leitor de questões complexas.

CAIXA DE FERRAMENTAS

QGIS: software SIG livre para análise avançada de geodados.

Google Earth Engine: plataforma na nuvem com vasta biblioteca de imagens e dados, ideal para monitoramento via satélite.

Mapbox: serviços online de mapas customizáveis, usados para relatórios interativos (permitindo controle de estilo, camadas e dados).

Datawrapper ou Flourish: oferecem módulos fáceis de usar para mapas básicos a partir de dados tabulares.

APIs de mapas: como Google Maps, OpenStreetMap ou Leaflet (JavaScript) para visualizações web responsivas.

Bases de dados: cadastro espacial do Instituto Brasileiro de Geografia e Estatística (IBGE), Instituto Nacional de Pesquisas Espaciais (INPE), Infraestrutura Nacional de Dados Espaciais (INDE), Global Forest Watch, entre outros.

REFERÊNCIAS

FALEIROS, Gustavo. **Por que repórteres e redações devem usar geodados**. International Journalists' Network. Disponível em: <https://ijnet.org/pt-br/story/por-que-repórteres-e-reda%C3%A7%C3%B5es-devem-usar-geodados>. Acesso em: 01 maio 2025.

LARANJEIRA, Antônio Heleno Caldas; AGUIAR, Sonia. **Cartografias jornalísticas**: apontamentos para uma agenda de pesquisa no campo da Comunicação e nos Estudos de Jornalismo. *Questões Transversais*, v. 12, n. 23, p. 35–51, 2024. Disponível em: <https://revistas.unisinos.br/index.php/questoes/article/view/26630>. Acesso em: 18 maio 2025.

DE GROOT, L. **QGIS basics for Journalists**. Berkeley Advanced Media Institute, 2013. Disponível em: <https://multimedia.journalism.berkeley.edu/tutorials/qgis-basics-journalists/>. Acesso em: 01 maio 2025.

OTTER, A. **9 Essential Mapping Tools for Journalists in 2022**. Global Investigative Journalism Network, 2022. Disponível em: <https://gijn.org/resource/9-essential-mapping-tools-for-journalists-in-2022/>(<https://gijn.org/resource/9-essential-mapping-tools-for-journalists-in-2022/>). Acesso em: 01 maio 2025.

Global Investigative Journalism Network. **Resources for Finding and Using Satellite Images**. GIJN, 2023. Disponível em: <https://gijn.org/resource/resources-for-finding-and-using-satellite-images/>. Acesso em: 01 maio 2025.

PINTOR, Ben Hur; ADOCH, Ketty; ILIE, Codrina Maria. **Geodados**: uma introdução gentil. Tradução e edição em português: Bernardo Loureiro. Revisão e editoração: Adriano Belisario e Anicely Santos. Realização: School of Data/Open Knowledge Brasil. Disponível em: <https://gis.escoladedados.org/ebook-gis.pdf>. Acesso em: 19 maio 2025.

Dados em rede: *social listening* e métodos digitais como recursos para o jornalismo

Marcela Canavarro

Métodos digitais e computacionais chegam ao jornalismo por diferentes portas. Além do Jornalismo de Dados, a *expertise* em disciplinas como Ciência das Redes, mineração de texto e visualização de dados contribuem para o avanço de abordagens interdisciplinares entre Comunicação e dados. No ecossistema de agências de comunicação, publicidade e relações públicas, o protagonismo na área de dados é dos departamentos de *business intelligence*, também conhecidos como “B.I.”.

Embora a tradução literal do termo em inglês seja “inteligência de negócios”, ela não é tão utilizada na área de Comunicação como nas disciplinas de Exatas. Em geral, as agências de comunicação e publicidade adotam a expressão anglo-saxã, o que pode gerar confusão. Isso porque *business intelligence* é, há bastante tempo, utilizado na área de tecnologia para designar atividades exercidas por desenvolvedores para coletar, transformar e disponibilizar dados para usuários finais, em um processo conhecido como ETL (*Extract – Transform – Load*).

Em definição que se aproxima daquela aplicada na área de Comunicação, B.I. é um processo para estruturar, acessar e explorar informações para melhorar processos de decisão. Mas há uma diferença fundamental na prática do B.I. nas duas áreas: enquanto em tecnologia e negócios o processo se baseia, principalmente, em dados internos das empresas (como registros de vendas e dados de CRM), na Comunicação, as principais fontes são externas, sendo os dados de redes sociais os mais utilizados.

Nesse contexto, a massificação das redes sociais, principalmente a partir da primeira metade dos anos 2010, trouxe novas fontes de informação relevante para o trabalho profissional de comunicação. Nesse período, as pessoas passaram a interagir mais com empresas, marcas, organizações e pautas políticas ou identitárias, tornando o moni-

toramento de redes essencial para gerir a percepção pública sobre uma marca ou pessoa. Hoje, o *business intelligence* está consolidado nas grandes agências do Brasil, mas ainda em fase de planejamento ou maturação em boa parte das empresas de pequeno e médio portes.

Presença digital e monitoramento de redes

“As pessoas passaram a interagir mais com empresas, marcas, organizações e pautas políticas ou identitárias, tornando o monitoramento de redes essencial para gerir a percepção pública sobre uma marca ou pessoa”

Do ponto de vista corporativo e organizacional, as redes sociais inauguraram novas formas para empresas, organizações e governos interagirem com seus públicos. Perfis em plataformas de rede social se tornaram canais diretos com clientes, públicos e audiências. Isso também inaugurou o chamado SAC 2.0 (serviço de atendimento ao cliente em redes sociais), em que interações como comentários em posts e mensagens privadas ao perfil corporativo passaram a ser sistematicamente analisadas e a contar com um sistema dedicado para gerir o fluxo de resposta aos usuários.

Além disso, também passou a ser fundamental criar e valorizar narrativas positivas da marca nas redes sociais. Por um lado, foram contratadas equipes para criar perfis corporativos fortes e atuantes nas redes. De outro, passou a ser importante identificar menções espontâneas em grupos públicos, publicações de grande alcance e perfis de influenciadores.

Esse tipo de monitoramento é ainda mais crítico em situações de crise reputacional, em que acompanhar o que se diz e o quão rápido se espalha é apenas o primeiro passo para definir uma estratégia eficiente de resposta institucional e frear efeitos perversos de narrativas negativas e, por vezes, desinformativas.

Acompanhar a atividade de influenciadores que abordam temas de interesse de organizações e governos também é parte relevante do

trabalho de B.I. A abordagem tem várias vantagens, já que identifica oportunidades para uma interação positiva com o perfil em questão, monitora canais para a contratação de *publi posts* e identifica rapidamente eventuais críticas ou ataques, ajudando a conter ou mitigar crises de reputação. Em um contexto de crescente politização das redes a partir dos anos 2010, perfis dedicados a difundir opiniões políticas mencionam, com frequência, marcas, organizações, empresas ou figuras políticas, podendo gerar faíscas de crise que rapidamente incendeiam as redes.

No contexto da Comunicação, influenciadores são frequentemente classificados como detratores ou apoiadores, facilitando a contextualização das mensagens e a compreensão do cenário. Também é importante identificar o grau de influência desses perfis, geralmente classificados em influenciadores micro, pequenos, médios ou grandes. As métricas que embasam essa classificação variam de acordo com os objetivos do monitoramento e seu alcance geográfico.

Assim, rapidamente a cultura de dados ganhou fôlego na Comunicação. Diagnósticos, projetos, processos e decisões estratégicas são cada vez mais baseados em análise aplicada e estratégica de dados, principalmente de redes sociais. Nesse cenário de rápida transformação, a comunicação corporativa investe tempo e recursos para atuar de forma permanente em redes sociais, em torno de três eixos: produção de conteúdo estratégico, uso tático das redes para difundir narrativas positivas, e uso aplicado de dados para medir resultados e tornar conteúdos e estratégias mais eficientes, de acordo com os objetivos do negócio. O B.I. pode ajudar em todas elas.

Social listening: a "escuta" das conversas nas redes

Assim, emergiram diferentes abordagens que compõem a caixa de ferramentas de times de B.I. em Comunicação. O *social listening* ou escuta social é, provavelmente, a mais utilizada na Comunicação e consiste no acompanhamento sistemático (geralmente, 24 x 7) de temas e perfis de interesse. Os posts coletados e seus metadados são reunidos em *dashboards* interativos, atualizados quase em tempo real, que podem ser acessados de forma remota e fornecem visualizações práticas para análises táticas e estratégicas.

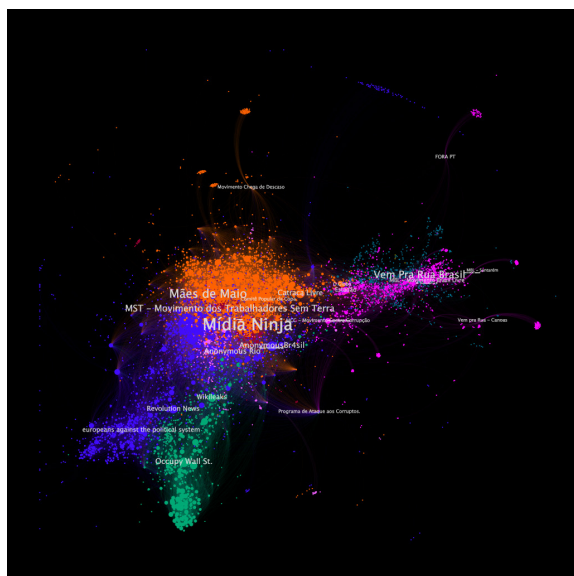
Uma empresa de carros, por exemplo, pode optar por acom-

panhar influenciadores que falem de suas marcas, mas também o debate sobre o impacto de veículos automotores no meio ambiente e o comportamento de gerações mais jovens em relação à compra e à utilização de carros. Ainda que não acompanhe os temas de forma permanente, pode monitorá-los por algum tempo para planejar ou lançar uma campanha específica, de modo a trazer *insights* úteis para alcançar determinados grupos sociais. A definição desse escopo é feita por meio de palavras-chave, *hashtags* e perfis selecionados para a coleta de dados, e que são configurados pelos times de B.I., de acordo com os objetivos de comunicação.

Quanto à análise dos dados coletados, uma das abordagens mais comuns é a de agrupamentos ou *clusters*, usada para conhecer melhor os públicos envolvidos nas conversas das redes. Com um viés quantitativo, algoritmos de aprendizado não-supervisionado identificam grupos de perfis que interagem muito entre si e geram os chamados grafos sociais.

Os grafos sociais representam muitos pontos ligados por algum tipo de conexão. Os pontos podem ser posts ou perfis e a conexão se refere a qualquer tipo de interação, como curtida, compartilhamento ou comentário, entre outros. Com essa visualização, analistas de B.I. identificam que pontos de vista circulam mais em cada grupo, munindo o time com informações relevantes para segmentar a comunicação ou reagir a narrativas negativas para a marca. Junto às métricas de rede, como alcance e engajamento, os grafos sociais ajudam a entender o tamanho do debate nas redes e o quão coeso ou difuso ele está.

“Esse tipo de monitoramento é ainda mais crítico em situações de crise reputacional, em que acompanhar o que se diz e o quão rápido se espalha é apenas o primeiro passo para definir uma estratégia eficiente de resposta institucional e frear efeitos perversos de narrativas negativas e, por vezes, desinformativas”.



Exemplo de grafo social. Elaboração da autora
com dados extraídos do Facebook

A partir da sistematização destes processos, alguns produtos de comunicação se estabeleceram como um padrão em departamentos de B.I., como:

- relatórios de monitoramento do debate público nas redes sobre temas de interesse;
- SAC 2.0 para atender a demandas de clientes e cidadãos via redes sociais;
- relatórios de menções, ou de presença social, para avaliar quantitativa e qualitativamente o quanto a marca é mencionada e em que contextos;
- relatórios de performance para medir alcance, engajamento e repercussão de posts de canais oficiais;
- protocolos de crise com critérios objetivos baseados em

métricas de rede para a identificação precoce de crises de reputação e do conjunto de ações possíveis ou necessárias, de acordo com a gravidade de cada cenário.

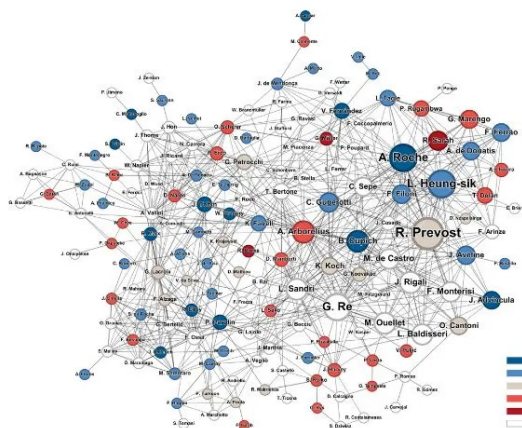
Aplicação no jornalismo

O que une as abordagens do B.I. na comunicação corporativa à prática jornalística são os métodos utilizados, que podem transcender o universo das agências e dos interesses corporativos para somar à produção de informação de interesse público. Já é comum que os veículos recebam relatórios de redes produzidos por agências e grupos de pesquisa para embasar a cobertura jornalística. Mas os jornalistas de redação também podem e devem se apropriar destes métodos para fazer suas próprias análises.

Na ocasião da escolha do papa Leão XIV, em 2025, uma análise de rede destacava a relevância de Robert Prevost, que de fato foi o escolhido, como um dos nós mais relevantes na rede de cardeais da Igreja Católica. Já em 2021, a monarquia britânica reagiu à bombástica entrevista em que o príncipe Harry e sua esposa Megan apontaram racismo na Família Real inundando as redes com uma semiótica que mostrava outra perspectiva. Em outra frente, análises qualitativas de publicações

“ O social listening ou escuta social é, provavelmente, a mais utilizada na Comunicação e consiste no acompanhamento sistemático (geralmente, 24 x 7) de temas e perfis de interesse. Os posts coletados e seus metadados são reunidos em dashboards interativos, atualizados quase em tempo real, que podem ser acessados de forma remota e fornecem visualizações práticas para análises táticas e estratégicas”.

nas redes embasaram ação da Polícia Federal brasileira contra desinformação no contexto das inundações que castigaram o Rio Grande do Sul, em 2024.



Análise de redes prevê corretamente o nome mais cotado para ser o novo papa. Fonte: Soda, Iorio, Rizzo, 2025. Disponível em: <https://www.unibocconi.it/en/news/network-conclave>



trfemoveis.uel.com.br - notícias - redeacao - 2021/03/10 -
Meghan e Harry: Príncipe Charles posa com Negros no ...
 3 dias atrás — Após Meghan falar de racismo, redes de Charles postam sé fotos com negros. Príncipe Charles posa com paciente durante visita à Jesus House, em Londres ...

gntv.ig.com.br - celebridades - charles-posita-fotos-apos-...
Charles posta fotos apenas com negros após Meghan Markle ...

revistaquem.globo.com - notícia - 2021/03 - charles-p...
Charles posa com agentes de saúde e pacientes negros em ...
 3 dias atrás — LONDON, ENGLAND - MARCH 09: Prince Charles, Prince of Wales visits Jesus House. Príncipe Charles durante visita à Jesus House, em Londres (Foto: Getty ...

www.diariodotercemundomundo.com.br - essencial - ap...
Após acusação de racismo, redes do príncipe Charles postam ...
 4 dias atrás — As redes sociais foram inundadas na Jesus House, uma igreja transformada em centro de acolhimento em Londres, e mostram o príncipe interagindo com funcionários negros ...

mundonegro.inf.br - ULTIMAS NOTÍCIAS -
Após Megan falar de racismo na família real, Charles posta ...
 4 dias atrás — Foto de racismo na família real, Charles posta sequência de fotos só com negros ... Na última segunda-feira (8) Meghan Markle e o príncipe Harry concederam ...

Família Real britânica publica em seus perfis de redes sociais fotos de convívio com pessoas negras para rebater acusação de racismo.

Fonte: Reprodução resultados de busca no Google



Ao conhecer os métodos digitais e computacionais usados por profissionais de B.I., os jornalistas podem não apenas produzir suas próprias análises, como também identificar e interpretar estes movimentos nas redes de forma mais qualificada.

Guiado por dados, mas ainda é jornalismo

Beatriz Farrugia

Nos últimos anos, o Jornalismo de Dados conquistou espaço como uma das mais inovadoras vertentes da prática jornalística contemporânea. Ao permitir o cruzamento de grandes volumes de informações com rapidez e precisão, ele transformou o modo como repórteres investigam, apuram e contam histórias. No entanto, em meio a tantas nuances tecnológicas, é fácil esquecer que o coração do processo ainda deve ser o jornalismo, com sua busca por histórias relevantes, apuração rigorosa e compromisso com o interesse público. A fonte pode ter mudado — de uma entrevista para uma base de dados, de um documento físico para uma API —, mas o objetivo permanece o mesmo: informar com rigor, clareza e responsabilidade.

O fascínio pelas ferramentas, códigos e visualizações pode, por vezes, ofuscar o que realmente importa. Não é o uso de Python, R ou SQL que garante a relevância de uma pauta, mas sim a capacidade do jornalista de identificar histórias com valor público e tratá-las com profundidade, contexto e ética. É preciso evitar que a técnica se sobreponha à narrativa, ou que a estética da visualização substitua a força do conteúdo.

Um projeto de Jornalismo de Dados costuma ter sete etapas: **1- hipótese**, que se refere à elaboração da pauta; **2- coleta dos dados**, que é busca pela melhor base de dados para a pauta; **3- limpeza dos dados**, que é o processo de correção de inconsistências na base de dados; **4- análise dos dados**, quando a análise é executada, com cálculos ou métodos específicos para identificar padrões nos dados e encontrar respostas para a hipótese; **5- visualização de dados**, para a criação dos melhores gráficos e visualizações para a reportagem; **6- produção da reportagem**, que se refere à apuração com especialistas e outras fontes para enriquecer e dar contexto à reportagem; **7- publicação da reportagem**, que é o momento de decidir os formatos e plataformas para veiculação.

Se você reparar bem, dentre essas sete etapas, apenas três delas estão mais relacionadas à ciência de dados e exigem habilidades técnicas: a coleta, limpeza e análise dos dados. As outras etapas já fazem - ou deveriam fazer - parte da mentalidade jornalística: o de pensar em temas relevantes para reportagens, de buscar as melhores fontes de informação, refletir sobre como trabalhar visualizações para apoiar a narrativa da reportagem, e sobre o melhor formato de publicação. Ou seja, são habilidades editoriais. Concorde?

HABILIDADES EDITORIAIS

Habilidades Técnicas



1. Hipótese



2. Coleta de Dados



3. Limpeza de Dados



4. Análise de Dados



5. Publicação



6. Produção da Reportagem



7. Visualização de Dados

O Jornalismo de Dados, portanto, é um meio, uma maneira de se atingir o objetivo de produzir uma reportagem relevante, comprometida com o interesse público. E não deve ser entendido como o fim: produzir uma reportagem utilizando linguagem de programação ou composta por gráficos e visualizações.

É por isso que muitos profissionais costumam dizer que o Jornalismo de Dados *ainda* é jornalismo. Ou seja, é apenas um método de se fazer reportagens. No final das contas, fazer Jornalismo de Dados é **fazer jornalismo**.

Os princípios que norteiam a prática do jornalismo tradicional devem estar presentes nas reportagens guiadas por dados também. Vou compartilhar com você um check-list que criei para alunos de jornalismo e o qual também me ajuda a manter o foco nos princípios do jornalismo durante a execução de um projeto envolvendo dados.

- ☐ Relevância
- ☐ Veracidade
- ☐ Contextualização dos dados
- ☐ Personagens
- ☐ Entrevistas
- ☐ Pluralidade

A relevância se refere à relevância da hipótese. Ou seja, o interesse público na pauta. Você já leu sobre a força da hipótese no capítulo escrito pela jornalista Bárbara Libório. A relevância de uma pauta no Jornalismo de Dados não deve estar apenas na disponibilidade de uma base ou na curiosidade que ela possa despertar, mas sim, no seu potencial de interesse público. Em um cenário em que o volume de dados disponíveis é imenso e crescente, é tentador escolher um tema apenas pela facilidade de acesso às informações ou pela estética promissora de uma visualização. No entanto, uma pauta só se justifica quando está conectada a questões significativas para a sociedade, capazes de provocar reflexão, gerar transparência ou estimular mudanças concretas.

Por exemplo, imagine que você seja repórter de cultura de um veículo de imprensa de atuação nacional. Certo dia, navegando por portais de transparência de alguns municípios, você percebe que a secretaria de cultura de Blumenau, em Santa Catarina, acabou de disponibilizar uma base de dados sobre apresentações circenses ocorridas na cidade, no ano passado. Em um primeiro momento, você pode ficar tentado a trabalhar em uma pauta sobre o tema só porque está diante de uma base de dados fresquinha. Mas será que essa pauta seria relevante para o público do seu veículo?

No jornalismo, devemos escolher nossas pautas pelo interesse público, e não apenas pela facilidade de acesso a uma informação. Portanto, o primeiro item do check-list é se certificar de que a hipótese,

a ideia de pauta, está sendo guiada pelo interesse público.

O segundo item do check-list é a veracidade.

Mesmo que uma análise de dados tenha apresentado conclusões relevantes para a pauta, ainda assim é necessário verificar se a informação está correta. O jornalismo trabalha com a verdade, e não se pode correr o risco de noticiar algo errado, incompleto

ou distorcido. É preciso verificar, utilizando outras fontes e recursos, se a conclusão da análise de dados é coerente. Trata-se do mesmo processo que um repórter faria em uma investigação tradicional: caso encontrasse uma informação valiosa, certamente, iria se certificar de sua veracidade através de outras fontes.

O terceiro item do check-list é a contextualização dos dados.

Eu costumo dizer que um dado não diz nada sozinho. Por exemplo, se eu falar que “comi três maçãs na semana passada”, qual conclusão você tiraria? Nenhuma, provavelmente, porque falta contexto para você entender a situação. O mesmo acontece no Jornalismo de Dados. Não adianta realizar todo o trabalho de coleta e análise de dados se, na hora de comunicar a informação, você não contextualizá-la, dando ao público as ferramentas para entender a relevância daquela informação.

Se eu disser que “comi três maçãs na semana passada, após ficar 30 anos sem ingerir uma fruta”, o seu entendimento mudaria, correto? Você entenderia o contexto e a importância que esse fato (a ingestão de uma fruta) teve na minha rotina. Obviamente, esse é um exemplo fictício, mas vale para você perceber a importância de dar contexto aos dados.

A reportagem **As pensões e os bilhões da família militar**, escrita pela equipe da Fiquem Sabendo e publicada na Revista Piauí em 12 de julho de 2021, é um exemplo de como os dados foram bem contextualizados para facilitar o entendimento - e a relevância - da reportagem.

“ Não é o uso de Python, R ou SQL que garante a relevância de uma pauta, mas sim a capacidade do jornalista de identificar histórias com valor público e tratá-las com profundidade, contexto e ética”

No primeiro parágrafo, a reportagem diz que, *"em 2020, o governo federal pagou pensões a 529 mil pessoas – em geral, parentes de servidores civis e militares que já morreram. Somadas, essas pensões custaram R\$ 36,5 bilhões, valor maior que o orçamento do Bolsa Família, que atende cerca de 14 milhões de famílias"*.

Repare que, se a reportagem se limitasse a dizer apenas o valor total de R\$ 36,5 bilhões, seria mais difícil para o leitor dimensionar o montante. Ao fazer uma comparação com o orçamento do Bolsa Família, a reportagem fornece contexto, permitindo que o leitor entenda melhor a grandeza das pensões.

O quarto item do check-list diz respeito aos personagens da reportagem – ou seja, às pessoas diretamente impactadas ou envolvidas na situação que está sendo narrada. No jornalismo guiado por dados, é essencial lembrar que números não substituem histórias humanas. Ao contrário, os dados devem servir como ponto de partida para identificar personagens e ilustrar, de forma concreta, como determinada realidade afeta a vida cotidiana das pessoas. Seja para evidenciar um problema social, uma injustiça, ou até mesmo os efeitos positivos de uma política pública, a presença de vozes humanas confere profundidade, empatia e legitimidade à narrativa jornalística.

No projeto Caminhos da Alimentação, da Gênero e Número, as histórias de quatro mulheres negras das regiões Norte e Nordeste são o fio condutor de uma reportagem que explora dados sobre insegurança alimentar no Brasil. O relato delas humaniza os dados, mostrando que sempre há pessoas e vidas por trás de cada estatística.

O quinto item do check-list são os entrevistados. Além dos personagens, a reportagem guiada por dados precisa incluir entrevistas com especialistas ou outras fontes que possam complementar os dados, explicar cenários futuros ou fazer correlações entre certos fenômenos.

Por fim, **o último item da lista é a pluralidade**. É a reflexão sobre a pluralidade da reportagem, se ela está ouvindo todos os lados da história, representando todos (ou o máximo possível) de pontos de vista, contemplando diferentes fontes de informação, opiniões e perspectivas.

Diante da abundância de dados disponíveis hoje em dia e das inúmeras possibilidades técnicas de análise, é fundamental que o

jornalista mantenha o foco na essência do ofício: contar histórias que importam, com compromisso com a verdade e o interesse público. Porque, apesar das ferramentas, dos códigos e dos gráficos, o Jornalismo de Dados ainda é — e sempre será — jornalismo.

Os ângulos narrativos mais utilizados no jornalismo guiado por dados

Muitas reportagens guiadas por dados chamam a atenção do público pelos belos gráficos e elementos de interatividade. Porém, a narrativa das reportagens guiadas por dados vão além desses dois elementos, e é possível produzir conteúdos diversos, dando enfoque e fazendo recortes específicos para cada história.

O jornalista britânico e professor de Jornalismo de Dados na Birmingham City University (BCU) Paul Bradshaw fez uma pesquisa analisando 100 reportagens guiadas por dados para entender quais eram as narrativas mais utilizadas pelos repórteres. Ele chegou à conclusão de que sete ângulos - ou seja, sete tipos de narrativas e enfoques - predominavam: escala, mudança, classificação, variação, explorar, correlação, e falta de dados.

O ângulo de “escala” se refere a reportagens que identificam um grande problema ou exibem o tamanho de um problema/acontecimento/fenômeno. Foi um ângulo de narrativa muito usado durante a pandemia de COVID-19, quando diariamente os veículos de imprensa noticiavam a média móvel de mortos com o intuito de comunicar a dimensão da pandemia.

Um exemplo do ângulo de “escala” é a reportagem publicada em 2023 pelo O Joio e O Trigo em parceria com a InfoAmazonia. A matéria Ferrogrão afetará pelo menos 6 terras indígenas, 17 unidades de conservação e 3 povos isolados utiliza dados geográficos para mostrar o tamanho do impacto da construção da Ferrogrão (EF-170), um projeto que prevê construir uma ferrovia de quase mil quilômetros atravessando o Brasil.

O ângulo “mudança”, por sua vez, retrata justamente uma mudança em alguma situação: algo vinha acontecendo de uma forma, mas, agora, está diferente. Em maio de 2024, o jornal Folha de S.Paulo publicou

a matéria Registros de ansiedade entre jovens e crianças superam os de adultos pela 1ª vez no Brasil, que é um ótimo exemplo do uso do ângulo “mudança”. Utilizando dados do SUS, a reportagem mostra uma mudança de tendência nos casos de ansiedade, com o número de jovens entre 10 e 20 anos de idade recebendo atendimento médico passando o número de adultos. Trata-se de um novo cenário, representando a mudança.

O ângulo “classificação” tem como objetivo mostrar o ranking de alguma situação, identificando quem/o que está no topo ou quem/o que está nas últimas colocações. É um ótimo ângulo para fazer comparações. Geralmente, essas reportagens trazem títulos com o verbo “liderar”, ou expressões do tipo “é o mais”, “é o melhor”, ou “fica em último”, como essa publicada pelo jornal O Estado de S.Paulo em abril de 2023 sobre o Big Brother Brasil: “BB23: Inteligência Artificial revela quem mais apareceu nos VTs do programa”. A equipe de reportagem coletou dados relacionados ao tempo de exposição de cada participante na TV durante o programa e montou um ranking dos que tiveram mais e menos aparições.

No ângulo “correlação”, tenta-se mostrar a correlação entre dois ou mais fatores, por exemplo, uma redução no orçamento de segurança pública e um aumento no índice de roubos em uma determinada cidade. O site Repórter Brasil mostrou a correlação entre o enriquecimento de pecuaristas e o desmatamento da Amazônia na reportagem especial “Nome aos Bois”, em que se analisou a atividade de 10 fazendeiros que estão entre os maiores pecuaristas do Brasil.

Já o ângulo “variação” narra as discrepâncias em acessos a serviços ou direitos que deveriam ser igualitários e homogêneos. É

“Diante da abundância de dados disponíveis hoje em dia e das inúmeras possibilidades técnicas de análise, é fundamental que o jornalista mantenha o foco na essência do ofício: contar histórias que importam, com compromisso com a verdade e o interesse público”

uma narrativa usada em reportagens que querem apresentar desigualdades ou disparidades. Um exemplo desse ângulo é uma reportagem publicada em 2019 pela rede britânica BBC sobre infraestrutura para carros elétricos. Os dados mostram discrepâncias entre as localidades do Reino Unido em termos de infraestrutura para os veículos e motoristas.

O ângulo “explorar” se refere às reportagens interativas, aquelas em que o público é convidado a clicar e interagir com o conteúdo, tendo controle total sobre a ordem e a profundidade das informações que acessa. Por exemplo, a reportagem Amazon Underworld, da InfoAmazonia, que apresenta um mapa da América do Sul para que o leitor vá clicando e explorando as histórias e os dados relacionados à presença de crime organizado e grupos armados na Amazônia.

Esses ângulos são os mais frequentes no jornalismo guiado por dados, segundo a análise de Bradshaw, mas não são os únicos. Há também abordagens que exploram o uso de simulações ou a reconstituição de eventos com base em dados temporais e espaciais. O que todos estes formatos têm em comum é a aplicação de técnicas narrativas clássicas do jornalismo. Afinal, esses mesmos ângulos, como escala, mudança, classificação, correlação e variação, também estão presentes no jornalismo tradicional, em reportagens baseadas em entrevistas, documentos ou observações diretas.

O que muda, portanto, não é o jornalismo, mas os recursos usados para exercê-lo. Mesmo em narrativas guiadas por dados, o trabalho jornalístico continua exigindo curadoria editorial, apuração criteriosa, consulta a fontes confiáveis, contextualização dos fatos e a presença de personagens que representem a realidade que os números tentam descrever. Em última análise, o que essas reportagens comprovam é que o Jornalismo de Dados é apenas uma nova forma de fazer o que sempre fizemos: contar histórias relevantes, com profundidade, clareza e responsabilidade.

REFERÊNCIAS

Amazon Underworld. **Infoamazonia**. Disponível em: <https://amazonunderworld.org/pt.html>

BRADSHAW, Paul. **Os ângulos mais usados por jornalistas para contar histórias com dados**. Online Journalism Blog, 8 jun. 2021. Disponível em: <<https://onlinejournalismblog.com/2021/06/08/os-angulos-mais-usados-por-jornalistas-para-contar-historias-com-dados/>>

Caminhos da Alimentação. **Gênero e Número**. Disponível em <https://alimentacao.generonumero.media/>

MELITO, L. **Ferrogrão afetará pelo menos 6 terras indígenas, 17 unidades de conservação e 3 povos isolados**. Infoamazonia e O Joio e o Trigo. 5 jul.2023. Disponível em <https://infoamazonia.org/2023/07/05/ferrograo-afetara-pelo-menos-seis-terras-indigenas-17-unidades-de-conservacao-e-tres-povos-isolados/>

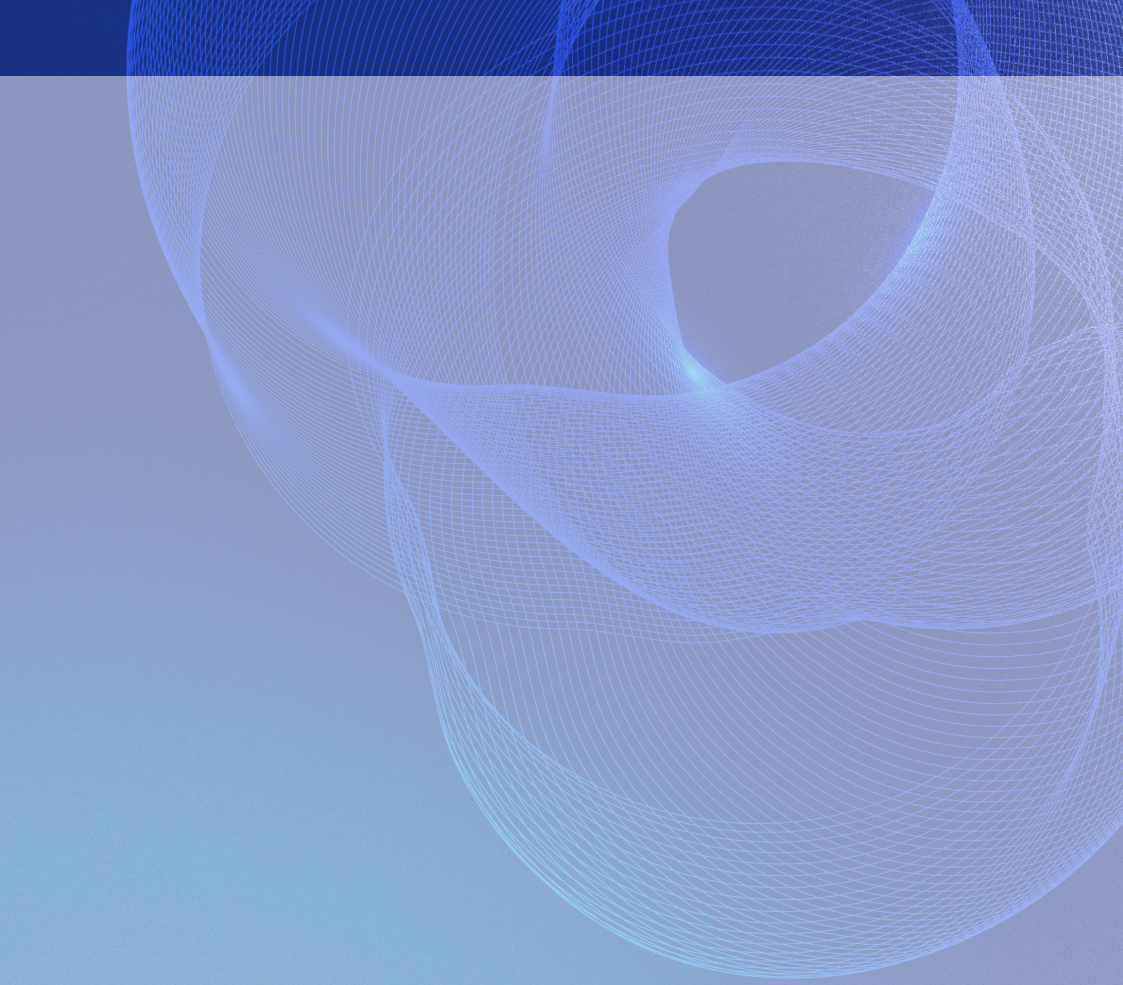
Registros de ansiedade entre crianças e jovens superam os de adultos pela 1ª vez no Brasil. **Folha de S.Paulo**. 31 de maio de 2024. Disponível em <https://www1.folha.uol.com.br/folhateen/2024/05/registros-de-ansiedade-entre-criancas-e-jovens-superam-os-de-adultos-pela-1a-vez.shtml>

ROSSI, M. **Nome aos Bois**. Repórter Brasil. 31 de nov. de 2022. Disponível em <https://nomeaosbois.reporterbrasil.org.br/>

SEIBT, Taís et al. **As pensões e os bilhões da família militar**. Revista Piauí. 12 jul.2021. Disponível em <https://piaui.folha.uol.com.br/as-pensoes-e-os-bilhoes-da-familia-militar/>

STANTON, Aime. **Lack of electric car charging points 'putting off drivers'**. BBC. 4 abril 2019. Disponível em <https://www.bbc.com/news/uk-47696839>

THAYNAN, Lucas et al. **BBB 23: Inteligência Artificial revela quem mais apareceu nos VTs do programa**. O Estado de S.Paulo. 19 de abril de 2023. Disponível em <https://www.estadao.com.br/emails/tv/bbb-23-inteligencia-artificial-revela-quem-mais-apareceu-nos-vts-do-programa/>



Parte IV

Um passo a mais

Visualização de dados para contar histórias mais envolventes

Lucas Thaynan

A visualização de dados, também chamada de *data visualization* ou *dataviz*, é uma área que vem ganhando cada vez mais espaço no jornalismo, impulsionada especialmente pela digitalização das redações, pelo acesso ampliado a grandes volumes de dados públicos e à uma maior disponibilidade de novas ferramentas digitais. Trata-se da representação visual de informações e números de forma a facilitar sua interpretação e análise, seja por meio de gráficos, mapas, tabelas interativas ou infográficos.

Mais do que apenas uma tendência, a visualização de dados se tornou uma necessidade. Em um cenário jornalístico no qual a quantidade de dados disponíveis atualmente vem crescendo gradativamente, transformar números brutos em representações visuais claras e acessíveis permite que os leitores compreendam melhor as histórias que os dados contam.

Imagine, por exemplo, uma reportagem sobre a evolução populacional dos 5.570 municípios brasileiros com base nos dados do Censo Demográfico de 2022. Inserir todas estas informações, ou parte delas, apenas no corpo do texto tornaria a leitura confusa e cansativa. É justamente para evitar esse tipo de problema que a visualização entra em cena: ela organiza a informação e facilita seu entendimento.

Nesse caso hipotético, algumas alternativas de visualização poderiam incluir:

- Uma tabela interativa e buscável, listando todos os municípios e seus respectivos dados populacionais;
- Um gráfico de barras com os 10 municípios que mais cresceram em população;
- Um ranking visual por capitais, destacando as maiores variações;
- Um mapa com cores indicando crescimento ou queda populacional para cada município brasileiro.

Saber usar os recursos visuais corretamente em uma reportagem pode não ser um trabalho fácil, mas sem dúvidas, enriquecerá a sua história.

Diferentes tipos de gráficos e mapas

Devemos sempre ter em mente que um gráfico, nada mais é que um resumo visual dos meus dados. Ao escolher um determinado tipo de gráfico, estou definindo qual história quero contar. Isso nos leva ao grande ponto de atenção: a escolha de uma visualização errada pode prejudicar a história que quero contar.

A seguir, separei os principais tipos de gráficos e mapas, e em quais situações devemos usá-los (na maioria dos casos):

Gráfico de linhas

Utilizado, geralmente, para representar dados ao longo do tempo (como por dia, ano ou década), esse tipo de gráfico é ideal para mostrar tendências. Por exemplo, pode ser usado para exibir a evolução de casos de dengue ao longo dos meses ou a variação de acidentes de trânsito nos últimos anos.

Na Figura 1, vemos um gráfico de linhas que mostra a inflação acumulada de alimentos durante os mandatos dos presidentes Lula e Bolsonaro. A linha do tempo evidencia com clareza as diferenças entre os períodos.

Inflação dos alimentos nos governos Lula e Bolsonaro

% acumulada desde o início do mandato

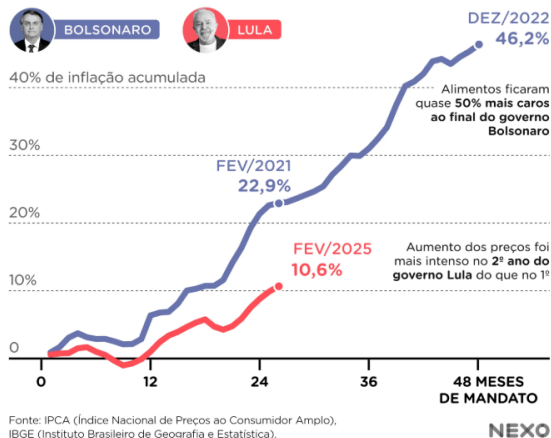


Figura 1 – Gráfico de linhas sobre a evolução da inflação acumulada de alimentos durante os primeiros anos de mandato dos presidentes Lula (terceiro mandato) e Bolsonaro. Crédito: Gabriel Zanlorenssi e Giovanna Hemerly.

Fonte: Nexo Jornal. Acessar a reportagem em: <https://www.nexojournal.com.br/grafico/2025/03/25/inflacao-alimentos-comparativo-governo-lula-e-bolsonaro-grafico>

Um recurso interessante nessa visualização é o uso de anotações sobre o gráfico, que ajudam a destacar ou explicar pontos importantes que o leitor pode não perceber sozinho. Além disso, os eixos X e Y incluem rótulos informativos, como “40% de inflação acumulada” e “48 meses de mandato”.



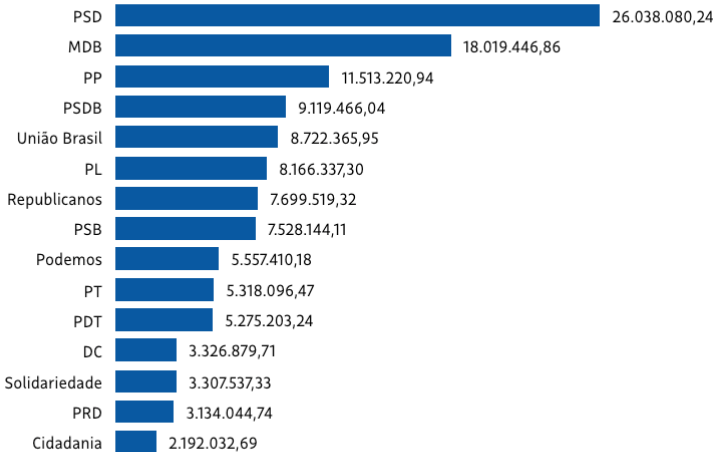
Figura 2 – Principais tipos de gráficos de linha. Fonte: Data Viz Project

Gráficos de barras

Este é o tipo de gráfico mais comum, e não é à toa. O gráfico de barras tem seu uso variado e funciona bem em diversos casos, sobretudo quando se quer comparar mais facilmente a diferença de valores monetários, números absolutos ou mesmo categorias. Ele pode ser formado por barras verticais, horizontais ou mesmo barras compostas.

Partidos que receberam mais doações diretas em 2024

Em R\$



Observação: Dados coletados às 15h51 do dia 8.nov.2024.

Fonte: TSE (Tribunal Superior Eleitoral)

Figura 3 – Gráfico de barras horizontais com dados sobre as doações diretas recebidas pelos partidos em 2024. Crédito: Natália Santos.

Fonte: Folha de São Paulo. Acessar a reportagem em: <https://www1.folha.uol.com.br/poder/2024/11/controlador-da-cosan-e-maior-financiador-eleitoral-pela-4a-campanha-seguida.shtml>

No gráfico de barras verticais publicado pela Folha de São Paulo (Figura 3), é possível comparar, facilmente, como os partidos políticos receberam doações diretas nas eleições de 2024.

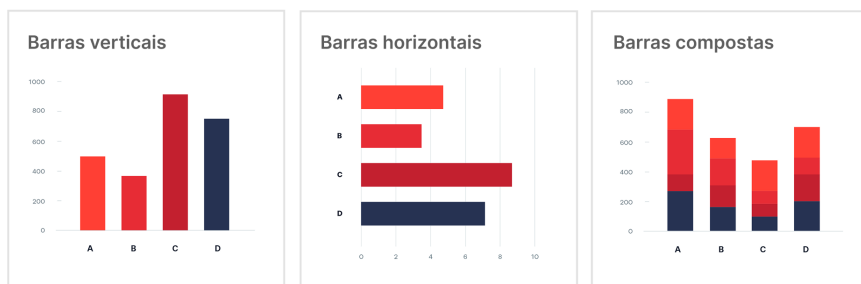


Figura 4 – Principais tipos de gráficos de barras. Fonte: Data Viz Project

Gráfico de dispersão

Este tipo de gráfico é indicado quando temos duas ou três variáveis e queremos correlacioná-las. Como podemos ver na Figura 5, o Núcleo elaborou um gráfico de dispersão, com linha de tendência, para evidenciar quais perfis de mídia tiveram os maiores engajamentos no X/Twitter.

Foram plotadas três variáveis: a média de engajamento por post (eixo Y), a taxa percentual de engajamento (eixo X) e a classificação dos perfis (em cores). Os pontos mais destacados na parte superior direita são os chamados *outliers*, que indicam desempenhos fora do padrão.

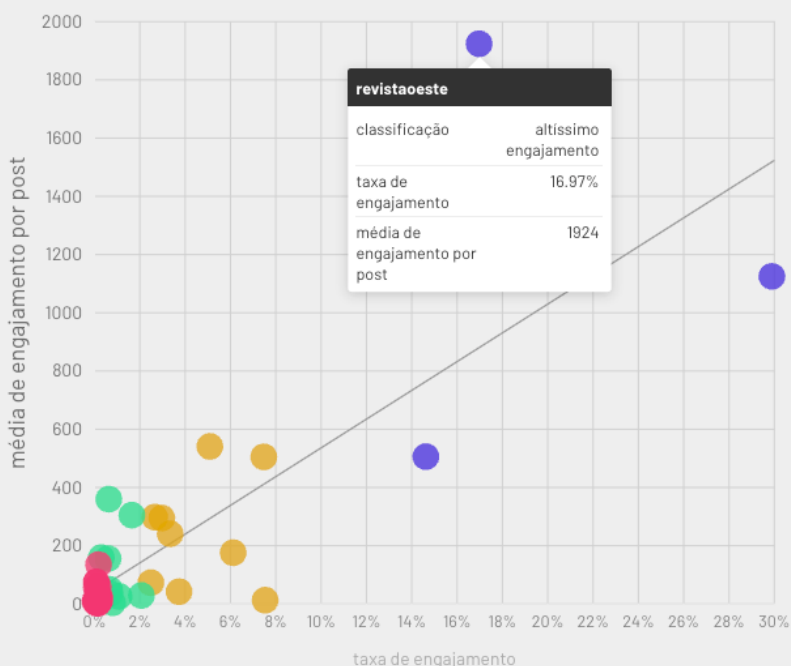
ENGAJAMENTO NA VEIA

Taxa de engajamento de cada publicação em relação à média de engajamento por posts dos últimos 500 tweets

COMO LER O GRÁFICO: Os pontos acima da linha têm a tendência de registrar mais interações por post em relação à sua taxa de engajamento

Todas as classificações ▼

classificação ● altíssimo engajamento ● alto engajamento ● em linha ● baixo engajamento



FONTE: Análise Núcleo Jornalismo

A taxa de engajamento é calculada pela divisão da média de interações por post pelo número de seguidores. A classificação foi feita pelo cálculo de quartil 75% **NÚCLEO**

Figura 5 – Gráfico de dispersão a respeito do engajamento de perfis de média no X/Twitter. Crédito: Sérgio Spagnuolo.

Fonte: Núcleo Jornalismo. Acessar a reportagem em: <https://nucleo.jor.br/interativos/2023-06-01-revista-oeste-e-metropoles-disputam-por-views-no-twitter/>

A partir da visualização é possível claramente ver que o perfil da Revista Oeste desponta com um dos maiores engajamento dentre todos os analisados. Além disso, o gráfico conta com *tooltips* (pequenos balões de informação que surgem ao passar o cursor sobre cada ponto), permitindo ao leitor explorar detalhes como o nome do perfil, os valores exatos das variáveis e sua categoria. Esse recurso torna a análise mais interativa e acessível, enriquecendo a experiência de quem deseja examinar os dados com mais profundidade.



Figura 6 – Principais tipos de gráficos de dispersão.

Fonte: Data Viz Project

Gráfico de pizza

Evite, sempre que possível. Apesar de visualmente atrativos, os gráficos de pizza costumam atrapalhar a interpretação dos dados. Isso porque o ser humano tem dificuldade de comparar áreas visuais e ângulos, além do excesso de cores, que devem ser usadas para diferenciar categorias e não valores da mesma variável.

Por esses motivos, os gráficos de pizzas, em geral, não são boas opções de visualização. Na maioria dos casos, eles podem ser facilmente substituídos por gráficos simples, como o de barras e linhas.

Mapas

Os mapas são recursos visuais valiosos para apresentar dados espaciais. Sempre que você dispuser de informações associadas a lugares (bairros, municípios, estados ou países) eles se tornam uma excelente alternativa para facilitar a interpretação. Em geral, as pessoas têm curiosidade sobre o contexto do local onde vivem, trabalham ou circulam, e o mapa coloca esses dados diretamente sobre o território de interesse, tornando os mais intuitivos.

Existem diferentes tipos de mapas, cada um voltado a um objetivo específico. Você pode saber mais sobre dados geográficos no capítulo específico sobre este tema, escrito por Thays Lavor, mas para não deixar passar, veja abaixo, os principais tipos de visualizações de dados em mapas:

- **Mapa coroplético:** Nesse formato, os limites geográficos (como fronteiras de estados ou bairros) estão claramente delineados e cada área é preenchida por uma cor que representa o valor de um indicador. Tons mais claros costumam sinalizar valores menores, enquanto tons mais escuros indicam valores maiores. Exemplo: um mapa do Brasil por estado em que a intensidade da cor reflete a taxa de alfabetização.
- **Mapa de símbolos (ou mapa de bolhas):** Aqui, um símbolo, geralmente um círculo, é posicionado em coordenadas específicas (latitude e longitude). O tamanho e/ou a cor do símbolo varia conforme o dado representado. Exemplo: círculos nas capitais brasileiras cujo diâmetro é proporcional ao tamanho da população e cuja cor reflete a taxa de crescimento demográfico.
- **Mapa de localidade (ou de pontos):** Útil para exibir informação em escala mais detalhada, marcando cada ocorrência individual no espaço. Exemplo: mapear todos os museus de um determinado município com um ponto para cada instituição, permitindo identificar concentrações e lacunas de oferta cultural.

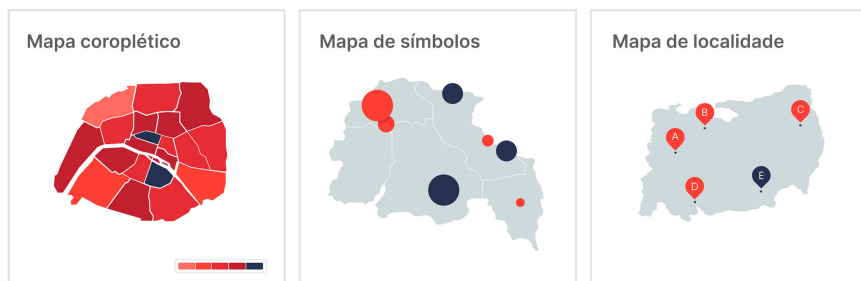


Figura 7 – Principais tipos de mapas. Fonte: Data Viz Project

Dica extra

Além dos exemplos apresentados neste capítulo, existem dezenas de outras possibilidades de visualizações. Se estiver em dúvida sobre qual gráfico ou mapa usar, o site [Data Viz Project](#) pode ser um excelente guia. Ele permite filtrar os gráficos pela função (comparação, distribuição, parte de um todo, etc.), estilo ou tipo de dado. O [The Data Visualisation Catalogue](#) é uma outra opção de plataforma igualmente útil para lhe auxiliar.

Para se inspirar

Nos últimos anos tenho feito muitos projetos voltados para narrativas visuais, tanto na Editoria de Arte do Jornal Estado de São Paulo, como na Agência Tatu. A seguir, compartilharei um pouco do processo de elaboração de um deles que acho mais relevantes:

Radar da Criminalidade

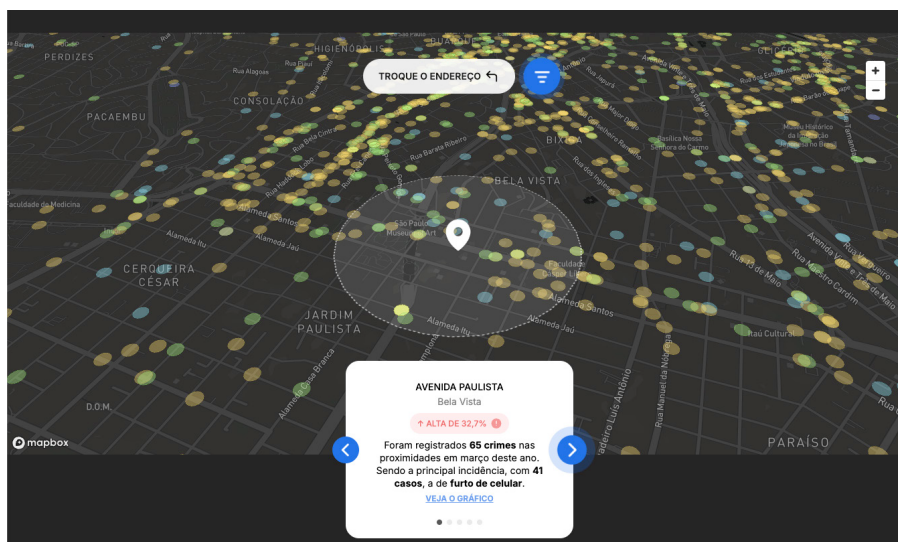


Figura 8 – Radar da Criminalidade mostrando o volume de crimes ocorridos ao redor da Avenida Paulista.

Crédito: Lucas Thaynan, Cindy Damasceno e Bruno Ponceano.

Fonte: Estadão.

“A minha vizinhança é segura?”. Esta foi a pergunta que suscitou o desenvolvimento do projeto, em um momento em que a Editoria de Arte do Estadão tinha o desafio de pensar em reportagens voltadas para o contexto da criminalidade no estado de São Paulo, a partir de uma demanda do jornal em ter uma cobertura mais ampla sobre esta temática.

Com a base de dados disponibilizada pela Secretaria de Segurança Pública de São Paulo (SSP), milhares de registros sobre diversos crimes - tais como, roubos e furtos de aparelhos celulares e veículos, sequestros, homicídios, entre outros - puderam ser coletados.

Todo esse volume de informação está organizado em planilha de dados (no formato CSV) e não possui uma leitura amigável para parte da população. O nosso desafio com o Radar da Criminalidade foi trazer à tona estes registros criminais - em visualizações aces-

síveis - e mostrar como eles estão tão próximos do nosso dia a dia.

No entanto, publicar apenas gráficos e mapas não bastaria, pois cada pessoa vive realidades distintas em uma metrópole como São Paulo. Queríamos ir além: criar uma narrativa personalizada para cada usuário.

No Radar, o leitor é o centro da história. Ao digitar um endereço da capital, ele descobre quantos crimes ocorreram no último mês em sua vizinhança. A ferramenta traça um raio de 500 metros ao redor do ponto informado e destaca no mapa todas as ocorrências dentro desse perímetro.

O design intuitivo conduz o leitor pelos dados sem barreiras de interpretação. Com um simples clique no botão "próximo", o Radar mostra como os registros criminais impactam diretamente o seu entorno, desde a área mais próxima até o distrito policial correspondente.

Erros para não cometer jamais

Há anos coleciono o que chamo de "gráficos infames", ou seja, verdadeiras aberrações visuais que desinformam em vez de esclarecer. A seguir, apresento alguns exemplos clássicos dessa coleção para entendermos o que não devemos fazer ao criar um gráfico.

No primeiro exemplo (Figura 9), vemos um gráfico de barras publicado no perfil do PSDB sobre uma pesquisa eleitoral para o governo de São Paulo em 2018. O problema salta aos olhos: a barra do então candidato João Doria, que deveria representar 22%, foi esticada até se aproximar da marca dos 90%. Além disso, essa barra aparece maior que a de "brancos e nulos", que tem 40%.



Figura 9 – Perfil do PSDB publica gráfico manipulado sobre intenções de votos ao governo de São Paulo de 2018.

Fonte: Perfil do PSDB (Facebook).

A lição é simples: respeite sempre a escala e a proporcionalidade do gráfico. Qualquer distorção compromete a credibilidade do trabalho e engana o leitor.

Neste segundo caso (Figura 10), temos um gráfico exibido na GloboNews sobre a variação da inflação brasileira entre 2009 e 2013. Assim como no exemplo anterior, o tamanho das barras não condiz com os dados, levando a interpretações equivocadas, já que a barra de 2013 aparece maior que a de 2011, embora a inflação de 2013 tenha sido 5,91% e a de 2011, 6,50%. Há ainda outras distorções: a linha da meta oficial (4,5%) e a barra de 2009 também estão em proporções incorretas.



Figura 10 – GloboNews exibe gráfico com distorções sobre dados da inflação entre 2009 e 2013.

Fonte: GloboNews.

Por último, apresento um caso emblemático. O governo do Maranhão divulgou nas redes sociais, em 2023, números sobre o total de alunos em escolas de tempo integral no estado, ilustrando os com um gráfico de barras (Figura 11). Nesse exemplo, o gráfico nem deveria existir, pois em 2014, o valor era zero, portanto não haveria barra. E exibir apenas a barra de 2022 não caracteriza uma comparação válida.



Figura 11 – Governo maranhense publica card errado sobre número de alunos em escolas integrais.

Fonte: Perfil do Governo do Maranhão no Instagram.

Além disso, o intervalo temporal de 2014 a 2022 parece arbitrário, já que a publicação não explica por que esses anos foram escolhidos como comparação. Nessas condições, o gráfico não acrescenta informação, apenas confunde.

Dica extra

Sempre que precisar recriar um gráfico para peças de redes sociais ou infográficos, utilize os gráficos gerados automaticamente por ferramentas como Google Sheets ou Excel. Dessa forma, você assegura que a escala esteja correta e evita erros de proporcionalidade.

CAIXA DE FERRAMENTAS

Há centenas de ferramentas de visualizações de dados disponíveis, mas há duas especialmente poderosas e relativamente simples de usar: Datawrapper e Flourish. Para se ter ideia, grandes jornais brasileiros como Estadão, Folha de São Paulo, G1 e outros internacionais, a exemplo do The New York Times, as utilizam frequentemente. Abaixo, deixo estas e outras sugestões de ferramentas de *dataviz*:

Ferramenta	Plano	Recursos principais	Nível de dificuldade
Datawrapper	Gratuito generoso	Cria gráficos e mapas interativos	Fácil
Flourish	Gratuito generoso	Produz gráficos, mapas e infográficos interativos	Fácil
Mapbox	Gratuito generoso	Permite desenvolver mapas customizados e interativos (exige conhecimento em JavaScript, HTML e CSS)	Difícil
QGIS	Gratuito	Software robusto para tratamento e análise de dados geoespaciais	Difícil
Mapshaper	Gratuito	Alternativa mais amigável ao QGIS para simplificar e converter arquivos geográficos	Moderado
RawGraphs 2.0	Gratuita e open source	Converte rapidamente tabelas em gráficos vetoriais exportáveis (SVG/PDF)	Fácil
Tableau Public	Gratuito (plano online)	Cria dashboards interativos com gráficos e mapas	Moderado

(Continua)

Ferramenta	Plano	Recursos principais	Nível de dificuldade
Infogram	Plano gratuito limitado; assinaturas pagas com recursos avançadas	Cria infográficos, dashboards, gráficos e mapas interativos com modelos prontos	Fácil

REFERÊNCIAS

FALEIROS, Gustav GIANNELLA, Júlia Rabetti; MEDEIROS, Rodrigo Pessoa. **Dataviz em perspectiva**: ensino e prática profissional da visualização de dados no design brasileiro. Rio de Janeiro: 2AB Editora, 2021.

HUFF, Darrell. **Como mentir com estatísticas**. Rio de Janeiro: Globo Livros, 2014.

KNAFLIC, Cole Nussbaumer. **Storytelling com dados**: um guia sobre visualização de dados para profissionais de negócios. São Paulo: Alta Books, 2017.

INOVAÇÃO EM GOVERNO E NO CONTROLE. **Visualização de dados para a transparência e comunicação do setor público**. Medium, 28 jan. 2021. Disponível em: <https://medium.com/inovacao-em-governo-e-no-controle/visualizacao-de-dados-para-a-transparencia-e-comunicacao-do-setor-publico-37b28733bb0f>. Acesso em: 2 maio de 2025.

DATA FRAME. **Data Visualization**: que tipo de gráfico devo escolher para apresentar meus dados? Medium, 2 abr. 2021. Disponível em: <https://medium.com/data-frame/data-visualization-que-tipo-de-grafico-devo-escolher-para-apresentar-meus-dados-a35df040ff18>. Acesso em: 2 maio de 2025.

Programação ajuda a navegar por dados além da superfície

Judite Cypreste

Quando recebi a missão de escrever um capítulo sobre programação para jornalistas, meu primeiro pensamento foi que, finalmente, eu teria a oportunidade de responder a uma das perguntas que mais escuto na vida: “Como posso aprender a programar?”.

A verdade é que não existe um curso certo, uma maneira correta ou um manual capaz de responder a essa questão de forma satisfatória. Sempre que me perguntam isso, costumo devolver com outra pergunta: “Depende. Você quer aprender a programar para quê?”.

E, quase sempre, a resposta revela que o jornalista ainda nem sabe exatamente as razões para querer aprender uma linguagem de programação.

Essa situação é inusitada, principalmente porque muitos de nós escolhemos o jornalismo justamente para fugir dos números e da lógica matemática. Meus boletins escolares do ensino médio que o digam: eu jamais imaginei que um dia estaria escrevendo códigos e muito menos ensinando programação para outras pessoas.

Mas a vida, especialmente a vida na redação, tem um jeito curioso de nos empurrar para fora da zona de conforto. E foi exatamente isso que aconteceu comigo.

A programação entrou na minha vida por necessidade. Eu queria tirar ideias de pauta do papel, mas, em muitos momentos, me deparava com problemas humanamente impossíveis de resolver. Foi aí que percebi que, se eu quisesse contar aquelas histórias, precisaria aprender a programar. Naquele momento, a programação deixou de ser uma opção e virou uma necessidade.

Além disso, todo estudante de jornalismo aprende desde cedo que precisa desenvolver técnicas para identificar, abordar e entrevistar **fontes humanas**. Mas, no mundo de hoje, com o volume de informações digitais crescendo a todo momento, é preciso também aprender a lidar com **fontes de dados**.

Bases de dados se tornaram fontes tão importantes quanto as pessoas e, muitas vezes, é a programação que vai abrir o caminho para você encontrar boas histórias escondidas.

Talvez você ainda esteja se perguntando se realmente precisa aprender a programar. No próximo capítulo, vou contar como eu mesma enfrentei essa dúvida e como a necessidade me levou a descobrir um novo jeito de fazer jornalismo.

Primeiros passos

Quando comecei, confesso que achei que programar fosse fácil. Imaginei que, depois de aprender meia dúzia de comandos, já estaria dominando tudo o que precisava. Logo percebi que não era bem assim. E faço questão de contar isso porque, até hoje, encontro colegas que fizeram um ou dois cursos e saíram frustrados, achando que não aprenderam nada ou dizendo que esqueceram tudo logo depois.

A realidade é que aprender a programar não é como decorar fórmulas. É mais parecido com aprender um **novo idioma**. No começo, você decora algumas palavras e expressões básicas. Entende mais ou menos o que está escrito, mas ainda se sente perdido numa conversa de verdade. Parece difícil, estranho, mas como acontece com qualquer língua, a fluência só vem com a prática.

O meu primeiro contato com programação foi em um curso rápido de Python, uma linguagem muito versátil, capaz de lidar com dados, estruturar informações e realizar análises. Me lembro que uma das primeiras funções que aprendi foi usar o Python como calculadora. Mas convenhamos que, se fosse só para somar e subtrair números, a boa e velha calculadora do celular/computador/da lojinha de R\$ 1,99 já seria suficiente. Achei uma perda de tempo.

Naquela época, vi também alguns outros comandos soltos, mas sem um objetivo para utilizá-los, nada ficou gravado na minha cabeça. Foi então que percebi que somente trabalhando com exemplos reais esse conhecimento realmente se consolidaria.

Meu primeiro projeto foi organizar o cadastro completo dos beneficiários do Bolsa Família no país. Eu queria identificar possíveis irregularidades, cruzar dados e buscar histórias que estivessem escondidas ali. Não tinha um lide, nem história para buscar. Eu só

sabia que, se eu tivesse esses dados na mão, com certeza algo eu encontraria.

Acabei não fazendo nada com isso, mas o aprendizado ficou: estruturar aquele volume de dados manualmente seria impossível. Ali entendi que programar não era somente somar números, era implementar uma lógica para resolver problemas que, de outra forma, seriam insolúveis.

Foi aí que entendi que para aprender programação, tudo fica mais fácil quando você tem um **objetivo concreto**.

No meu caso, a vontade de investigar pautas por meio da análise de dados foi o que me levou ao aprendizado da programação. E a verdade é que o processo de aprender nunca acaba: quanto mais diferentes forem as pautas e as perguntas que você quiser responder, mais você vai precisar expandir sua lógica de programação, aprendendo novas técnicas, novas soluções, enfrentando problemas cada vez mais complexos.

Aprender a programar é um processo contínuo

No jornalismo, isso é uma das maiores vantagens: a cada novo problema, você pode criar uma nova solução. Mas antes de sair escrevendo linhas de códigos por aí, existe uma etapa fundamental, e que muita gente esquece: **entender a lógica dos dados**.

Programar sem dominar a lógica de tabelas, números e bases de dados é como tentar escrever um texto em uma língua que você não entende: você até digita comandos, mas não compreende de verdade o que está fazendo, e isso pode levar a erros graves na análise e na interpretação da informação.

Por isso, antes de avançarmos para o que a programação pode fazer no jornalismo, é fundamental voltar um passo: **nenhum jornalista**

“A realidade é que aprender a programar não é como decorar fórmulas. É mais parecido com aprender um **novo idioma**. No começo, você decora algumas palavras e expressões básicas. Como acontece com qualquer língua, a fluência só vem com a prática”

deve ignorar a base, o domínio de planilhas, a interpretação crítica de dados e o entendimento das estruturas de informação.

A base antes das bases

Se você quer usar programação para investigar dados no jornalismo, precisa primeiro dominar algo ainda mais básico: **entender o que são os dados.**

Antes de sair escrevendo códigos, é essencial saber trabalhar bem com **tabelas, planilhas, números e estruturas de dados.** Sem essa base, aumenta muito o risco de cometer erros graves na interpretação das informações.

Você não precisa ser um gênio em matemática, mas precisa entender conceitos como:

- O que significa cada linha e coluna de uma tabela;
- Saber o que é uma média, uma mediana ou uma regra de três;
- Como identificar inconsistências e padrões numéricos.

Dominar a lógica de planilhas te dá algo mais importante do que saber programar: **te dá consciência sobre o que você está analisando.** É essa consciência que vai te guiar para usar a programação de maneira crítica e eficiente.

Antes de querer automatizar tarefas ou cruzar bases imensas, pergunte-se:

- Eu entendo o que esses dados significam?
- Sei como eles foram coletados?
- Entendo como eles se relacionam?

Se a resposta for “não”, ainda é cedo para programar.

Possibilidades no jornalismo com programação

Agora que já falamos sobre a importância da base de conceitos ao manipular dados, chegou a hora de entender por que aprender programação pode transformar a maneira como você faz jornalismo.

Raspagem de dados (web scraping)

Às vezes, você está no meio da apuração de uma pauta, pesquisa muito, encontra o site certo com as informações que você precisa, mas os dados estão estruturados de um jeito ruim, que impossibilita analisá-los diretamente no Excel. É nesse momento que a **raspagem de dados**, ou **web scraping**.

Em vez de copiar manualmente dado por dado (e correr o risco de levar meses nesse processo), você pode automatizar essa tarefa e coletar tudo de uma vez só. A raspagem permite extrair informações de sites e documentos públicos que não foram disponibilizados de forma organizada.

O g1 usou Python para raspar dados do Banco Nacional de Mandados de Prisão (BNMP) e cruzá-los com informações das candidaturas disponíveis no site de dados abertos do TSE. A apuração¹ analisou mais de 764 mil registros e identificou 61 candidatos nas eleições de 2024 que tinham mandados de prisão em aberto. Depois, os jornalistas checaram documentos públicos para confirmar os casos.

Análises e cruzamentos de grandes bases de dados

Com os dados em mãos, começa outra etapa: a análise. Programar permite identificar tendências, encontrar padrões e descobrir inconsistências em bases volumosas, o que muitas vezes não é viável manualmente.

Você pode até pensar em usar o Excel ou o Sheets para analisar essas informações. Mas, se o volume de dados for muito grande, esses programas podem travar ou nem conseguir abrir a base. Além disso, eles não são as ferramentas ideais para cruzar bases de dados diferentes.

Com programação, é possível calcular variações ao longo do tempo, cruzar bases distintas para encontrar sobreposições e até criar seus próprios indicadores para as reportagens. O Globo²

1 <https://g1.globo.com/politica/eleicoes/2024/eleicao-em-numeros/noticia/2024/09/18/exclusivo-eleicoes-2024-candidatos-alvos-de-mandados-de-prisao-em-aberto.ghtml#levantamento3>.

2 <https://oglobo.globo.com/brasil/mais-de-14-milhoes-de-brasileiros-gastam-ao-menos-duas-horas-ate-uma-uti-para-tratamento-da-covid-19-24442488>.

usou a linguagem R para calcular a distância entre todas as cidades brasileiras e a UTI mais próxima, durante a pandemia de Covid-19. Com apoio da API do Google Maps, a reportagem estimou o tempo das rotas de carro até os hospitais, revelando desigualdades no acesso à saúde no país em meio ao surto da doença.

“ Se você quer usar programação para investigar dados no jornalismo, precisa primeiro dominar algo ainda mais básico: **entender o que são os dados**”

Aplicação de Inteligência Artificial

Após raspar e organizar os dados, é possível ir além da análise tradicional, aplicando Inteligência Artificial (IA) para identificar tendências e padrões com maior precisão. Modelos de IA são capazes de, por exemplo, encontrar padrões em imagens ou textos que poderiam passar despercebidos em uma análise manual.

A DW usou Python para analisar como mulheres de diferentes nacionalidades são representadas em resultados de buscas do Google. Consultas como “mulheres brasileiras” foram feitas em inglês e em línguas locais e as imagens retornadas foram analisadas com IA para identificar conteúdo sexualizado. A reportagem³ mostrou padrões de objetificação em buscas por certas nacionalidades.

Produtos jornalísticos

A programação não serve apenas para coletar e analisar dados. Ela também permite criar produtos jornalísticos interativos. Esses projetos vão além das reportagens tradicionais, oferecendo formas mais dinâmicas e inovadoras de contar histórias e se conectar com o público.

3 <https://www.dw.com/pt-br/como-o-google-contribui-para-perpetuar-estere%C3%B3tipos-sexistas/a-56789034>.

A revista AzMina criou o Elas no Congresso⁴, um projeto que monitora a atuação de parlamentares em pautas de gênero. A ferramenta usa dados públicos para informar o público via Twitter e em uma plataforma com rankings atualizados. O veículo também publica reportagens e uma newsletter sobre o tema.

PARA SABER MAIS

Python

- Python's Not (Just) For Unicorns – por Jonathan Soma. Introdução prática ao Python com foco em aplicações em dados. Disponível em: <https://littlecolumnns.com/learn/python/>
- Python para usuárias de Excel – vídeo do PyLadies São Paulo que ensina Python a partir de comparações com o Excel. Disponível em: https://www.youtube.com/watch?v=hlTaYS6K-A8&ab_channel=PyLadiesS%C3%A3oPaulo

SQL

- SQL Tutorial – curso da plataforma Mode com foco em análise de dados usando SQL. Disponível em: <https://mode.com/sql-tutorial>
- SQL Murder Mystery

Jornalismo de Dados

- Jornalismo de Dados para universitários - Curso introdutório aplicado aos estudantes de graduação em jornalismo na UNB. Disponível em: <https://juditecypreste.com/data%20journalism/2022/07/02/introdu%C3%A7%C3%A3o-jdd.html>
- Curso online sobre Jornalismo de Dados – da Escola de Dados, voltado para jornalistas que querem aprender a trabalhar com dados Disponível em: <https://escoladedados.org/curso-online-sobre-jornalismo-de-dados/>

⁴ <https://www.elasnocongresso.com.br/>.

Reportagens

- Eleições 2024: candidatos alvos de mandados de prisão em aberto (G1, 2024): Disponível em: <https://g1.globo.com/politica/eleicoes/2024/eleicao-em-numeros/noticia/2024/09/18/exclusivo-eleicoes-2024-candidatos-alvos-de-mandados-de-prisao-em-aberto.ghtml#levantamento>
- Mais de 14 milhões de brasileiros gastam ao menos duas horas até uma UTI para tratamento da Covid-19 (O Globo, 2020). Disponível em: <https://oglobo.globo.com/brasil/mais-de-14-milhoes-de-brasileiros-gastam-ao-menos-duas-horas-ate-uma-uti-para-tratamento-da-covid-19-24442488>
- Como o Google contribui para perpetuar estereótipos sexistas (Deutsche Welle, 2021). Disponível em: <https://www.dw.com/pt-br/como-o-google-contribui-para-perpetuar-estere%C3%B3tipos-sexistas/a-56789034>
- Elas no Congresso (Azmina). Disponível em: <https://www.elasnocongresso.com.br/>

Fact-checking vai aos dados para comprovar fatos

Marta Alencar

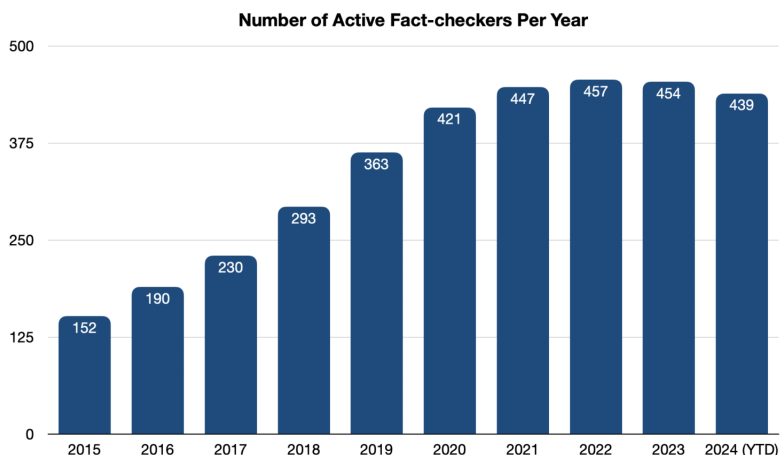
Antes do século XXI, os jornalistas tinham o costume de ir às ruas diariamente para apurar os acontecimentos, geralmente entrevistando os dois lados de uma história. Isso era um trabalho mais árduo e demorado. Até porque jornalistas, por natureza, buscam apurar dados e acontecimentos. Mas com a imersão dos computadores nas redações brasileiras a partir de 1995 e a migração dos grandes jornais para a internet, o processo de apuração que era mais manual, passou a ser mais acelerado e automatizado.

Já houve diversos momentos na história em que boatos e declarações manipuladas impactaram governos, mudaram o percurso de políticos e afetaram a vida de milhões de pessoas. Tais acontecimentos eram registrados em páginas de jornais impressos. Embora, no início, os jornais transmitissem o que era do impresso para a internet, a ascensão do jornalismo para o digital permitiu um maior acesso à pesquisa de dados para aprofundamento de investigações. Esse processo marcou as cinco gerações do jornalismo digital. Com o passar do tempo, também trouxe novas ferramentas e o aprimoramento da verificação no dia a dia do jornalista.

Por volta de 1991, o jornalista Brooks Jackson, que atuava na emissora CNN, pensava em como aprimorar o processo de verificação diante de declarações e promessas de candidatos norte-americanos. O jornalista decidiu checar as declarações de pré-candidatos nos Estados Unidos na TV. Mais tarde, em 2003, ele lançou o *Fact-check.org*, juntamente com o Centro de Políticas Públicas da Universidade da Pennsylvania. Esse processo que Jackson e outros intitulam de *fact-checking* pareceu uma novidade na época, mas há registros de que já na década de 1920, nos Estados Unidos, algumas redações tinham um departamento de verificação das notícias, incluindo redações de revistas como a *Time* (Mantizarlis, 2019). No entanto, com a crise econômica que abalou o mundo em 1929, conhecida como a

Grande Depressão, após o colapso da Bolsa de Valores de Nova York, muitos desses departamentos de verificação fecharam e permaneceu a prática tradicional de apuração jornalística nas redações.

Anos mais tarde, o jornalista Bill Adair, do “Tampa Bay Times”, lançou uma nova seção em seu site, o *Politifact*, que chegou a ganhar um prêmio Pulitzer em 2009. Desde então, o número de sites especializados tem crescido no mundo. Na América Latina, o *Chequeado* foi o pioneiro, criado em 2010, e se tornou uma das maiores inspirações para a implantação de sites de *fact-checking* no Brasil, como a Agência Lupa e o Aos Fatos, iniciativas pioneiras no país, lançadas em 2015, enquanto no mundo existiam 152 iniciativas de verificação de fatos. A prática foi crescendo nos anos seguintes, e chegou a 439 iniciativas em 2024, segundo o censo anual do Duke Reporters Lab.



Originalmente, o *fact-checking* checava principalmente declarações de políticos e boatos em processos eleitorais, utilizando etiquetas para guiar o público sobre o processo de verificação e o porquê da conclusão daquele conteúdo ser falso ou verdadeiro, por exemplo. Nesse processo de checagem de fatos, o mais comum é que os jornalistas lancem mão de competências do Jornalismo de Dados para contrapor declarações públicas com base em evidências encontradas em bases de dados do poder público ou de outras fontes de pesquisa.

Outra prática, conhecida como *debunking*, vem sendo cada vez mais incorporada pelas iniciativas de checagem no contexto das plataformas digitais. O *debunking* analisa e desmascara boatos na internet, acionando ferramentas computacionais para verificação de autenticidade das informações. Com base no terceiro guia publicado pelo First Draft, intitulado “Essential Guide to Understanding Information Disorder”, produzido pela diretora e cofundadora do First Draft nos EUA, Claire Wardle (2019), são sete os tipos mais comuns de desinformação:

- **Conteúdo fabricado:** Novo conteúdo 100% falso, projetado para enganar e causar danos.
- **Conteúdo impostor:** Quando fontes genuínas são representadas, incluindo aqueles que utilizam marcas de empresas jornalísticas ou imagens de repórteres, para mascarar um conteúdo falso.
- **Conteúdo manipulado:** Informações ou imagens genuínas manipuladas para enganar.
- **Contexto falso:** Quando o conteúdo genuíno é compartilhado com informações falsas.
- **Conteúdo incorreto:** Uso enganoso de informações para enquadrar um problema ou indivíduo.
- **Conexão falsa:** Quando manchetes, imagens ou legendas são focadas no sensacionalismo para atrair audiência, mas promovem ruídos ou enganos ao público.
- **Paródia/sátira:** Nenhuma intenção de causar danos, mas tem potencial para enganar, pois à medida que é compartilhada, mais pessoas perde a conexão com a mensagem original e não conseguem entendê-la como sátira.

A partir disso, o *fact-checking* é considerado uma prática de verificação de dados mais detalhada e que questiona declarações públicas (Graves, 2013), mesclando aprendizados tradicionais do jornalismo (apuração) com novas práticas, como a automação, além de não se basear em opiniões. O uso de dados no *fact-checking* é fun-

damental para compreensão e busca em bancos de dados públicos, relatórios e outros documentos, tanto que diversas iniciativas utilizam práticas do jornalismo guiado por dados (JGD) para tal prática, sejam projetos universitários e organizações jornalísticas (Fonseca *et al.*, 2018).

Dourado (2016) defende que o *fact-checking* é um gênero que ascendeu nas coberturas políticas e assume uma nova forma de mediação por meio de uma rotina de checagem e confrontação dos dados, o que pode resultar em uma nova notícia. Seguindo o mesmo pensamento, Seibt (2019) ressalta que o *fact-checking* reposiciona o jornalismo no mercado e corrobora para estabelecer novos parâmetros de transparência que desconstróem a notícia na sua forma convencional. Para a pesquisadora, o momento atual em que a desinformação atinge patamares altos, o *fact-checking* reforça a credibilidade jornalística e a prática da verificação, que vai além das declarações oficiais.

No Brasil, diversas iniciativas foram criadas ao longo dos anos, como Uol Confere (site Uol), Fato ou Fake (site G1) e Estadão Verifica (site Estadão). Iniciativas locais também têm impactado comunidades e regiões, como o Verifica RS, do Rio Grande do Sul, e a COAR Notícias, do Piauí. Ainda, sites que já praticavam o *debunking* há mais tempo, como Boatos.org e E-farsas, hoje participam de discussões conjuntas com *fact-checkers* no combate à desinformação.

É importante salientar que, na última década, houve um surto significativo de crescimento nos projetos de verificação na África, Ásia, Europa e América do Sul. Essa proliferação de agências de checagem nesse período reforçou uma nova oportunidade de negócios, desde a apuração e simples venda de checagens para empresas de mídia,

“ No processo de checagem de fatos, é comum que os jornalistas lancem mão de competências do Jornalismo de Dados para contrapor declarações públicas com base em evidências encontradas em bases de dados do poder público ou de outras fontes de pesquisa”

até oferecimento de cursos e solicitação de doações individuais (Palacios, 2019). Em grande parte do mundo, o período da pandemia, em 2020, certamente contribuiu para o crescimento dessas iniciativas de *fact-checking*. Stencil, Ryan e Luther (2024) identificam que, no ano de 2023, foi a primeira vez que houve mais equipes de verificação que encerraram suas atividades, no total 18. Isso também se deve à falta de financiamentos para manutenção de equipes internas de *fact-checking* em redações ou projetos independentes de checagem.

Alguns desses sites integram a *International Fact-Checking Network* (IFCN), que é uma unidade do Poynter Institute dedicada a reunir verificadores de dados do mundo todo. A IFCN classifica em signatários, aquelas organizações de *fact-checking* que obedecem a seus Códigos de Princípios, com base em relatórios imparciais sobre a veracidade de declarações de figuras públicas e instituições proeminentes, bem como outras alegações amplamente divulgadas relacionadas a questões de interesse público. Esses códigos são: Não partidarismo e justiça; Padrões e a Transparência das Fontes; Transparência do financiamento e da organização; Compromisso com Padrões e a Transparência da Metodologia; Política de Correções Aberta e Honesta. Atualmente, 154 organizações são classificadas como signatários da IFCN.

Transparência e conteúdos com IA

O colapso do antigo modelo de negócios que sustentou o jornalismo desde o início do século XX e o avanço das *big techs* no mercado da informação tem permitido que a desinformação ganhe maior notoriedade do que a notícia verificada e analisada com base no *ethos* jornalístico. Isso é mais notório nos últimos anos, com a ascensão das redes sociais digitais, pois observa-se um excesso de informação intermediado por algoritmos, em que as *big techs* disputam a atenção dos usuários por meio de conteúdos viralizáveis para ampliar lucros com publicidade digital. Essa disputa tem se tornado o principal motor articulador das narrativas de “verdades alternativas” e desinformação (Rêgo; Barbosa, 2020).

Outra questão é que os usuários consomem mais redes de vídeo como YouTube, TikTok e Instagram, conforme apontou o Relatório de

Notícias Digitais de 2024 do Instituto Reuters (Newman; Cherubini, 2025), e que continuará sendo tendência no consumo de notícias nos próximos anos, segundo o mesmo relatório. Nesse cenário, o uso de Inteligência Artificial generativa também tem preocupado pesquisadores e comunicadores, principalmente quanto aos dados.

Carvalho (2024) aponta que a fragilidade do uso de dados pela Inteligência Artificial precisa ser compreendida em três níveis principais: primeiro, sobre as bases de dados utilizadas; segundo, sobre a adulteração de dados por humanos; e terceiro, sobre a adulteração de dados programada. Por exemplo, o ChatGPT da OpenAI, Gemini do Google e Bing da Microsoft são pouco transparentes sobre as fontes de informação consultadas. Nota-se também a ausência de princípios éticos como o reconhecimento da autoria em conteúdos divulgados nessas plataformas de Inteligência Artificial.

A *NewsGuard*, organização estadunidense dedicada a fornecer ferramentas de Inteligência Artificial para combate à desinformação, identificou, até maio de 2025, 1.254 sites de informações não confiáveis gerados por IA (NewsGuard, 2025). Esses sites têm pouca ou nenhuma supervisão humana e publicam artigos, geralmente, contando com alegações falsas, inclusive sobre líderes políticos ou mortes de celebridades.

Assim como IA Generativa tem sido utilizada para desinformar, também tem sido instrumento de iniciativas de *fact-checking* no mundo para esclarecer dados complexos e automatizar as checagens de conteúdos.

Desde 2016, a iniciativa argentina *Chequeado* trabalha com Inteligência Artificial, inicialmente com o *Chequeabot* para acelerar o processo de verificação de dados, incluindo declarações públicas em processos eleitorais na Argentina. Desde 2024, o *Chequeado* tem intensificado novas ferramentas com uso de IA generativa para verificar os dados de maneira mais rápida e com maior alcance (Higuera, 2024). A iniciativa criou um laboratório de Inteligência Artificial, composto por uma equipe interdisciplinar, com a ajuda do fundo Engage, apoiado pela *International Fact-Checking Network*. E lançou, em março de 2025, o guia *Inteligência Artificial generativa en redacciones de fact-checking* para auxiliar outras redações na América Latina (Chequeado, 2025).

A Lupa, reconhecida como primeira agência de checagem do Brasil, foi fundada em 2015, no Rio de Janeiro, e lançou em 2018 o Lupe!,

um assistente digital, que ajudava eleitores brasileiros a checar informações em tempo real por meio do Facebook (Alencar; Aquino, 2023). Assim como o Chequeado, a Lupa também trabalha com IA para coleta e apuração de tarefas complexas de análise, comparação e identificação de padrões em conjuntos de

dados. Também emprega IA para adaptar conteúdos para múltiplos formatos de publicação em redes sociais e aplicativos de mensagens, além da criação de projetos na área de Educação (Lupa, 2025). Entre os projetos com uso de IA, o monitor de discurso nas redes sociais chamado LupaScan, criado em 2022. A ferramenta verifica conteúdos publicados por autoridades públicas, com filtragens específicas para temas particularmente ligados a narrativas desinformativas e discurso de ódio.

Aos Fatos, que é uma organização jornalística que alinha verificação de fatos com tecnologia, incluindo o uso de IA, desenvolve projetos patrocinados de inteligência artificial e *fact-checking* automatizado. Um exemplo é o Radar Aos Fatos, ferramenta em que os usuários podem acompanhar em tempo real como conteúdos estão se espalhando na internet; e a Fátima, ferramenta de automação que funciona como um *chatbot* no WhatsApp. A iniciativa também tem um laboratório de inovação, o *Lab*, que foi criado em 2022 para desenvolver produtos que mesclam tecnologia e *fact-checking*. Os projetos desenvolvidos nesse laboratório já receberam algumas das premiações mais importantes do jornalismo, incluindo o Prêmio Gabo na categoria Inovação, o Digital Media Awards Américas e o Prêmio Cláudio Weber Abramo de Jornalismo de Dados (Aos Fatos, 2025).

Assim como o Chequeado, Lupa e Aos Fatos e várias outras iniciativas de checagem no mundo lançaram políticas editoriais de uso de Inteligência Artificial para reforçar que as verificações de dados com tecnologia contribuem para agilizar as atividades, mas que as ferramentas de IA não substituem o trabalho humano: todos os conteúdos

“ Em grande parte do mundo, o período da pandemia, em 2020, certamente contribuiu para o crescimento de iniciativas de *fact-checking* ”

são sempre revisados, editados e ajustados por humanos. Esse tipo de política editorial é importante porque reforça a transparência das iniciativas de *fact-checking* assim como também a credibilidade jornalística.

Iniciando na década de 1920, nos Estados Unidos, como um processo interno nas redações, o *fact-checking* se desenvolveu como formato de publicação jornalística quase um século depois, surgindo como resposta às transformações digitais na sociedade plataformizada. Agora, a checagem se reposiciona no mercado jornalístico como uma oportunidade de aprimorar e automatizar o processo de verificação de dados, com apoio da inteligência artificial, permitindo que o jornalista tenha novas habilidades digitais para lidar com a desinformação, cada vez mais viral e rápida.

Ainda que o *fact-checking* não seja a solução para as crises enfrentadas pelo jornalismo nos últimos anos, incluindo a de credibilidade e de sustentabilidade dos negócios jornalísticos, essa prática contribui para reduzir os ruídos informacionais no mundo e traz mais transparência aos dados públicos. Com aplicação das competências de Jornalismo de Dados, a checagem de fatos contribui para um ecossistema informacional mais saudável, baseado em evidências.

REFERÊNCIAS

ALENCAR, Marta Thaís; AQUINO, Maria Clara. Algoritmos e desinformação: automação da checagem na América Latina. **Animus**: Revista Interamericana de Comunicação Midiática, Santa Maria, v. 21, n. 47, p. 207–223, jan./abr. 2023. Disponível em: <https://periodicos.ufsm.br/animus/article/view/72211>. Acesso em: 4 maio 2025.

AOS FATOS. **Aos Fatos Lab**. Disponível em: <https://www.aosfatos.org/lab/>. Acesso em: 2 maio 2025.

CARVALHO, Guilherme. **Inteligência Artificial e desinformação**: fragilidades e riscos. Disponível em: <https://www.observatoriodaimprensa.com.br/tecnologia/inteligencia-artificial-e-desinformacao-fragilidades-e-riscos/>. Acesso em: 29 abr. 2025.

CHEQUEADO. **Inteligência Artificial generativa en redacciones de fact-checking**. Disponível em: <https://chequeado.com/inteligencia-artificial-generativa-en-redacciones-de-fact-checking/>. Acesso em: 15 abr. 2025.

DOURADO, Tatiana. Fact-checking como possibilidade de accountability do jornalismo sobre o discurso político: as três iniciativas brasileiras. In: **ENCONTRO ANUAL DA ANPOCS, 40., (ST17 Mídias, política e eleições)**. 2016, Caxambu, Minas Gerais. Anais [...]. Caxambu, MG: ANPOCS, 2016. Disponível em: <https://www.anpocs.com/index.php/encontros/papers/40-encontro-anual-da-anpocs/st-10/st17-8>. Acesso em: 4 abr. 2019.

FONSECA, Virgínia et al. **Revista Comunicação & Inovação**, v. 19, n. 41, p. 67–83, set./dez. 2018. Disponível em: https://seer.uscs.edu.br/index.php/revista_comunicacao_inovacao/article/view/5491/2553. Acesso em: 4 maio 2025.

GRAVES, L. **Deciding what's true: the rise of political fact-checking in american journalism**. 2013. Thesis. (Doctor of Philosophy)-Graduate School of Arts and Sciences, Columbia University, New York, 2013.

HIGUERA, Silvia. **Jornalistas podem confiar na IA para simplificar temas complexos? Este site argentino de fact-checking quer saber**. LatAm Journalism Review, 24 jul. 2024. Disponível em: <https://latamjournalismreview.org/pt-br/articles/jornalistas-podem-confiar-na-ia-para-simplificar-temas-complexos-este-site-argentino-de-fact-checking-quer-saber/>. Acesso em: 2 maio 2025.

LUPA. **Um hub de soluções de combate à desinformação**. Disponível em: <https://lupa.uol.com.br/institucional>. Acesso em: 18 fev. 2025.

MANTZARLIS, Alexios. Verificação dos fatos. In: IRETON, Cherilyn; POSETTI, Julie (org.). *Jornalismo, fake news e desinformação: manual para educação e treinamento em jornalismo*. Brasília: UNESCO, 2019. p. 87–102.

NEWMAN, Nic; CHERUBINI, Federica. **Journalism, media, and technology trends and predictions 2025**. Disponível em: <https://reutersinstitute.politics.ox.ac.uk/journalism-media-and-technology-trends-and-predictions-2025>. Acesso em: 20 fev. 2025.

NEWSGUARD. **AI Tracking Center**. Disponível em: <https://www.newsguardtech.com/special-reports/ai-tracking-center/>. Acesso em: 1 maio 2025.

PALACIOS, M. Fake news e a emergência das agências de checagem: terceirização da credibilidade jornalística?. In: MARTINS, M.; MACEDO, I. (ed.). **Políticas da língua, da comunicação e da cultura no espaço lusófono**. Vila Nova de Famalicão: Edições Humus, 2019. p. 77-92. Disponível em: <https://core.ac.uk/download/pdf/268088765.pdf>. Acesso em: 23 fev. 2022.

RÊGO, Ana Regina; BARBOSA, Marialva. **A construção intencional da ignorância: o mercado das informações falsas**. 1 ed. Rio de Janeiro: Mauad X, 2020.

SEIBT, Taís. **Jornalismo de verificação como tipo ideal: A prática de fact-checking no Brasil**. 2019. Orientadora: Virginia Pradelina da Silveira Fonseca. 2019. 265 f. Tese (Doutorado) - Universidade Federal do Rio Grande do Sul, Porto Alegre. Disponível em: <http://hdl.handle.net/10183/193359>. Acesso em: 29 nov. 2022.

SEIBT, Taís; FONSECA, Virginia Pradelina da Silveira. Transparência como princípio normativo do jornalismo: a prática de fact-checking no Brasil. **Revista Intexto**, Porto Alegre, n. 48, p. 1–25, jan./abr. 2019. Disponível em: <https://lume.ufrgs.br/handle/10183/201020>. Acesso em: 29 abr. 2021.

STENCEL, Mark; RYAN, Erica; LUTHER, Joel. **With half the planet going to the polls in 2024, fact-checking sputters**. Disponível em: <https://reporterslab.org/2024/05/30/wi-th-half-the-planet-going-to-the-polls-in-2024-fact-checking-sputters/>. Acesso em: 28 fev. 2025.

WARDLE, Claire. Essential Guide to Understanding Information Disorder. London: First Draft News, 2019. Disponível em: https://firstdraftnews.org/wp-content/uploads/2019/10/Information_Disorder_Digital_AW.pdf?x76701. Acesso em: 8 jan. 2020.

Jornalismo de Dados sob o signo da Inteligência Artificial

Mayane Batista

A Inteligência Artificial, enquanto tecnologia e regime epistemológico, tornou-se um dos grandes vetores de reconfiguração do Jornalismo de Dados. Mas o que é Inteligência Artificial?

Chama-se de Inteligência Artificial aquilo que, nas últimas décadas, parece ocupar o lugar da promessa e do fracasso técnico. Ela sobe e desce no interesse público conforme o financiamento, atravessando estações de entusiasmo e decepção que ficaram conhecidas como os “invernos” e as “primaveras” da IA (Jones, 2008; Russell, 2013, 2019). Quando a máquina não cumpre o que promete, os recursos minguam e a pesquisa desacelera; basta um avanço, mais processamento, mais dados, menos custo, para que o entusiasmo retorne, e com ele, os investimentos e o discurso da inovação. Mas a IA nunca foi uma coisa só. Sempre foi um amontoado de estratégias, algoritmos, tentativas e acidentes.

Não há um momento inaugural limpo, tampouco uma linearidade que leve do fracasso à redenção técnica. O que há são diferentes modos de fazer máquinas aprenderem, ajustarem-se, improvisarem diante do inesperado (Lima, 2020; Cozman, 2018). Hoje, a IA se espalha onde houver dado, ruído e gente com tempo ou dinheiro para treinar um modelo. Se tornou menos uma invenção progressista do que um regime cotidiano, um jeito de lidar com problemas antigos usando ferramentas novas, ainda que nem sempre saibamos muito bem o que estamos realmente entregando à máquina e o que ainda tentamos salvar da nossa própria inteligência.

A Inteligência Artificial (IA), enquanto subárea da ciência da computação, busca construir sistemas capazes de realizar tarefas outrora restritas ao domínio humano. Não há consenso sobre o que seria, em termos estritos, uma “inteligência de nível humano”, tampouco sobre a definição última de uma Inteligência Artificial (Coppin, 2004; McCarthy, 2007; Jones, 2008). Esse desacordo se mantém como

uma espécie de ruído de fundo permanente no debate desde seu advento.

No plano tecnoprático, convencionou-se entender inteligência como a faculdade de lidar com situações inéditas, resolver problemas e, não menos importante, fazer perguntas. Uma IA, então se ocupa de executar tarefas complexas, assim sendo, identificar padrões, especializar-se em processamentos simbólicos, reconhecer imagens, aprender e manipular atividades inspiradas no pensamento humano (Weizenbaum, 1976; Rich, 2010; Poole, 2010; Rosa, 2011).

Certas abordagens, em vez de simular uma mente adulta, treinam arquiteturas que mimetizam o desenvolvimento cognitivo de uma criança, permitindo que a máquina percorra, a seu modo, uma trajetória de aprendizado. À medida que a IA se aproxima do mundo concreto, as próprias definições do campo se expandem, tornando-se mais plurais, elásticas e contingentes (Luger, 2013; Russell & Norvig, 2013, 2022).

O cenário dos estudos em Inteligência Artificial ganhou novos contornos em 1950, quando o artigo seminal de Alan Turing introduziu uma dimensão até então inexplorada à computação, a capacidade das máquinas de simular processos de pensamento, experimentada no célebre jogo da imitação. Nesse experimento, três participantes, identificados como A, B e C, eram desafiados por um interrogador a distinguir entre humano e máquina. Turing foi provavelmente o primeiro a vislumbrar que as investigações sobre Inteligência Artificial avançariam mais programando computadores do que tentando construir máquinas físicas (McCarthy, 2007). Sua indagação se, com base no desempenho no jogo, seria possível considerar uma máquina pensante não apenas deslocou as fronteiras do que se entendia por

“ Não se trata, portanto, de hierarquias fixas entre humanidade, natureza e máquina, mas de um campo de interdependências e sobreposições. Inteligências culturais, naturais e artificiais não se excluem, mas se reinventam mutuamente em uma espécie de fractal dinâmico”

computação, como lançou os fundamentos para o campo que viria a ser reconhecido como Inteligência Artificial (Lima & Albuquerque, 2018; Lima & Rodrigues, 2019).

Embora Turing tenha sido responsável por popularizar os debates em torno da inteligência das máquinas, o termo "Inteligência Artificial" só seria cunhado em 1956 por John McCarthy, durante a conferência de Dartmouth, episódio frequentemente tratado como o marco inaugural da área. Poucos anos depois, em 1966, Joseph Weizenbaum desenvolveu o programa ELIZA, inspirado em interações com psicoterapeutas, capaz de conduzir conversas plausíveis com usuários a partir das informações recebidas.

Se hoje ELIZA nos parece um protótipo rudimentar do que agora chamamos de *chatbot*, à época sua capacidade de reconhecer padrões em linguagem natural e responder de forma plausível surpreendeu diversos especialistas e leigos no MIT, convencendo muitos de que um novo limiar havia sido cruzado. Paradoxalmente, foi justamente o realismo das interações com ELIZA (antepassada do Chatgpt) em sua habilidade de simular uma conversa, que levou Weizenbaum a se tornar, mais tarde, um crítico contundente das ambições atribuídas à Inteligência Artificial (Jones, 2008, p. 7).

Viveiros de Castro (2024) propõe que inteligência não se deixa capturar por categorias estanques, mas pode ser pensada como uma tríade interligada, sendo elas a cultural, natural e a artificial. O que se observa, nesse modelo, é menos uma separação entre domínios do que um jogo de justaposições, deslocamentos e disputas por centralidade. Essas inteligências se entrelaçam, por vezes colaboram, por vezes colidem, em uma dinâmica que mais lembra zonas de contato do que fronteiras bem demarcadas. A cada instante, há uma disputa ou nem tanto pela primazia epistemológica e ontológica, cada inteligência tenta se afirmar como continente, relegando as demais à condição de províncias. Mas essa tensão, longe de ser destrutiva, pode ser pensada como produtiva; é nesse atrito que surgem novas configurações do saber (Viveiros de Castro, 2024). Não se trata, portanto, de hierarquias fixas entre humanidade, natureza e máquina, mas de um campo de interdependências e sobreposições. Inteligências culturais, naturais e artificiais não se excluem, mas se reinventam mutuamente em uma espécie de fractal dinâmico.



A IA não se sustenta sem a cultura e os dados humanos que a alimentam; por outro lado, torna-se cada vez mais difícil conceber a experiência humana desligada da mediação das máquinas. Não há um lado de cá e um lado de lá, mas uma camada sobrepondo a outra, como já antecipava McLuhan (1964) ao pensar as tecnologias como extensões do humano. Agora, a IA é não só ferramenta, mas zona de interpenetração entre mundos, regimes de existência e modos de conhecimento.

Jornalismo de Dados entre ferramentas e a construção do fato

Se a Inteligência Artificial delinea múltiplas camadas de sentido atravessando regimes de existência técnico, natural e cultural, sua infiltração no Jornalismo de Dados intensifica as ambiguidades e potencializa novas formas de ver, narrar e interrogar o mundo. O Jornalismo de Dados, já é um campo que depende de mediações maquinicas: extrair padrões, cruzar bancos de dados, construir visualizações, automatizar rotinas e lidar com volumes informacionais inabarcáveis para o olho e o gesto humano. O que muda com a IA é o grau de autonomia dessas operações, a escala do processamento e o horizonte de possibilidades para investigação, análise e narrativa.

Como argumenta Bounegru e Gray (2021), esse campo se configura como um "experimento coletivo" que, torna mais porosas e indaga continuamente, que dados, de quem, para quê e com quais ferramentas? O dado não é dado, mas feito e a própria ideia de dado carrega disputas, escolhas, apagamentos e potencialidades que se desdobram em práticas, visualizações, narrativas e controvérsias públicas. A questão central, portanto, é menos sobre a aplicação de ferramentas e mais sobre o regime de práticas, como os jornalistas manipulam bancos de dados, *scripts*, APIs e softwares livres para investigar corrupção, desmatamento, desigualdades ou epidemias, de que maneira a objetividade jornalística é rearticulada nos métodos e na replicabilidade das análises e, sobretudo, como o fato jornalístico se torna construção situada e negociada.

Spotlight: arquitetura algorítmica de investigação contemporânea

Tomemos como exemplo, a investigação do caso Spotlight, investigação jornalística que revelou a rede de abusos sexuais cometidos por membros da Igreja Católica em Boston e serviu de base para o filme Spotlight, vencedor do Oscar. Se ocorresse hoje, o ponto de partida seria, como no início dos anos 2000, acesso ao acervo físico, processos judiciais, listas de transferências, cartas internas, documentos que indisponíveis, mas cuja abertura dependia de batalhas jurídicas travadas por Walter Robinson e sua equipe do Boston Globe.

Após acesso aos documentos, a digitalização do acervo físico, processos judiciais, listas de transferências, cartas internas, tudo seria convertido em texto por tecnologias de reconhecimento óptico de caracteres, como o Tesseract. Esses documentos digitalizados formariam uma base pesquisável, para a etapa seguinte, ferramentas de processamento de linguagem natural, como spaCy, NLTK ou mesmo modelos como o ChatGPT, poderiam ser utilizados. Depois, os sistemas vasculhariam as milhares de páginas, extraíndo nomes de padres, das vítimas, dos bispos, localidades, datas, além de detectar relações explícitas ou implícitas e identificar padrões linguísticos que sugerissem acobertamento, tudo isso dependeria do prompt e do treinamento do modelo porque tais documentos possuem termos eufemísticos e linguagem ambígua.

Com os dados extraídos e estruturados, seria possível alimentar bancos de dados, planilhas, nos quais cada evento, pessoa ou localidade estaria associado a documentos, datas e contextos, após esse procedimento se utilizaria ferramentas de limpeza e padronização, como o OpenRefine ou bibliotecas de análise de dados como o Pandas, permitiriam eliminar duplicidades e inconsistências. Uma vez estruturados, esses dados poderiam ser cruzados e analisados por softwares

“ A IA é não só ferramenta, mas zona de interpenetração entre mundos, regimes de existência e modos de conhecimento ”

de análise de grafos, como Neo4j, dessa forma as conexões entre padres, vítimas, superiores hierárquicos e trajetórias institucionais, seria demonstrado ao digitar e verificar os padrões recorrentes de transferência, zonas de concentração de denúncias e possíveis redes de proteção interna, as visualizações ficariam a cargo de ferramentas como Flourish ou Gephi, mapas de transferências, clusters temporais e geográficos, fluxos de denúncias e dinâmicas de deslocamento. Além disso, a busca automatizada em textos, com ferramentas como Elasticsearch que facilitaria a checagem cruzada de nomes, termos e locais em múltiplos contextos, para a identificação de recorrências e vínculos suspeitos.

Todo esse aparato técnico, entretanto, serve sobretudo como suporte ao trabalho investigativo de jornalistas, e não como substituição. Repórteres utilizariam as redes, mapas e listas geradas pela Inteligência Artificial para orientar as entrevistas, buscar confirmações, identificar lacunas e aprofundar as hipóteses geradas. Por mais sofisticados que sejam os algoritmos, é na escuta, no contato com as pessoas vítimas, fontes, interlocutoras e a análise dos contextos que permanecem como tarefas insubstituíveis de jornalistas. O fator humano ainda é o que confere sentido e responsabilidade ao uso de qualquer tecnologia.

REFERÊNCIAS

BOUNEGRU, L., & GRAY, J. (Eds.). **The Data Journalism Handbook Towards a Critical Data Practice**. Amsterdam University Press. 2021

COPPIN, B. **Artificial intelligence illuminated**. 1st ed. Jones and Bartlett Publishers, 2004.

COZMAN, Fabio. **Inteligência Artificial: uma utopia, uma distopia**. TECCOGS: Revista Digital de Tecnologias Cognitivas, São Paulo, n. 17, p. 32-42, jan./jun. 2018.

JONES, T. **Artificial Intelligence: A Systems Approach**. Jones and Bartlett Publishers, 2008.

LIMA, Mayane Batista. **Perspectivismo Maquínico**: Sobre um ponto de vista heurístico concernente aos ecossistemas comunicacionais. Orientador: Renan Albuquerque. Co Orientadora: Denize Piccolotto Carvalho. 74f. Dissertação (Mestrado em Ciência da Comunicação) – Universidade Federal do Amazonas (UFAM), 2020.

LIMA, MB. ; RODRIGUES, R. A. **Perspectivismo maquínico à luz dos ecossistemas comunicacionais**. Mayane Batista Lima e Renan Rodrigues, RELEM – Revista Eletrônica Mutações, janeiro – julho, 2018.

LIMA, MB; ALBUQUERQUE, R. **Perspectivismo Maquínico e Ecossistemas Comunicacionais**, Mayane Batista Lima em: Olhares comunicacionais - Renan Albuquerque, Noélio Martins Costa e Georgio Ítalo Oliveira,, Alexa Cultural: São Paulo, Manaus: EDUA, 2019.

LUGER, George R. **Inteligência Artificial**. 6 ed. São Paulo: Pearson Education do Brasil, 2013.

MCCARTHY, John. **What Is Artificial Intelligence?** Computer Science Department Stanford University, 2007. Acesso em: 22 dez. 2022.

MCLUHAN, Marshall. **Understanding Media**: The Extensions of Man, McGraw-Hill Book Company 1964.

POOLE, D. L.; MACKWORTH, A. K. **Artificial Intelligence**: Foundations of Computational Agents. Cambridge University Press, 2010.

RICH, E.; KNIGHT, K. **Artificial Intelligence**. 3.ed. s. l.: McGraw Hill, 2010.

ROSA, João Luís Garcia. **Fundamentos da Inteligência Artificial**. Rio de Janeiro: LTC, 2011.

RUSSELL, S. Human Compatible: **Artificial Intelligence and the Problem of Control**. Penguin Books, 2019.

RUSSELL, S.; NORVIG, P. **Inteligência Artificial**. Tradução Regina Célia Simille. 3rd ed. Rio de Janeiro: Elsevier, 2013.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence**: A Modern Approach. 4th Edition. Pearson Education, 2022.

USBANDS, Philip; HOLLAND, Owen. **The Ratio Club**: A Hub of British Cybernetics. In: The Mechanical Mind in History, edited by Philip Husbands, Owen Holland, and Michael Wheeler, The MIT Press, 2008.

VIVEIROS DE CASTRO, Eduardo. **Máquinas sobrenaturais e outros habitantes da tríplice fronteira antropológica**. Aion Journal of Philosophy and Science, n. 1, 2024.

WEIZENBAUM, Joseph. **Artificial Intelligence**. In: Computer Power and Human Reason: From Judgement to Calculation. W.H.Freeman & Co Ltd, 1976.

Perspectivas profissionais: desafios e oportunidades do mercado de trabalho para jornalistas de dados

Reinaldo Chaves

O Jornalismo de Dados emerge como uma especialidade cada vez mais crucial na contemporaneidade, oferecendo ferramentas e métodos para transformar a crescente oferta de dados no mundo em narrativas compreensíveis, relevantes e de impacto social. No Brasil, esta área está em franca expansão, refletindo tanto transformações estruturais no jornalismo quanto a necessidade de novas competências para lidar com a convergência digital (Mick, 2015).

Este capítulo propõe-se a ser um guia para estudantes e jovens jornalistas, explorando o panorama do mercado de trabalho nacional, as posições mais comuns e os requisitos essenciais, os desafios inerentes e as vastas oportunidades.

Tomamos como base estudos recentes sobre o perfil do profissional de dados no Brasil (Oliveira & Angeluci, 2019; Träsel, 2014), os desafios no uso de dados abertos governamentais (Silva et al., 2023), as novas dinâmicas de trabalho (Mick, 2015) e as implicações éticas de novas tecnologias como a Inteligência Artificial (García de Torres et al., 2025), contextualizando-os com tendências globais como as identificadas pelo "The State of Data Journalism Survey 2023" (EJC, 2023). Além disso, usamos algumas entrevistas com proeminentes jornalistas de dados do Brasil.

O objetivo é oferecer uma visão didática e prática das perspectivas que se abrem, num campo que se move das ciências sociais à Inteligência Artificial.

Onde residem as oportunidades

O mercado para jornalistas de dados no Brasil é dinâmico e multifacetado. As oportunidades não se limitam às redações tradi-

cionais; estendem-se a novas configurações midiáticas e até mesmo a organizações não midiáticas (Mick, 2015). Podemos falar em veículos tradicionais, como grandes jornais, emissoras e portais, que continuam sendo atores relevantes, com a criação de editorias e núcleos especializados, como o Estadão Dados (Träsel, 2014). Estes espaços buscam profissionais capazes de agregar valor às reportagens por meio de análises e visualizações.

Também são importantes as mídias nativas digitais e independentes. Este segmento tem sido particularmente inovador no uso de Jornalismo de Dados para investigações de fôlego e cobertura de nichos. Veículos como Agência Pública, Repórter Brasil, e Fiquem Sabendo demonstram a potência do jornalista de dados para a fiscalização e a produção de conhecimento aprofundado.

As Organizações Não Governamentais (ONGs), Think Tanks e Setor Público também têm um papel. A necessidade de comunicar pesquisas, monitorar políticas públicas e promover a transparência tem aberto frentes para jornalistas com habilidades em dados nestas instituições. Mick (2015) já apontava a arregimentação de saberes típicos do jornalismo por empresas e organizações não midiáticas.

Podem ser citados também *freelancing*, consultoria e novas configurações, isto é, a especialização em Jornalismo de Dados permite a atuação autônoma, prestando serviços para diversas mídias ou projetos. Surgem também novas funções ligadas à convergência digital, como analista de mídias sociais, pesquisador de conteúdo e gestor de canais online (Mick, 2015).

Cita-se ainda as colaborações. Embora a implementação prática seja um desafio, há uma atitude positiva em relação a projetos colaborativos entre diferentes redações e com a audiência, especialmente em contextos locais (García de Torres et al., 2025).

“ A exigência
demasiada para
o jornalista de dados
ser multitarefa pode
levar à sobrecarga e à
precarização, dificultando
o aprofundamento
necessário para o
Jornalismo de Dados
de qualidade”

A pesquisa *Perfil profissional do jornalismo brasileiro* de 2012 já indicava que cerca de 9,5% dos jornalistas de mídia e 10,9% dos que atuavam fora dela ocupavam novas denominações de funções, muitas delas ligadas ao ambiente digital (Mick, 2015). A convergência digital também intensificou a exploração do trabalho, exigindo produção para múltiplas mídias simultaneamente (Mick, 2015). O estudo também apontava uma tendência de reconfiguração profissional que foi drasticamente aprofundada na década seguinte, como revelam os dados de 2021 (Barros, Nicoletti & Lima, 2023). Se antes cerca de 10,9% dos jornalistas que atuavam fora da mídia ocupavam novas funções ligadas ao ambiente digital, a pesquisa mais recente sobre o perfil do jornalista brasileiro mostra que a produção de conteúdo para meios digitais e o gerenciamento e monitoramento de redes sociais se consolidaram como as atividades mais frequentes para este grupo, logo após a assessoria de imprensa.

Da mesma forma, a intensificação do trabalho pela convergência digital evoluiu para um quadro de precarização estrutural, onde a exigência de produção para múltiplas mídias se tornou a norma para o “jornalista multifunção”, com 42,2% dos profissionais reportando jornadas superiores a oito horas e um crescimento alarmante de vínculos de trabalho precários (freelancers, PJ e MEI), que já representam 24% das formas de contratação.

Decifrando os códigos do mercado

O “jornalista de dados” é um profissional híbrido, cujas funções podem variar enormemente. Oliveira & Angeluci (2019) identificam um conjunto de competências e habilidades chave, enquanto Träsel (2014) explora o ethos desse profissional, marcado pela tecnofilia (forte atração ou fascínio pela tecnologia) e cooperativismo, mas ainda ancorado no objetivismo jornalístico.

O jornalista de dados pode desempenhar papéis como:

- **Repórter de Dados/Analista de Dados:** profissional focado na apuração, coleta, limpeza, análise e interpretação de dados para produzir reportagens ou insights.

- **Editor de Dados:** lidera projetos, orienta a equipe, valida metodologias e garante a qualidade final das publicações.
- **Desenvolvedor/Designer de Visualização de Dados:** cria as interfaces gráficas, mapas interativos e outras formas de apresentar os dados de maneira clara e engajadora.
- **Especialista em Mineração de Dados, Raspagem e Automação:** focado em extrair e processar grandes volumes de dados de fontes diversas.
- **Cargos emergentes:** Coordenador de mídias digitais, analista de conteúdo digital, gestor de mídias sociais (Mick, 2015).

Para se colocar nesse mercado, o jornalista de dados geralmente atende a alguns requisitos, que de certa forma nortearam a elaboração deste e-book, como recuperamos a seguir:

- **Habilidades jornalísticas fundamentais:** apuração rigorosa, ética, pensamento crítico, curiosidade, storytelling e a capacidade de "fazer as perguntas certas aos dados". O capítulo escrito pela jornalista Beatriz Farrugia explora essa questão mais a fundo.
- **Raciocínio lógico-analítico:** fundamental para entender e aplicar métodos estatísticos e matemáticos, como já abordado no capítulo redigido pela professora Liliane Ito neste e-book.
- **Coleta e tratamento de dados:** domínio da Lei de Acesso à Informação (LAI) e conhecimento de portais de transparência, abordados por Jamile Santana anteriormente, além de técnicas de raspagem de dados (web scraping) e habilidade com planilhas estão no escopo.
- **Análise e estatística:** compreensão de estatística descritiva e inferencial, como noções básicas dessas operações apresentadas por Marcelo Fontoura em capítulo anterior deste e-book.
- **Programação:** conhecimento em linguagens como Python e R é cada vez mais valorizado para análise e automação, embora tal exigência dependa do foco e do propósito de cada jornalista de dados na área, como Judite Cypreste já nos contou no capítulo sobre programação.

- **Visualização de dados:** uso de ferramentas como Datawrapper, Flourish, Tableau e outros recursos para criar narrativas visuais eficazes são diferenciais, como trabalhado por Lucas Thaynan no capítulo de visualização de dados e também por Thays Lavor sobre o uso de mapas.
- **Cultura hacker e colaborativa:** disposição para o cooperativismo, compartilhamento de conhecimento (típico da ética hacker) e trabalho em equipes multidisciplinares (Träsel, 2014; Oliveira & Angeluci, 2019).
- **Soft skills:** resiliência, capacidade de aprendizado contínuo (essencial em um campo em constante evolução), comunicação eficaz, trabalho em equipe e adaptabilidade.

Obstáculos para novos profissionais

Apesar do crescimento, a área de Jornalismo de Dados no Brasil enfrenta obstáculos significativos. Existem pesquisas que apontam a baixa qualidade e acesso a dados governamentais abertos, especialmente em esferas estaduais e municipais. Este é um dos maiores entraves. Silva et al. (2023) apontam que “arquivos desatualizados”, “falta de inteligibilidade” e “informação incompleta” são problemas recorrentes para mais de 60% dos jornalistas entrevistados. Cerca de 94% já desistiram de pautas devido à má qualidade dos datasets. A falta de documentação adequada e formatos impróprios também são barreiras (Silva et al., 2023).

A curva de aprendizado também pode ser um entrave. A formação universitária tradicional muitas vezes não acompanha as demandas do mercado (Oliveira & Angeluci, 2019). O estudo *State of Data Journalism Survey 2023* (EJC, 2023) mostra que 53% dos profissionais globais são autodidatas online, indicando uma lacuna no treinamento formal. Träsel (2014) discute a crise no jornalismo que impulsiona a busca por novas especialidades como o Jornalismo de Dados, ou seja, há uma corrida por conhecimento.

E essas situações também acabam causando uma exigência demasiada para o jornalista de dados ser multitarefa, o que pode levar à sobrecarga e à precarização, dificultando o aprofundamento

necessário para o Jornalismo de Dados de qualidade. Há tarefas no Jornalismo de Dados que exigem conhecimento de matemática, estatística, tecnologia, geografia etc., que em um cenário ideal demandam mais tempo de aprendizagem e testes.

Pelo lado da infraestrutura e cultura organizacional também podem existir problemas. Nem todas as redações investem adequadamente em ferramentas, treinamento e, crucialmente, tempo para investigações com dados. A pressão por produtividade imediata muitas vezes se choca com a natureza mais demorada do Jornalismo de Dados (EJC, 2023; Träsel, 2014). E da mesma forma o jornalista de dados freelancer muitas vezes não têm recursos para investir em toda essa preparação.

Nos avanços recentes de implantar um cenário de colaboração mundial entre jornalistas de dados e no uso de Inteligência Artificial também há desafios. Embora haja interesse, a implementação de colaborações e o uso de IA ainda são baixos, muitas vezes por falta de conhecimento técnico, diretrizes claras ou por preocupações éticas (García de Torres et al., 2025). O uso de IA, apesar de promissor para automação, traz consigo preocupações sobre vieses algorítmicos, transparência, verificação e o papel humano no jornalismo (García de Torres et al., 2025).

Horizontes promissores para o Jornalismo de Dados brasileiro

Mas as dificuldades coexistem com um campo vasto de oportunidades. No âmbito da transparência, o Jornalismo de Dados é uma ferramenta poderosa para a fiscalização do poder público e privado, contribuindo para o combate à corrupção. A própria LAI (Lei de Acesso à Informação), apesar dos desafios em sua aplicação (Silva et al., 2023), já é um instrumento amplamente usado por jornalistas de dados no Brasil para diversas reportagens. Há muito espaço para inovação narrativa e de produto. A capacidade de criar visualizações interativas, aplicativos e novas formas de contar histórias engaja o público e agrega valor ao conteúdo jornalístico (Oliveira & Angeluci, 2019).

As coberturas de temas complexos entram nas possibilidades, já que o Jornalismo de Dados permite desvelar e explicar temas complexos

como políticas públicas, meio ambiente, saúde e economia de forma mais aprofundada e baseada em evidências. Inteligência Artificial e automação oferecem potencial para automatizar tarefas de coleta, processamento e até mesmo geração de partes de narrativas, liberando jornalistas para análises mais complexas e investigações (García de Torres et al., 2025). Conforme o

State of Data Journalism Survey 2023, embora a adoção de IA ainda seja incipiente - cerca de 1/3 dos jornalistas globais a utilizaram -, há um grande interesse em desenvolver habilidades nessa área.

Dentro do empreendedorismo e novos modelos de negócio, o conhecimento em Jornalismo de Dados pode ser a base para a criação de novos veículos, agências de dados, ou consultorias especializadas. Por isso, também pode existir uma valorização do profissional híbrido - profissionais que combinam habilidades jornalísticas com competências em tecnologia, análise de dados e visualização são cada vez mais demandados e valorizados. O ethos cooperativo e a tecnofilia identificados por Träsel (2014) são características desse novo perfil.

O jornalista de dados, imbuído de um espírito investigativo e de uma crença na capacidade da tecnologia para desvendar verdades, não apenas analisa planilhas, mas “entrevista” os dados. Ele busca superar as limitações do jornalismo declaratório, mergulhando em bases de dados para extrair fatos e narrativas que, de outra forma, permaneceriam ocultas, agindo com um misto de curiosidade técnica e compromisso com o objetivismo (Träsel, 2014).

“ Há muito espaço para inovação narrativa e de produto. A capacidade de criar visualizações interativas, aplicativos e novas formas de contar histórias engaja o público e agrega valor ao conteúdo jornalístico”

Desbravando o caminho

Para estudantes e jovens profissionais que almejam uma carreira em Jornalismo de Dados, há diversas maneiras para começar ou se aprimorar. Sempre é importante investir em formação contínua e au-

todidatismo, dada a rápida evolução da área e as lacunas na formação tradicional, o aprendizado contínuo é essencial (Oliveira & Angeluci, 2019). Uma opção é buscar cursos online (Escola de Dados, Abrají, Knight Center), congressos e conferências de jornalismo, tutoriais, e explore a vasta documentação disponível.

No aprendizado de Jornalismo de Dados, sem dúvida, existe um grupo de competências tidas como desejáveis, das mais simples até mais avançadas:

- Entenda os fundamentos de estatística e matemática para não apenas usar dados, mas compreendê-los e interpretá-los corretamente.
- Domine planilhas (Excel, Google Sheets) e familiarize-se com ferramentas de visualização (Datawrapper, Flourish, Tableau).
- Comece com SQL, Python ou R. Existem muitos recursos gratuitos para iniciantes na programação.

Ter atenção para construir um portfólio prático é importante também. Desenvolva projetos pessoais utilizando dados abertos. Analise dados de sua cidade, investigue temas de seu interesse e publique suas descobertas em plataformas como Medium, GitHub Pages ou blogs pessoais.

Engaje-se com a comunidade de dados. Participe de eventos (Coda. Br, Congresso da Abrají), hackathons, grupos de discussão online e siga profissionais da área. O networking e a troca de conhecimento são fundamentais.

Domine a Lei de Acesso à Informação (LAI), entenda seus direitos e deveres e utilize a LAI como ferramenta de apuração. Cultive o pensamento crítico sobre dados, lembre-se que dados não são neutros. Questione suas fontes, metodologias de coleta e possíveis vieses (Silva et al., 2023; Träsel, 2014).

Vale sempre ressaltar: a ética do jornalismo deve vir em primeiro lugar. Especialmente com o avanço da IA, a reflexão ética sobre coleta, uso, interpretação e apresentação de dados é primordial (García de Torres et al., 2025).

O que dizem profissionais da área

Para escrever este capítulo, realizei algumas entrevistas com cinco jornalistas de dados de diferentes idades e regiões do Brasil, além de um profissional que já migrou do jornalismo de redação para outro setor graças às possibilidades do Jornalismo de Dados circular entre vários temas. A intenção não é produzir uma amostra científica ou estatística do mercado, mas oferecer opiniões mais atuais e práticas.

Sobre formação, há consenso de que a carreira em Jornalismo de Dados exige qualificação além da academia. Luiz Fernando Toledo, por exemplo, recomenda fazer muitos cursos online e estudar inglês. Isso reforça a necessidade de autoaprendizado e do domínio da língua inglesa, já que muitos materiais estão disponíveis apenas nesse idioma.

Em relação ao mercado de trabalho, a demanda por jornalistas de dados nas redações é restrita. Rodrigo Menegat considera que a oferta já foi melhor. Quando surgem vagas ou projetos envolvendo dados, o reconhecimento profissional e financeiro nem sempre acompanha a qualificação exigida. Além disso, a maioria dos entrevistados relataram dificuldade para reter talentos técnicos, pois quem domina programação frequentemente migra para o setor de tecnologia em busca de salários melhores, o que não deixa de ser uma oportunidade para jornalistas que se aprimoram nessas técnicas. Esse cenário de pouca valorização e vagas restritas em veículos jornalísticos, de certa forma, reforça desafios estruturais da própria profissão.

A familiaridade com múltiplas áreas técnicas ressaltada pelos entrevistados confirma o perfil multidisciplinar do jornalista de dados. Sobre os principais desafios, os entrevistados destacaram a falta de cultura de dados entre as chefias. Uma das jornalistas entrevistadas, comenta que a direção nem sempre acompanha ou entende o modo de produção do jornalista de dados. Essa percepção reforça a dificuldade de integrar esses especialistas nas pautas diárias.

Apesar desses percalços, os jornalistas ressaltam que sempre há oportunidades no campo. Julie Pereira destaca que persistência é importante e que é possível aprender. Unir capacitação contínua e habilidades técnicas ao espírito jornalístico e à resiliência é a lição compartilhada pelos colegas de profissão para prosperar no Jornalismo de Dados em um mercado em constante transformação.

PARA SABER MAIS

Manuais e livros

- **Manual de Jornalismo de Dados** - tradução brasileira do The Data Journalism Handbook (European Journalism Centre e Google News Initiative). Disponível em: <https://datajournalism.com/read/handbook/two>
- **Guia completo do CruzaGrafos e investigações baseadas em evidências**. Disponível em: https://drive.google.com/file/d/1GMBYzqpFg4TkJ_NCevAiMO4f1uVF9pFx/view?usp=sharing
- **Rumo a uma prática crítica dos dados** - edição revista pela Escola de Dados, com capítulos sobre ética algorítmica, racismo e clima. Disponível em: <https://escoladedados.org/ebooks/>
- **Fluxo de trabalho com dados: do zero à prática** - guia passo-a-passo para todo o data-pipeline, também da Escola de Dados. Disponível em: <https://escoladedados.org/ebooks/>
- **Verification handbook (3ª ed.)** – referência sobre checagem de dados, imagens e redes sociais. Disponível em: <https://datajournalism.com/read/handbook/verification-3>

Cursos e plataformas

- **Cursos on-line da Abraji** – edições regulares de Jornalismo de Dados (básico, SQL, LAI etc.). Disponível em: <https://abraji.org.br/formacao>
- **Escola de Dados** – formações abertas, trilha Dados 360 e workshops sob demanda. Disponível em: <https://escoladedados.org/>
- **DataJournalism.com** – comunidade global com MOOCs gratuitos, fóruns e bibliotecas. Disponível em: <https://datajournalism.com/>

- **Knight Center / JournalismCourses.org** – catálogo de 50+ cursos autodirigidos (português, inglês e espanhol). Disponível em: <https://journalismcourses.org/>

Comunidades e eventos

- **Domingo de Dados** - dia especial dentro do Congresso Internacional da Abraji voltado ao Jornalismo de Dados. Disponível em: <https://congresso.abraji.org.br/>
- **Coda.Br** – principal conferência brasileira de Jornalismo de Dados e métodos digitais (anual, híbrida). Disponível em: <https://escoladedados.org/coda/>
- **NICAR** – maior conferência internacional de dados (IRE, março em Minneapolis em 2025). Disponível em: <https://schedules.ire.org/nicar-2025>
- **Sigma Awards Slack** – hub mundial para troca de técnicas e projetos de data journalism. Disponível em: <https://www.sigmaawards.org/>
- **Data Visualization Society Slack** – 14 mil+ membros trocando dicas de visualização, design e storytelling. Disponível em: <https://www.datavisualizationsociety.org/slack-community>

Podcasts

- **Pizza de Dados** – primeiro podcast brasileiro sobre ciência e dados; episódios sobre Jornalismo de Dados e carreira. Disponível em: <https://podcast.pizzadedados.com/>
- **The Data Journalism Podcast** – Alberto Cairo, Simon Rogers & Scott Klein entrevistam repórteres de dados do mundo todo. Disponível em: <https://podcasts.apple.com/us/podcast/the-data-journalism-podcast/id1564666308>

- **Newsroom Robots** – debates semanais sobre IA e automação em redações, com especialistas globais. Disponível em: <https://podcasts.apple.com/us/podcast/newsroom-robots/id1681331324>
- **Data Hackers Podcast** – comunidade brasileira de analytics; episódio 68 detalha projetos de data journalism no Estadão. Disponível em: <https://www.datahackers.news/p/o-podcast-de-hoje-sobre-jornalismo-de-dados-no-estado>

REFERÊNCIAS

BARROS, J. V. et al. (org.). **O trabalho de jornalistas no Brasil: Desigualdades, identidades e precariedades**. Florianópolis: Editora Insular, 2023.

EJC. The state of data journalism survey 2023. [S.l.]: EJC, 2023. Disponível em: <http://datajournalism.com>. Acesso em: 31 out. 2025.

GARCÍA DE TORRES, E. et al. **The use and ethical implications of artificial intelligence, collaboration, and participation in local Ibero-American newsrooms**. *Frontiers in Communication*, [S.l.], v. 10, 2025.

MICK, J. **Trabalho jornalístico e convergência digital no Brasil: um mapeamento de novas funções e atividades**. *Revista Pauta Geral - Estudos em Jornalismo*, [S.l.], v. 2, n. 1, p. 15-37, 2015.

OLIVEIRA, A. P.; ANGELUCI, A. C. B. **Competences and skills in data journalism: perceptions on the profile of brazilian professionals**. *Brazilian Journalism Research*, [S.l.], v. 15, n. 2, p. 380-399, 2019.

SILVA, S. P. et al. **Jornalismo de Dados e a qualidade dos Dados Abertos governamentais no Brasil: problemas, desafios e obstáculos nos processos de apropriação profissional**. *Revista Compolítica*, [S.l.], v. 12, n. 2, p. 5-32, 2023.



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

Conheça autores e autoras deste e-book

Adriano Belisario

Pesquisador e jornalista investigativo com ampla experiência no uso de inteligência de código aberto (OSINT). Mestre em Social Data Science pela Universidade de Oxford e em Comunicação Social pela Universidade Federal do Rio de Janeiro, atua na interseção entre comunicação, tecnologias abertas e ciência de dados. Já realizou trabalhos para organizações nacionais e internacionais como Open Knowledge Brasil, UNHCR, Bellingcat, Forensic Architecture e Agência Pública. Atualmente, atua como consultor para organismos internacionais e como pesquisador no NetLab/UFRJ. Mais informações em: <https://belisario.website>.

Bárbara Libório

Jornalista especializada em investigação e dados. É co-diretora do Instituto AzMina, coordenadora do MBA em Jornalismo de Dados do IDP e professora na ESPM-SP. Foi editora do Aos Fatos, da Época e do Canal Meio. Como repórter, passou por IstoÉ, iG e Folha de S.Paulo. É doutoranda em Comunicação pela Universidade Metodista de São Paulo e mestre em Mídias Criativas pela Universidade Federal do Rio de Janeiro. Ganhou prêmios internacionais como o Prêmio de Jornalismo Gabriel García Márquez e nacionais como o Cláudio Weber Abramo de Jornalismo de Dados. É fellow do Edward Murrow Program e foi bolsista da Fundación Carolina.

Beatriz Farrugia

Jornalista de dados, professora e pesquisadora. Atua com investigações on-line e com projetos de combate à desinformação. É coautora do livro **1950: O Preço de Uma Copa** (Ed.Letras do Brasil) e fundadora do projeto **O Preço de Uma Copa** (www.oprecodeumacopa).

com), voltado à abertura de dados sobre os Mundiais de futebol da FIFA no Brasil. Beatriz tem mestrado em Jornalismo de Dados pela Birmingham City University (BCU) e é professora na Fundação Armando Alvares Penteado (FAAP), em São Paulo, e na University of Derby, na Inglaterra. Com mais de 15 anos de carreira, tem passagens por Agência ANSA, Digital Forensic Research Lab (DFRLab), Debunk.org e Lupa.

Francisco Amorim

Jornalista, mestre e doutor em Sociologia pela UFRGS. É bolsista de pós-doutorado empresarial do Programa Doutor Empreendedor (Fapergs/Sebrae) na mesma universidade. Fundou a Soda Numérica, ateliê especializado em aplicar inteligência artificial e ciência de dados ao jornalismo, com parcerias com Agência Lupa, Fiquem Sabendo e The Intercept Brasil. Com mais de 25 anos de profissão, atuou como repórter em Zero Hora e no Grupo Sinos e foi editor-chefe da newstech franco-brasileira Headline. Colaborador convidado do Pulitzer Center e do jornal britânico The Guardian. Como docente, foi coordenador do Comitê de Ética em Pesquisa da UniRitter, onde lecionou as disciplinas Metodologia de Pesquisa e Jornalismo de Dados. Integra os grupos de pesquisa Violência e Cidadania e Jornalismo Digital (JorDi/UFRGS), onde desenvolve investigações sobre narcotráfico e territorialidade, jornalismo de dados e inteligência artificial.

Jamile Santana

Jornalista de dados e ativista em Transparência Pública em municípios. Com 11 anos de experiência em jornalismo diário, atuou em redações como o G1 - Portal de Notícias da Globo no Alto Tietê, Portal Terra, entre outros. Já coordenou a Escola de Dados, programa de alfabetização em dados da Open Knowledge Brasil e atualmente é gestora de Transparência e Proteção de Dados na Prefeitura de Mogi das Cruzes. Também é fellow do Programa de Governo Aberto e Mudanças Climáticas da OEA/CGU (2025). Em 2021, recebeu a medalha Zumbi dos Palmares, por sua atuação ativa em transparência pública e redação da minuta do projeto de lei que regulamentou a Lei de Acesso à Informação em Mogi das Cruzes (SP).

Judite Cypreste

Jornalista de dados com mais de sete anos de experiência em investigações baseadas em dados públicos, com foco em reportagens sobre transparência, segurança pública e direitos humanos. Já foi reconhecida por prêmios nacionais e internacionais, como o Sigma Awards. Passou por redações do país como o Portal G1, Metrôpoles, GloboNews, UOL, Aos Fatos e Folha de S.Paulo. Atuou também no setor público, na Prefeitura do Rio de Janeiro, onde foi responsável por criar visualizações e ferramentas a partir de dados públicos da cidade. Foi fundadora do Colaboradados, primeiro veículo de dados abertos voltada à transparência pública no Brasil. Lecionou jornalismo de dados na Universidade de Brasília e foi convidada pela Enap para capacitar servidores no uso estratégico de dados. É especialista em jornalismo de dados pela Columbia University.

Liliane de Lucena Ito

Jornalista e professora do Departamento de Jornalismo da Universidade Estadual Paulista (Unesp). Professora do Mestrado Acadêmico em Comunicação Digital pelo IDP (Brasília/DF). Pós-doutora em Comunicação e Consumo (ESPM). Doutora e mestre em Comunicação pela Unesp. Vice-coordenadora do GP Comunicação e Desinformação da Intercom.

Lívia de Souza Vieira

Jornalista e professora da Faculdade de Comunicação e do Programa de Pós-Graduação em Comunicação e Cultura Contemporâneas da Universidade Federal da Bahia (UFBA). Vice-Coordenadora do Grupo de Pesquisa em Jornalismo On-Line (GJOL) do Póscom/UFBA. Doutora e mestre em Jornalismo pela Universidade Federal de Santa Catarina (UFSC). Editora da newsletter Farol Jornalismo.

Lucas Thaynan

Jornalista de dados, designer, desenvolvedor Python e programador web. Formado em Jornalismo pela Universidade Federal de



Alagoas (Ufal) e pós-graduado em Jornalismo de Dados, Automação e Data Storytelling, pelo Insper. Cofundador da startup Agência Tatu, que desenvolve conteúdos e produtos jornalísticos voltados para a região Nordeste. Hoje também atua no núcleo da Editoria de Artes do jornal Estadão, com enfoque na produção de reportagens visuais baseadas em dados.

Marcela Canavarro

Jornalista de dados no DeltaFolha, editoria de Jornalismo de Dados da Folha de São Paulo, já produziu conteúdos em impresso, online, televisão e rádio, em mais de 20 anos no jornalismo. É doutora em Mídias Digitais pela Faculdade de Engenharia da Universidade do Porto, com pesquisa no campo das Ciências Sociais Computacionais e tem pós-doutorado pelo NetLab UFRJ / IBICT / ECO. É mestra em Comunicação e Cultura (Escola de Comunicação da UFRJ) e especialista em Mídias Digitais e Interativas (Senac-Rio). Atualmente, também é pós-graduanda em Ciência de Dados (UTFPR).

Marcelo Crispim Fontoura

Jornalista e professor adjunto de jornalismo de dados e estudos de desinformação na PUCRS. Também atua como pesquisador da região Sul para o Atlas de Notícias e como avaliador para a International Fact Checking Network. Trabalha como consultor para projetos de inovação em jornalismo digital, tendo coordenado iniciativas na News Product Alliance, na Meedan e na Ajor. Tem doutorado em comunicação pela PUCRS, focado na expansão das fronteiras do jornalismo, e mestrado em comunicação pela mesma instituição, focado em hackers de dados abertos no Brasil.

Marcelo Soares

Um dos pioneiros do Jornalismo de Dados no Brasil, entrevista planilhas desde o final dos anos 90. Hoje, dirige o estúdio de inteligência de dados Lagom Data. Foi o primeiro editor de audiência e dados da imprensa brasileira, na Folha de S. Paulo. Também teve passagens

pelo Correio do Povo, MTV e Los Angeles Times. Desde 2010, é membro do International Consortium of Investigative Journalists (ICIJ). Foi cofundador e primeiro gerente da Associação Brasileira de Jornalismo Investigativo (Abraji), além de cocriador do projeto Excelências, da Transparência Brasil (prêmio Esso de Melhor Contribuição à Imprensa em 2006). Kursou Jornalismo na Universidade Federal do Rio Grande do Sul (UFRGS) e fez mestrado em divulgação científica e cultural no Labjor-Unicamp; lecionou Jornalismo de Dados em especializações e cursos curtos na ESPM-SP, PUC-RS, UFSC e outras.

Marília Gehrke

Professora de Jornalismo e Estudos de Mídia na Universidade de Groningen, na Holanda, com passagem pelo Centro de Democracia Digital (DDC), na Dinamarca, e estadia de pesquisa no Weizenbaum Institute, na Alemanha. Gehrke fez mestrado e doutorado na Universidade Federal do Rio Grande do Sul (UFRGS), período em que investigou jornalismo guiado por dados e transparência (e também descobriu o feminismo de dados). Atualmente, pesquisa desinformação com foco em gênero e raça. Coordena o projeto Forced to Quit, que mapeia mulheres que tiveram de deixar a política, o jornalismo e o ativismo por causa da violência. É uma das fundadoras do DataJor.

Marta Alencar

Doutoranda em Comunicação pela Unisinos, onde pesquisa desertos de notícias e desinformação local e regional. Mestra em Comunicação pela UFPI. Jornalista. Fundadora da COAR Notícias, primeira iniciativa de letramento midiático com linguagem regional no Brasil. Integra o Grupo de Pesquisa Práxis no Jornalismo (PráxisJor), vinculado à Universidade Federal do Ceará (UFC). Idealizadora dos Manuais de Checagem com linguagem regional no país, distribuídos especialmente em vazios noticiosos. Atualmente é membra da Comissão Editorial da Associação de Jornalistas de Educação (Jeduca) e integra a Rede Nacional de Combate à Desinformação (RNCD).

Mayane Batista Lima

Docente colaboradora do Programa de Pós-Graduação em Informação e Comunicação da Faculdade de Informação e Comunicação da Universidade Federal do Amazonas (PPGIC/FIC/Ufam). Docente do curso de Jornalismo na Faculdade Boas Novas. Doutora em Antropologia Social pela Universidade Federal do Amazonas, com estágio doutoral-sanduíche realizado na Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo (FFLCH-USP). Pesquisadora visitante do Centro de Inteligência Artificial da USP (C4AI-CID) e da Cátedra Oscar Sala (IEA-USP). É Bacharel em Jornalismo e Ciência da Computação.

Polinho Mota

Nascido e criado em Manaus, atualmente mora no Rio de Janeiro onde atua como coordenador de dados e DPO do data_labe, laboratório de dados com foco em gênero, raça e território. Possui mestrado e está no doutorado em Epidemiologia e ciência de dados populacionais na Universidade Federal do Rio de Janeiro (UFRJ). Polinho acredita e trabalha para a democratização das tecnologias, feitas por todas as pessoas de forma livre e soberana.

Samuel Pantoja Lima

Jornalista, professor do curso de Jornalismo da Universidade Federal de Santa Catarina (UFSC) e do Programa de Pós-Graduação em Jornalismo da UFSC (PPGJOR). Doutor em Mídia e Teoria do Conhecimento (UFSC, 2005), é pesquisador do Observatório da Ética Jornalística (objETHOS/UFSC) e do Laboratório de Sociologia do Trabalho (Lastro/UFSC). Coordena as pesquisas *Perfil dos Jornalistas Brasileiros* (2021) e *Novas institucionalidades de serviço público não-estatal para o jornalismo local em Santa Catarina* (2025/2026).

Taís Seibt

Professora no Mestrado em Comunicação Digital e no MBA em Jornalismo de Dados do Instituto Brasileiro de Ensino, Desenvolvimento

e Pesquisa (IDP), e na Escola da Indústria Criativa da Universidade do Vale do Rio dos Sinos (Unisinos). Doutora em Comunicação e Informação pela Universidade Federal do Rio Grande do Sul (UFRGS), é líder do Núcleo de Estudos em Jornalismo de Dados e Computacional - DataJor (CNPq/IDP) e colabora com o Projeto Temático Horus - Técnicas de Inteligência Artificial para Detecção e Análise de Realidades Sintéticas do Instituto de Computação da Universidade Estadual de Campinas (IC-Unicamp). Sua tese de doutorado "Jornalismo de verificação como tipo ideal: a prática de fact-checking no Brasil" recebeu menção honrosa no Prêmio Capes de Tese 2020.

Thays Lavor

Jornalista e pesquisadora, coordena a Escola de Dados, eixo educacional da Open Knowledge Brasil, e integra o conselho fiscal da Abraji. É doutoranda em Comunicação pela Universidade Federal da Bahia (UFBA), vinculada ao Grupo de Pesquisa em Jornalismo On-Line (GJOL), mestra pela Universidade Federal do Ceará (UFC) e especialista em Ciência de Dados pela Esalq/USP. Atua com jornalismo investigativo e de dados, com colaborações e publicações em veículos nacionais e internacionais, e na formação em temas como desinformação, transparência e uso de dados.

Reinaldo Chaves

Coordenador de projetos na Associação Brasileira de Jornalismo Investigativo (Abraji), foi finalista dos prêmios Pulitzer, Sigma Awards em 2023 e 2021 com trabalhos sobre investigações de relações de poder, meio ambiente e saúde pública com o estudo de milhões de dados de interesse público com uso de Python e IA.

