

# § Parágrafo

ISSN 2317-4919

Esta edição da revista Parágrafo representa o fechamento do ciclo de uma proposta editorial caracterizada, desde o segundo semestre de 2014, pela construção de dossiês temáticos - dedicados não apenas à compreensão de fenômenos jornalísticos, mas a questões mais amplas do campo da comunicação, em interface com pesquisadores internacionais. A partir do próximo número, a revista priorizará produções que contemplam o jornalismo como objeto de estudo, buscando aprofundar sua aderência ao programa de Mestrado Profissional em Jornalismo do FIAM-FAAM Centro Universitário, ao qual está vinculada. Esperamos, com isso, contribuir com o avanço da reflexão teórica, metodológica e empírica sobre as práticas jornalísticas, bem como a divulgação de relatos de pesquisa aplicada ao campo profissional, reafirmando a importância do diálogo interdisciplinar neste processo.

**FIAMFAAM**

**Revista do Mestrado Profissional em Jornalismo  
FIAM-FAAM – Centro Universitário**

**v.6, n.1, jan./abr. 2018**



# § Parágrafo

**Revista do Mestrado Profissional em Jornalismo do FIAM-FAAM – Centro Universitário**

São Paulo, Brasil, volume 6, número 1, janeiro/abril de 2018

ISSN: 2317-4919

The logo consists of a solid blue square with the text "FIAMFAAM" in white, uppercase, sans-serif font centered within it.

**FIAMFAAM**

# § Parágrafo

Rua Vergueiro, 101, Liberdade, São Paulo, SP, CEP 01504-001

website: [revistaseletronicas.fiamfaam.br/index.php/recicofi](http://revistaseletronicas.fiamfaam.br/index.php/recicofi)

e-mail: [paragrafo@fiamfaam.br](mailto:paragrafo@fiamfaam.br)

## Comitê Editorial

**Editora:** Cláudia Nonato

**Editora Adjunta:** Michelle Roxo

**Conselho Editorial:** Alciane Baccin, Francisco de Assis, Ivan Paganotti, Juliana Doretto, Silvio Anaz e Vicente Darde

**Conselho Científico:** Adriana Braga (Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, RJ, Brasil), Alice Mitika Koshiyama (Universidade de São Paulo, São Paulo, SP, Brasil), Alan Angeluci (Universidade Municipal de São Caetano do Sul, São Caetano do Sul, SP, Brasil), Álvaro Lorangeira (Universidade Tuiuti do Paraná, Curitiba, PR, Brasil), Ana Carolina Rocha Pessôa Temer (Universidade Federal de Goiás, Goiânia, GO, Brasil), Andreas Hepp (Universität Bremen, Bremen, Alemanha), Ângela Cristina Salgueiro Marques (Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil), Bárbara Heller (Universidade Paulista, São Paulo, SP, Brasil), Bruno Campanella (Universidade Federal Fluminense, Niterói, RJ, Brasil), César Ricardo Siqueira Bolaño (Universidade Federal de Sergipe, Aracaju, SE, Brasil), Cláudia Quadros (Universidade Federal do Pa-

raná, Curitiba, PR, Brasil), Dennis de Oliveira (Universidade de São Paulo, São Paulo, SP, Brasil), Egle Müller Spinelli (Escola Superior de Propaganda e Marketing, São Paulo, SP, Brasil), Eliza Bachega Casadei (Escola Superior de Propaganda e Marketing, São Paulo, SP, Brasil), Eugênio Trivinho (Pontifícia Universidade Católica de São Paulo, São Paulo, SP, Brasil), Fábio Pereira (Universidade de Brasília, Brasília, DF, Brasil), Fernando Firmino da Silva (Universidade Federal da Paraíba, Campina Grande, PB, Brasil), Francisco Rüdiger (Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brasil), Gislene Silva (Universidade Federal de Santa Catarina, Florianópolis, SC, Brasil), Graham Murdock (Loughborough University, Loughborough, Inglaterra), Göran Bolin (Södertörn University, Huddinge, Suécia), Helder Bastos (Universidade do Porto, Porto, Portugal), Ibrahim Saleh (University of Cape Town, Cape Town, África do Sul), Igor Sacramento (Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil), Jairo Ferreira (Universidade do Vale do Rio dos Sinos, São Leopoldo, RS, Brasil), João Carlos Masarolo (Universidade Federal de São Carlos, São Carlos, SP, Brasil), José Guibson Dantas (Universidade Federal de Alagoas, Maceió, AL, Brasil), Laan Mendes de Barros (Universidade Estadual Paulista, Bauru, SP, Brasil), Laura Cánepa

(Universidade Anhembi Morumbi, São Paulo, Brasil), Luciano Victor Barros Maluly (Universidade de São Paulo, São Paulo, SP, Brasil), Luís Mauro Sá Martino (Faculdade Cásper Líbero, São Paulo, SP, Brasil), Luiz Peres Neto (Escola Superior de Propaganda e Marketing, São Paulo, SP, Brasil), Marcelo Kischinhevsky (Universidade Estadual do Rio de Janeiro, Rio de Janeiro, RJ, Brasil), Marco Roxo (Universidade Federal Fluminense, Niterói, RJ, Brasil), Marcos Paulo da Silva (Universidade Federal do Mato Grosso do Sul, Campo Grande, MS, Brasil), Maria Ângela Mattos (Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, MG, Brasil), Maria Angela Pavan (Universidade Federal do Rio Grande do Norte, Natal, RN, Brasil), Maria Aparecida Baccega (Escola Superior de Propaganda e Marketing, São Paulo, SP, Brasil), Maria Immacolata Vassalo de Lopes (Universidade de São Paulo, São Paulo, SP, Brasil), Maria Ogécia Drigo (Universidade de Sorocaba, Sorocaba, SP, Brasil), Mark Deuze (University of Amsterdam, Amsterdam, Holanda), Marko Ampuja (University of Helsinki, Helsinki, Finlândia), Marli dos Santos (Faculdade Cásper Líbero, São Paulo, SP, Brasil), Marta Regina Maia (Universidade Federal de Ouro Preto, Mariana, MG, Brasil), Mayra Rodrigues Gomes (Universidade de São Paulo, São Paulo, SP, Brasil), Muniz Sodré (Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil), Nilda Jacks (Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil), Olga Tavares (Universidade Federal da Paraíba, João Pessoa, PB, Brasil), Oscar Westlund (University of Gothenburg,

Gotemburgo, Suécia), Paulo Roberto Figueira Leal (Universidade Federal de Juiz de Fora, Juiz de Fora, MG, Brasil), Pramod Nayar (University of Hyderabad, Hyderabad, Índia), Rafael Bellan (Universidade Federal do Espírito Santo, Vitória, ES, Brasil), Rafael Grohmann (Faculdade Cásper Líbero, São Paulo, SP, Brasil), Rogério Christofolletti (Universidade Federal de Santa Catarina, Florianópolis, SC, Brasil), Rogério Covalesk (Universidade Federal de Pernambuco, Recife, PE, Brasil), Rosana de Lima Soares (Universidade de São Paulo, São Paulo, SP, Brasil), Rose de Melo Rocha (Escola Superior de Propaganda e Marketing, São Paulo, SP, Brasil), Sérgio Luiz Gadini (Universidade Estadual de Ponta Grossa, Ponta Grossa, PR, Brasil), Sandra Reimão (Universidade de São Paulo, São Paulo, SP, Brasil), Stig Hjarvard (University of Copenhagen, Copenhagen, Dinamarca), Thiago Soares (Universidade Federal de Pernambuco, Recife, PE, Brasil), Valdir José Morigi (Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil), Veneza Mayora Ronsini (Universidade Federal de Santa Maria, Santa Maria, RS, Brasil) e Verônica Dantas (Universidade Federal do Tocantins, Palmas, TO, Brasil). Todos são doutores titulados em diferentes instituições no Brasil no exterior.

**Assistentes Editoriais:** Maria Tereza Blois (preparação de originais e revisão), Gerson Victor dos Santos (editoração eletrônica e programação) e Euclides Santos (projeto gráfico).



# Sumário

- 9 **Editorial**  
Cláudia Nonato e Michelle Roxo
- Dossiê: Mediações Algorítmas**  
*Dossier: Algorithmic Mediation*
- 11 **Apresentação**  
*Presentation*  
Carlos d'Andréa e Willian Araújo
- 17 **Questões e apontamentos para o estudo de algoritmos**  
*Issues and prompts to the study of algorithms*  
Amanda Chevtchouk Jurno e Sílvia DalBen
- 31 **Repensando big data, algoritmos e comunicação: para uma crítica da neutralidade instrumental**  
*Rethinking the big data, algorithms and communication: a critique of instrumental neutrality*  
Tales Tomaz e Guilherme Cavalcante Silva
- 43 **Mediações algorítmicas: o poder de modulação dos algoritmos do Facebook**  
*Algorithmic mediations: the power of modulation of Facebook's algorithms*  
Débora Franco Machado
- 57 **Topologias de espaços híbridos na era da Internet das Coisas**  
*Topologies of hybrid spaces in the Era of Internet of Things*  
Fernanda da Costa Portugal Duarte
- 71 **O uso de algoritmos na mídia programática**  
*Use of algorithmic in the programmatic media*  
Stefanie Carlan da Silveira e João Gabriel Danesi Morisso

- 83 **Um jornalismo para chamar de meu? Algoritmos e o fenômeno da customização de notícias**  
*Is there a journalism to call my own? Algorithm and the phenomenon of news customization*  
Raquel Dornelas
- 95 **A relevância dos algoritmos**  
*The relevance of algorithms*  
Tarleton Gillespie
- 123 **Examinando uma técnica algorítmica: o classificador de bayes como uma leitura interessada da realidade**  
*Scrutinizing an algorithmic technique: the Bayes classifier as interested reading of reality*  
Bernhard Rieder
- 143 **Por que empresas de mídia insistem que não são empresas de mídia, por que estão erradas e por que isso importa**  
*Why media companies insist they are not media companies, why they are wrong, and why it matters*  
Philip Napoli e Robyn Caplan

## **Entrevista**

### *Interview*

- 165 **Algoritmos como um devir: uma entrevista com Taina Bucher**  
Carlos d'Andréa e Amanda Jurno

## **Artigos**

### *Articles*

- 173 **Relações públicas e comunicação organizacional digital: empresas fronteiriças e suas estratégias**  
*Public relations and organizational digital communication: companies on the border and their strategies*  
Henrique Esper e Cristóvão Domingos de Almeida
- 185 **Convergência midiática e cultura participativa: possíveis interações entre novas tecnologias e agentes sociais no campo da comunicação**  
*Media convergence and participatory culture: possible interactions between new technologies and social agents in the field of communication*  
Antonio Francisco Magnoni e Giovani Vieira Miranda

# Editorial

Esta edição da revista **Parágrafo** representa o fechamento do ciclo de uma proposta editorial caracterizada, desde o segundo semestre de 2014, pela construção de dossiês temáticos – dedicados não apenas à compreensão de fenômenos jornalísticos, mas a questões mais amplas do campo da comunicação, em interface com pesquisadores internacionais. A partir do próximo número, a revista priorizará produções que contemplam o jornalismo como objeto de estudo, buscando aprofundar sua aderência ao Programa de Mestrado Profissional em Jornalismo do FIAM-FAAM – Centro Universitário, ao qual está vinculada. Esperamos, com isso, contribuir com o avanço da reflexão teórica, metodológica e empírica sobre as práticas jornalísticas, bem como com a divulgação de relatos de pesquisa aplicada ao campo profissional, reafirmando a importância do diálogo interdisciplinar nesse processo.

**Parágrafo** também recebe novo projeto gráfico, como parte de sua reestruturação editorial. Sua periodicidade torna-se quadrimestral. O fluxo de submissões espontâneas passa a ser priorizado e, com isso, o formato dos dossiês temáticos, adotado nos últimos quatro anos, será reservado a números especiais, organizados em parceria com editores convidados.

Neste número de transição, apresentamos o dossiê “Mediações Algorítmicas: olhares da pesquisa em comunicação e mídia”, organizado por Carlos d’Andréa, da Universidade Federal de Minas Gerais, e Willian Araújo, da Universidade de Santa Cruz do Sul. A eles, agradecemos pela edição cuidadosa apresentada a seguir, com textos e entrevista inéditos, além de artigos internacionais traduzidos sobre a temática em tela.

Agradecemos, em especial, a Rafael Grohmann, que conduziu o trabalho de edição nos últimos anos, articulando uma variedade de olhares e perspectivas teóricas em dossiês com forte presença de autores estrangeiros, os quais contribuíram para a projeção da revista. Nesta nova etapa, desejamos fortalecer a inserção da **Parágrafo** entre o conjunto de periódicos científicos especializados em Jornalismo – em sintonia com a proposta do nosso programa –, aprimorando-a constantemente, e trabalhar pelo progressivo reconhecimento da publicação em estratos do Qualis e em outros indexadores.

*Cláudia Nonato e Michelle Roxo*  
Editoras



# Dossiê

Mediações Algorítmicas:  
olhares da pesquisa em comunicação e mídia

# Apresentação

## Presentation

### **Carlos d' Andréa**

Universidade Federal de Minas Gerais  
Departamento de Comunicação Social  
Belo Horizonte, MG, Brasil

### **Willian Araújo**

Universidade de Santa Cruz do Sul  
Santa Cruz do Sul, RS, Brasil

As mediações algorítmicas têm se tornado, nos últimos anos, um importante tema de discussão para a sociedade contemporânea e para as pesquisas em comunicação e mídia. Ao selecionar, hierarquizar, visibilizar, controlar, os algoritmos instauram lógicas particulares de produção de conhecimento e estão hoje no centro do ecossistema informacional e midiático. A faceta mais visível deste fenômeno é a cada vez mais controversa atuação de plataformas como o Google e o Facebook, mas certamente as questões políticas, éticas, econômicas etc do tema não se esgotam na atuação das gigantes do Vale do Silício.

Nas pesquisas sobre comunicação e mídia, "algoritmo" pode ser reconhecido como um dos objetos empíricos mais ricos no crescente diálogo da área com abordagens teórico-metodológicas que enfatizam a dimensão sociotécnica e performativa dos diferentes atores nos ambientes digitais. Inicialmente muito centrada em apropriações da Teoria Ator-Rede (LATOUR, 2012), aos poucos a produção acadêmica no Brasil vai ampliando seu diálogo com os *Software Studies*, *Science and Technologies Studies (STS)* e outras abordagens ligadas à chamada "virada material" (STERNE, 2014).

Partindo deste cenário, este dossiê procurou reunir pesquisas empíricas e/ou discussões teóricas que abordassem questões como as relações dos algoritmos com práticas do jornalismo, da publicidade, do audiovisual, do entretenimento, da comunicação organizacional etc; as dimensões políticas e éticas das mediações algorítmicas contemporâneas (privacidade, por exemplo); apropriações táticas e de resistência às mediações algorítmicas; os desafios conceituais e metodológicos da pesquisa com algoritmos.

## Traduções e uma entrevista

Antes de apresentar os seis artigos inéditos publicados pelo dossiê e os artigos livres, chamamos a atenção para os três artigos traduzidos e para a entrevista inédita aqui publicada.

"*Algoritmos como um dever: uma entrevista com Taina Bucher*" é fruto de uma conversa com a pesquisadora da Universidade de Copenhagen (Dinamarca) durante a conferência da Association of Internet Researchers (AoIR) 2017. Temas como jornalismo, caixa-preta e a contribuição de autores como Ross Ashby, Alfred N. Whitehead e Annemarie Mol são abordados na entrevista realizada por Carlos d'Andréa e Amanda Jurno. Uma das referências fundamentais para os estudos de algoritmos (vide, por ex, artigo "*Want to be on the top? Algorithmic power and the threat of invisibility on Facebook*", publicado em 2012<sup>1</sup>), Taina Bucher está lançando, em junho de 2018, o livro "*IF... THEN: Algorithmic power and politics*" pela Oxford University Press<sup>2</sup>.

Publicado originalmente em 2014 no livro "*Media Technologies: Essays on Communication, Materiality, and Society*" (MIT Press, 2014), o artigo "*A relevância dos algoritmos*", de Tarleton Gillespie, é possivelmente o trabalho mais citado sobre o tema desde então – inclusive entre os *papers* inéditos

deste dossiê. Ao destacar e detalhar seis dimensões dos "algoritmos de relevância pública" (entre os quais "Ciclos de antecipação" e "A produção de públicos calculados"), o autor refuta as abordagens deterministas centradas na tecnologia e mostra como, na agência dos algoritmos, os discursos, as práticas e as formas de conhecimento se entrelaçam com as lógicas dos códigos de programação. Pesquisador-chefe da Microsoft Research New England e professor da Cornell University (EUA), Gillespie lança, em 2018, o livro "*Custodians of the Internet - Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*" (Yale Press). O artigo "*A relevância dos algoritmos*" foi traduzido por Amanda Jurno.

"*Examinando uma técnica algorítmica: o classificador de Bayes como uma leitura interessada da realidade*" foi publicado em 2017 por Bernhard Rieder (Universidade de Amsterdam) no dossiê "*The Social Power of Algorithms*" da revista "*Information, Communication & Society*"<sup>3</sup>. Mais conhecido por alguns pelo desenvolvimento de aplicativos como Netvizz (para extração de dados do Facebook) e DMI-TCAT (Twitter), Rieder mostra neste artigo uma leitura profundamente embasada sobre formas de conhecimento estatístico usadas recorrentemente no desenvolvimento de sistemas de acesso e classificação de informação. Ao analisar de forma minuciosa a técnica do classificador de Bayes, o autor desenvolve uma proposta que se posiciona entre a teorização ampla sobre algoritmos, predominante nos estudos emergentes das ciências sociais e humanas, e a investigação empírica de aplicações concretas. A tradução foi feita por Willian Araujo e Fernanda Pires de Sá.

A terceira tradução do dossiê é o artigo "*Por que empresas de mídia insistem que não são empresas de mídia, por que estão erradas e por que isso importa*", de Philip M. Napoli (Duke University, EUA) e Robyn Caplan (Data & Society Research Institute). O texto apresenta uma discussão crítica sobre os

1 <https://doi.org/10.1177/1461444812440159>

2 <https://global.oup.com/academic/product/ifthen-9780190493035?lang=en&cc=us>

3 <https://doi.org/10.1080/1369118X.2016.1181195>

discursos proferidos pelas plataformas de mídias sociais, denunciando a tentativa dessas empresas de se posicionar como *meros suportes tecnológicos*, em um sentido determinista e que opõe mídia e tecnologia. Através da discussão de diferentes casos, os autores sugerem que essa não é apenas uma distinção semântica, mas a busca por um espaço de atuação que escape das regulamentações que serviços de comunicação estão sujeitos. Publicado em 2017 na revista *First Monday*<sup>4</sup>, o artigo foi traduzido por Tarcízio Silva.

Bastante diferentes entre si, estes artigos agora disponíveis em português dão uma pequena mas, acreditamos, significativa amostra da intensa produção acadêmica sobre o tema fora do Brasil. Interessados em ir adiante podem encontrar as referências fundamentais na lista "*Critical Algorithm Studies*"<sup>5</sup>, organizada por Tarleton Gillespie e Nick Seaver no site "Social Media Collective". Esta lista, no entanto, não é atualizada desde 2016 - e talvez seja um indício de que é virtualmente impossível dar conta do crescimento exponencial da bibliografia sobre algoritmos e temas afins.

É preciso, também, ir além de uma visada anglo-saxã do tema. São crescentes (mas ainda incipientes) os esforços para situar as discussões sobre algoritmos a partir de "epistemologias alternativas" e os *data activismos* a elas associados (Milan e Velden, 2016), ou ainda na "relação despudorada e inventiva com os objetos técnicos" descrita por Bruno (2017) ao evidenciar as (famigeradas) "gambiaras". Neste movimento, uma das iniciativas a serem destacadas é a "rede latino-americana de estudos sobre vigilância, tecnologia e sociedade" (Lavits), que vem se articulando desde 2009. Outra iniciativa que merece atenção é o projeto "Big Data from the South", sediado na Universidade de Amsterdã - curiosamente, ou

sintomaticamente, um país do Norte Global<sup>6</sup>. Temas e desafios para futuros dossiês...

## Sobre os artigos inéditos

O primeiro artigo inédito do dossiê é "*Questões e apontamentos para o estudo de algoritmos*", de Amanda Chevtchouk Jurno e Silvia Dalben (UFMG). De forma didática, as autoras dialogam com autores-chave da área, como Rob Kitchin, Lucas Introna e Kate Crawford para discutir questões como os "algoritmos são entidades estáveis?". O objetivo aqui é apresentar um "panorama" que auxilie pesquisadores interessados na temática.

De modo completar, o artigo "*Repensando big data, algoritmos e comunicação: para uma crítica da neutralidade instrumental*", de Tales Tomaz (Unasp) e Guilherme Cavalcante Silva (Unicamp), também discute questões conceituais relativas ao tema do dossiê. Para "elucidar a noção etimológica e contextual de *data*", os autores resgatam sua relação com a cibernética, que já nos 1940 associavam comunicação e controle. Em diálogo com fenomenologia de Martin Heidegger, o artigo apresenta uma abordagem crítica aos discursos instrumentais sobre *big data* e algoritmos.

Em "*Mediações algorítmicas: analisando o poder de modulação dos algoritmos do Facebook*", Débora Franco Machado (UFABC) apresenta, através de duas patentes, como a plataforma de mídia social mais popular do mundo se organiza para modularizar as "emoções" de seus usuários. Baseada em Gilles Deleuze, a autora discute as implicações políticas do crescente processo de controle delegado aos algoritmos na sociedade contemporânea.

As relações entre os processos cada vez mais automatizados de coleta, processamento e análise dos dados e as diferentes topologias que (re)criam os espaços híbridos são o tema do artigo de Fernanda da Costa Portugal Duarte (UFMG). Em "*Topologias de espaços híbridos na era da In-*

<sup>6</sup> <https://data-activism.net/big-data-from-the-south/>

<sup>4</sup> <http://firstmonday.org/ojs/index.php/fm/article/view/7051/6124>

<sup>5</sup> <https://socialmediacollecive.org/reading-lists/critical-algorithm-studies/>

*ternet das Coisas*”, a autora apresenta fenômenos como a emergência de modelos preditivos e o monitoramento digital da performance fisiológica. A discussão se insere na “virada espacial” das teorias sociais e do campo das mobilidades.

O artigo seguinte do dossiê chama-se “O uso de algoritmos na mídia programática”, de Stefanie Silveira (UFSC) e João Gabriel Morisso (USP). Aqui os autores explicam como a publicidade digital se baseia, desde o início da internet, na busca por segmentação, monitoramento e metrificação. Um dos exemplos atuais explorados é a modalidade de compra e venda em “tempo real” que se baseia em informações sobre as audiências (*Real-Time-Bidding*). Ao final, o artigo discute os problemas relacionados à crescente automatização dos investimentos publicitários em plataformas como YouTube.

A crescente popularização de noticiários hiperindividualizados e suas consequências é o tema do artigo “Um jornalismo para chamar de meu? Algoritmos e o fenômeno da customização de notícias”, de Raquel Dornelas (UERJ). Serviços como “Apple News” e “Google News” são referenciados a partir do atual panorama no Brasil e no exterior. Ao final, a autora discute as consequências das bolhas informacionais à luz da privacidade dos usuários e da “financeirização” dos negócios de Comunicação.

## Agradecimentos

Este dossiê não seria possível sem o apoio efetivo de um conjunto de colegas que se empenharam durante os vários meses de trabalho. As/os pareceristas, por exemplo, contribuíram decisivamente ao avaliar, em um curto intervalo de tempo, os artigos submetidos ao dossiê. Suas sugestões, temos certeza, foram de grande valia para os autores cujos trabalhos foram aprovados ou reprovados. Em ordem alfabética, nossos agradecimentos a: Ana Brambilla, Carlos Henrique Falci, Elias Bittencourt, Fernando Firmino, Joana Ziller,

João Carlos Magalhães, Leonardo Foletto, Ludimila Mattos, Marcelo Freire, Marcelo Trasel, Sandra Montardo e Vanessa Brandão.

As/os pesquisadores que trabalharam duro para traduzir os artigos originais em inglês foram creditados acima. Só quem já traduziu um paper sabe o trabalho essa tarefa dá, a responsabilidade que ela implica e, pior, o quão desvalorizada é esta atividade intelectual nas atuais avaliações de produtividade acadêmica. Muito obrigado aos nossos tradutores e tradutoras *ad hoc*!

Por fim, agradecemos à equipe da revista **Parágrafo** pela confiança e pelo trabalho conjunto em todo o processo. A edição proposta em 2017 por Rafael Grohmann foi conduzida por Francisco de Assis e por Cláudia Nonato, sempre em diálogo tranquilo e qualificado.

## Referências

BRUNO, F. Objetos técnicos sem pudor: gambiarra e tecnicidade. *Revista ECO-Pós*; v. 20, n. 1 (2017): Gilbert Simondon, 2017.

LATOUR, Bruno. *Reagregando o social: uma introdução à teoria do ator-rede*. Salvador/Bauru Edufba/Edusc, 2012.

MILAN, S.; VELDEN, L. VAN DER. The Alternative Epistemologies of Data Activism. *Digital Culture & Society*, v. 2, n. 2, p. 57-74, 2016.

STERNE, J. “What Do We Want?” “Materiality!” “When Do We Want It?” “Now!” In: GILLESPIE, T.; BOCZKOWSKI, P. J.; FOOT, K. A. (orgs.). *Media Technologies*. MIT Press, 2014. p. 119-128.

**Carlos d' Andréa** – Professor do Departamento de Comunicação Social da Universidade Federal de Minas Gerais (UFMG). Doutor em Estudos Linguísticos (Poslin/UFMG), mestre em Ciência da Informação (UFMG) e graduado em Jornalismo (UFMG). Em 2017/2018, pesquisador visitante no Departamento de Media Studies da Universidade de Amsterdam (UVA), Holanda (bolsa CAPES de pós-doutorado). Pesquisador do Núcleo de Pesquisa em Conexões Intermediáticas. Suas pesquisas contam com o apoio da Fapemig e CNPq. **E-mail:** carlosfbd@gmail.com

**Willian Araújo** – Doutor em Comunicação e Informação na Universidade Federal do Rio Grande do Sul (2017), mestre em Processos e Manifestações Culturais na Universidade Feevale e graduado em jornalismo pela Universidade Federal de Santa Maria. Realizou estágio doutoral no Institute Interdisciplinary Internet (IN3) da Universitat Oberta de Catalunya (UOC, Barcelona). Pesquisa a mediação tecnológica em mídias digitais, principalmente nos debates sobre infraestruturas, softwares e algoritmos. Atualmente é professor da Universidade de Santa Cruz do Sul (Unisc). **E-mail:** willianfaraujo@gmail.com



# Questões e apontamentos para o estudo de algoritmos

## Issues and prompts to the study of algorithms

### **Amanda Chevtchouk Jurno**

Universidade Federal de Minas Gerais  
Programa de Pós-Graduação em Comunicação  
Belo Horizonte, MG, Brasil

### **Sílvia DalBen**

Universidade Federal de Minas Gerais  
Programa de Pós-Graduação em Comunicação  
Belo Horizonte, MG, Brasil

## **Resumo**

Controvérsias recentes envolvendo plataformas e suas políticas de circulação de conteúdo fizeram com que os algoritmos se tornassem pauta nos veículos de mídia. Neste trabalho, falamos um pouco sobre o que são os algoritmos e como eles agem nas plataformas, passando por algumas discussões fundamentais que emergem quando falamos sobre algoritmos. Nosso objetivo é, portanto, desenhar e apresentar um panorama para auxiliar pesquisadores que se interessam em estudar a temática.

## **Abstract**

Recent controversies concerning platforms and their policies of circulation of content have put the spotlight on algorithms in media vehicles. In this paper, we talk about what algorithms are and how they act on the platforms, going through some fundamental discussions that emerge when we talk about algorithms. Therefore, our goal is to design and present a panorama to help researchers who are interested in studying the subject.

## **Palavras-chave**

Algoritmos. Plataformas. STS.

## **Keywords**

Algorithms. Platforms. STS.

## Introdução

Controvérsias recentes envolvendo plataformas e suas políticas de circulação de conteúdo fizeram com que os algoritmos se tornassem pauta nos veículos de mídia. Como exemplo, podemos mencionar o caso da censura à foto “Terror of War” pelo Facebook, quando a famosa foto da guerra do Vietnã foi retirada de circulação porque “continha nudez” – nudez refere-se à garotinha que corre nua e assustada fugindo de uma bomba; ou o caso do vídeo do youtuber Logan Paul, na floresta Aokigahara no Japão, no qual o jovem mostra e debocha de corpos de suicidas pendurados nas árvores, em um local conhecidamente usado para este fim; ou sobre a multa de 2,42 bilhões de euros aplicada ao Google, pela Comissão Europeia, em função do abuso da sua posição dominante na ferramenta de comparação de preços para compras online - o Google Shopping. Apesar de não entrarmos em detalhes sobre os casos mencionados, eles servem como exemplo dos vários casos que trouxeram à tona discussões sobre a delegação da moderação, seleção e avaliação de conteúdo para algoritmos – e em que medida essas escolhas são legítimas e acertadas?

Para discutir essa questão é preciso entender primeiro sobre a natureza dos algoritmos e da sua ação nas redes sociotécnicas das quais participam. Neste trabalho, falamos sobre o que são os algoritmos e como eles agem nas plataformas, influenciando o acesso às informações e ao conteúdo de bilhões de pessoas ao redor do mundo. Nesse processo, passamos por algumas discussões fundamentais que emergem quando falamos sobre algoritmos, como por exemplo sobre sua neutralidade e objetividade, ou sobre sua estabilidade e palpabilidade.

Nosso objetivo é, portanto, desenhar e apresentar um panorama que auxilie pesquisadores interessados no estudo da temática, com questões e apontamentos que consideramos serem fundamentais para o estudo de algoritmos e que foram

organizados a partir do trabalho de importantes pesquisadores da área como Tarleton Gillespie, Taina Bucher, Malte Ziewitz, Rob Kitchin, Nick Diakopoulos, dentre outros.

Este trabalho divide-se em quatro tópicos que concentram questões e apontamentos que consideramos essenciais na discussão e estudo da agência dos algoritmos. São eles: ‘O que é um algoritmo?’; ‘Algoritmos são entidades estáveis?’; ‘Existe objetividade e/ou neutralidade algorítmica?’; ‘Algoritmos e sua lógica particular’. Acreditamos que a sistematização e os apontamentos aqui reunidos podem servir de auxílio para pesquisas e discussões futuras sobre a temática tão importante na atualidade.

## O que é um algoritmo?

Algoritmos são “duros de compreender, difíceis de descrever e virtualmente impossíveis de observar” (ZIEWITZ, 2011, p. 1). Assim como átomos, eles podem ser considerados a unidade básica da computação (DIAKOPOULOS, 2014), um termo que agrega em uma única palavra a lógica de funcionamento da linguagem de programação. Porém, apesar do termo ter se popularizado após o advento da computação, em um sentido mais amplo, algoritmo significa “uma sequência finita de passos que se usa para resolver um problema, e algoritmos são muito mais amplos – e mais antigos – do que o computador” (CHRISTIAN; GRIFFITHS, 2017, p. 13). Segundo os autores, inclusive, essa palavra deriva do nome de um importante matemático persa, que viveu no século IX, al-Khwarizmi, e os primeiros algoritmos datam de antes dele – há registros de um algoritmo sumério gravado em uma tábua de barro com 4 mil anos.

Assim, no seu sentido denotativo, a palavra refere-se a uma fórmula matemática, a procedimentos codificados ou softwares que, como define Gillespie (2014, p.167), “transformam o dado de um input em um output desejado”. Há tantos tipos de algoritmos quanto se pode imaginar, criados para automatizar

e sistematizar os mais diferentes tipos de problemas, e que funcionam com as mais diversas combinações de variáveis e dados. Para nós, interessa o que o autor chama de “algoritmo de relevância pública” que são rotinas de programação utilizadas para selecionar o que é mais relevante para determinada situação, trazendo ao encontro do usuário um conteúdo considerado de seu interesse. Eles nos interessam particularmente, porque são o principal tipo de algoritmo usado nas plataformas de redes sociais e sites de busca, como Facebook, Google e Twitter, por exemplo.

Esses algoritmos funcionam prevendo comportamentos e preferências a partir de dados sobre um determinado usuário. O trabalho desses agentes é escolher o que o usuário X gostaria de ver dentre a miríade de conteúdo disponível em uma determinada plataforma ou na web de forma geral, seja com base nas suas escolhas pessoais de X, seja com base em critérios como relevância e importância. Assim, eles escolhem quais posts esse usuário irá visualizar em seu perfil pessoal; em qual ordem eles estarão organizados e, portanto, serão visualizados; quais páginas merecem maior destaque em um determinado site; qual link oferece o conteúdo mais relevante para determinada busca; qual produto pode ser interessante para um usuário X; qual conteúdo está de acordo com os gostos do usuário X, e etc.

### **Algoritmos são estáveis?**

Não. De acordo com Rob Kitchin (2017), os algoritmos são realizações incertas, provisórias e frágeis, porque são constantemente refeitos e reeditados. Para detalhar um pouco mais nossa resposta, apresentamos uma das expressões comumente utilizadas em referência à natureza dos algoritmos e que traz consigo uma ideia de estabilidade: “algoritmos são como caixas-pretas”. Sugerimos então alterarmos a questão para: “os algoritmos são caixas-pretas?” E a resposta a essa questão é dupla: sim e não.

Antes de mais nada, é importante explicarmos que “caixa-preta” é uma expressão comumente usada em referência a entidades estáveis e das quais não se tem conhecimento além do seu exterior.

A expressão caixa-preta é usada em cibernética sempre que uma máquina ou conjunto de comandos se revela complexo demais. Em seu lugar, é desenhada uma caixinha preta, a respeito da qual não é preciso saber nada, senão o que nela entra e o que dela sai (LATOURE, 2011, p. 4).

Caixa-preta é também um importante conceito para autores dos Science and Technology Studies (STS) e da Teoria Ator-Rede (TAR). É usado em referência a redes sociotécnicas estabilizadas com difícil acesso às suas associações internas (JURNO; D'ANDRÉA, 2017a), e que agem em outras redes sociotécnicas como atores individuais, funcionando como uma unidade ainda que sejam compostos por várias entidades.

Ou seja, o conceito passa a ideia de uma associação de atores da qual não se conhece a composição, mas da qual é possível acompanhar a ação enquanto ator singular em uma determinada rede de atores. Apesar de não sabermos o que há ali dentro, conseguimos observar como aquele grupo age em determinadas situações - exatamente como os algoritmos das plataformas. Não temos acesso a todo seu conteúdo - e talvez esse acesso aos códigos não nos dissesse muito sobre a sua verdadeira natureza - mas conseguimos observar sua ação e os resultados dela.

Contudo, esse conceito também é usado em referência a controvérsias estabilizadas, em que é possível prever o que resultará de uma ação. Uma “caixa-preta” é um “caso encerrado, uma asserção indiscutível” (LATOURE, 2011, p. 43), é uma rede de actantes fortemente “convergente e irreversível” cujo comportamento é conhecido e previsível independentemente do contexto onde está inserida (CALLON, 1990, p. 152). Ou seja, muito diferente

dos algoritmos das plataformas - os algoritmos não são estáveis e tão pouco são previsíveis.

Embora a tecnologia possa ser (...) “encaixapretada” (LATOURET 1987, PINCH; BIJKER, 1984) por designers e usuários, isso não deveria nos levar a crer que ela permaneça estável. Na verdade, os algoritmos podem ser facilmente, instantaneamente, radicalmente e invisivelmente alterados. Embora grandes atualizações possam acontecer apenas ocasionalmente, os algoritmos estão regularmente sendo “ajustados” (GILLESPIE, 2014, p. 178).

Rob Kitchin (2017) explica que os efeitos dos algoritmos e do poder delegado a eles não podem ser previstos por três motivos. Primeiro, porque eles fazem parte de amplas redes de relações que medeiam e influenciam a sua ação. Segundo, a sua performance pode ter efeitos colaterais e consequências involuntárias, não antecipadas pelos seus desenvolvedores, e por isso eles podem agir de formas não previstas quando não são acompanhados de perto pelos programadores. Por fim, e não menos importante, as ações dos algoritmos não podem ser antecipadas porque eles podem conter erros ou bugs que modificam a sua ação (KITCHIN, 2017).

Além disso, os algoritmos não são só o que os programadores almejam, mas o resultado de como os usuários lidam com eles no dia a dia, subvertendo, reinventando e retrabalhando suas intenções iniciais. Como exemplo, podemos citar um caso relativo ao Feed de Notícias do Facebook. Em consequência de uso muito particular dado à opção “esconder post” por um determinado grupo de usuários, com um sentido muito diferente daquele imaginado por seus desenvolvedores, os engenheiros do Facebook precisaram adaptar os algoritmos para funcionarem de forma diferenciada naquela situação (OREMUS, 2016). É que ao invés de usarem a opção para retirar do Feed de Notícias o conteúdo que os desagradava - ação antecipada pelos desenvolvedores - esses usuários usavam a opção para

arquivar posts já visualizados, prejudicando a avaliação dos algoritmos.

Outro exemplo de ações dos algoritmos que não foram antecipadas pelos desenvolvedores, refere-se às sugestões fornecidas ao se fazer uma busca no Google, operadas pelos algoritmos da ferramenta AutoCompletar. Um caso emblemático, refere-se a momentos em que esses algoritmos “adotaram” uma visão sexista na plataforma. Quando os usuários digitavam a sentença “ela inventou” no campo de busca, a ferramenta sugeria uma correção no gênero pronominal presumindo que fora um erro de digitação - “você quis dizer ‘ele inventou?’”. A empresa justificou-se dizendo que a associação feita pelos algoritmos tomara como base a proporção das combinações dos termos em suas bases de dados - em que há muito mais sentenças - “ele inventou” do que “ela inventou”. Isso não impediu a emergência de uma ampla discussão sobre misoginia e preconceito na plataforma (ZIMMER, 2007).

Além disso, também devemos levar em conta as constantes alterações que são feitas nos critérios dos algoritmos a fim de priorizar determinado conteúdo em detrimento a outro, manter os usuários interessados na plataforma e manter as empresas que pagam pela exibição de suas informações satisfeitas. Exemplo disso são as constantes alterações feitas nos algoritmos responsáveis pelo Feed de Notícias do Facebook (JURNO; D'ANDRÉA, 2017a). Algumas delas são publicizadas de forma rasa pelos seus executivos através do site Newsroom, que reúne informações direcionadas especialmente para os interessados em investir dinheiro na plataforma. Dentro das “Notícias”, na categoria “Informações sobre o Feed de Notícias”, podemos perceber que os algoritmos e os critérios usados por eles são alterados pelo menos uma vez por mês pelos seus programadores. Isso levando em consideração as mudanças que são de alguma forma tornadas públicas pela empresa. Ou seja, os algoritmos são “provavelmente tão dinâmicos que uma fotografia instantânea nos daria poucas chances de acessar seus vieses” (PASQUALE, 2009, on-line).

Gillespie (2011, on-line) ressalta que há “uma importante tensão que emerge entre o que nós esperamos que esses algoritmos sejam, e o que eles são de fato”. Eles não são objetos concretos que podem ser colocados na palma da mão, por exemplo, mas também não deixam de ser entidades materiais. A sua composição está relacionada tanto aos bancos de dados, aos quais eles estão conectados, e às partes físicas dos computadores e processadores que os sustentam, quanto ao perfil e às ações do usuário que acessam o site em questão, à temporalidade e condições do acesso, etc. “Na verdade, o que nós nos referimos enquanto um algoritmo muitas vezes não é apenas um, mas são muitos algoritmos” que agem em conjunto, afirma Gillespie (2014, p. 178).

Assim, voltando à questão sobre essas entidades serem ou não “caixas-pretas”, podemos responder que: sim, de uma certa maneira os algoritmos são “caixas-pretas” porque são associações de atores que agem em conformidade, que se comportam em uma rede como se fossem entidades unas, e porque temos pouco (ou nenhum) acesso à sua verdadeira composição. Assim, considerá-los como tal, pode ser uma escolha metodológica importante (JURNO, 2016). Mas não, os algoritmos não podem ser considerados uma “caixa-preta” porque, ao contrário dela, suas ações não podem ser previstas e antecipadas completamente uma vez que eles são extremamente maleáveis e estão em constante transformação - em um constante devir.

## **Existe objetividade e/ou neutralidade algorítmica?**

Mais importante do que uma resposta positiva ou negativa a essa pergunta, é enumerarmos algumas questões cruciais que surgem quando estudamos os algoritmos e sua pretensa/possível neutralidade e/ou objetividade.

Os algoritmos são máquinas inertes até estarem conectados aos bancos de dados a partir

dos quais funcionam. Dessa forma, seria impossível pensar neles ou estudá-los de forma separada das informações que os fazem funcionar e das redes das quais fazem parte. Sem as informações de input, os algoritmos não podem agir e não geram output, não passando de uma série de variáveis e comandos organizados sob a forma de uma expressão matemática.

No caso dos algoritmos que nos interessam neste trabalho, os “algoritmos de relevância pública” (GILLESPIE, 2014, p. 168), nas suas bases e bancos de dados, eles encontram diversas informações sobre as preferências e as particularidades dos usuários que estão realizando aquele acesso ou aquela busca e, assim, podem trabalhar com o intuito de personalizar as suas ações ao perfil daquele usuário, prevendo qual o tipo de conteúdo pode ser considerado interessante por eles, selecionando-o e elencando em uma determinada ordem considerada relevante para eles. Esses bancos de dados são compostos por milhões de informações geradas nos acessos dos usuários e se referem a diversos aspectos; desde questões socioeconômicas e geográficas, até padrões de acesso, que também são geradas e armazenadas por outros algoritmos.

Todas as nossas ações na internet deixam rastros digitais (BRUNO, 2012) que são coletados e armazenados pelos algoritmos para gerar padrões de comportamento e tentar entender melhor a nossa navegação a fim de selecionar conteúdos considerados de interesse. Fernanda Bruno nos chama a atenção para o fato de que “se historicamente entendeu-se o esquecimento como o efeito mais ‘natural’ e o registro como gesto suplementar, vivemos na rede o inverso: para que o esquecimento se produza, é preciso uma ação deliberada” (idem, 2012, p. 7). Ou seja, a tendência na internet é que toda a ação feita seja registrada e armazenada. Assim, todas essas informações geradas nos nossos acessos são computadas junto às de milhares de outros usuários e usadas para criar “perfis de público”, prever comportamentos para serem usados na venda de espaços

de publicidade personalizada, relatórios de marketing para empresas interessadas ou sistemas de vigilância, dentre outros. Com essas informações, os algoritmos podem selecionar e exibir só o conteúdo considerado de interesse do usuário.

O que faz o Google ser tão dominante como meio de acessar a informação no mundo não é tanto ele achar nosso texto em centenas de milhões de webpages (...) mas o fato de ele ordenar essas webpages tão bem, e só nos mostrar as dez mais relevantes (CHRISTIAN; GRIFFITHS, 2017, p.101)

Na verdade, é essa a razão do destaque de todas as grandes plataformas: a seleção e moderação de conteúdo. De acordo com Gillespie (2018), a moderação é a principal commodity que as plataformas vendem, tanto para os usuários, quanto para os investidores. As plataformas prometem mostrar o que há de mais interessante e mais importante no universo de informações disponíveis online e organizá-las de forma compreensível. Essas informações são selecionadas a fim de atrair os usuários para a plataforma e mantê-los ali o maior tempo possível em troca de exposição a propagandas e disponibilização de mais dados pessoais e de navegação. A promessa das plataformas é que o conteúdo disponibilizado seja selecionado especialmente para aquele usuário, com base nos seus interesses pessoais. Mas, na verdade, os algoritmos trabalham principalmente a partir da criação de perfis de padrões de comportamento criados com base em dados de milhares de outros usuários. Os algoritmos enquadram cada usuário em um tipo de perfil e, com base nos padrões de comportamento daquele perfil, selecionam o conteúdo que será disponibilizado para eles.

Nesses bancos de dados sobre os usuários e seus acessos estão disponíveis informações sobre a rotina diária dos mesmos, os locais que eles frequentam, quais produtos consomem, quais são os seus trajetos diários, dentre outros. Por exemplo, ao permitir que sua conta seja usada em outro

site - como quando escolhemos nos logar a partir da conta do Twitter, Gmail ou Facebook, opção que implica na disponibilização dos dados da conta - as ações do usuário nas páginas e sites acessados passam a ser monitoradas e são traduzidas e registradas em grandes bancos de dados. O Facebook monitora inclusive os textos que os usuários digitam e apagam sem sequer os publicar ou enviar para outra pessoa (GROSSMAN, 2013) e o Google mantém um histórico detalhado sobre os últimos passos do usuário, com hora e local de visita, bem como tempo de permanência em cada lugar (VOLTOLINI, 2014). E esses são apenas alguns exemplos. De acordo com Julian Assange, fundador do Wikileaks, o Google e o Facebook juntos coletam mais dados que o governo dos Estados Unidos.

Todos esses dados, então, são categorizados, organizados, limpos e valorados a partir de critérios criados por pessoas que compõem a equipe de programação e cargos executivos da empresa. Esse tratamento é essencial para que os algoritmos possam trabalhar com os dados, processo que Gillespie (2014, p. 170) chama de "preparação para o algoritmo". Como os algoritmos são programados para agir de forma automática, ou seja, sem intervenção humana no processo, os dados precisam estar muito bem organizados e preparados para que os algoritmos consigam executar suas funções sem interrupção. Pois, como nos lembra Franklin Foer (2017, on-line), "o algoritmo foi desenvolvido para automatizar o pensamento, para remover decisões difíceis das mãos dos humanos, para resolver debates controversos". Dentre essas decisões, podemos mencionar a separação entre o que pode ser considerado discurso de ódio ou não, o que se configura como propaganda terrorista ou não, o que pode ser considerado conteúdo sexual ou não e etc.

Assim, como nos lembram Gitelman e Jackson (2013), não é possível falar em 'dados brutos', pois todos os dados são tratados e classificados

de alguma forma antes de serem armazenados e são programados com uma lista de critérios produzida parcialmente por um grupo de pessoas – para eles, dados brutos são um oxímoro. Diakopoulos (2014) chama a atenção para essa dimensão humana dos algoritmos que exercem influência nos critérios de escolha e interpretação dos dados.

A ética e transparência algorítmica devem, portanto, considerar os algoritmos como objetos criados por humanos e levar em conta as intenções, de grupos ou processos institucionais, que influenciaram o design, assim como a agência de atores humanos na interpretação do output em processos superiores de tomada de decisões (DIAKOPOULOS, 2014, p. 5).

Gillespie (2014) argumenta que uma das dificuldades de se entender como os algoritmos operam reside no fato deles imbuírem uma falsa ideia de imparcialidade, sustentada pelo discurso de grandes empresas de tecnologia – como o Facebook e o Google – que não revela as redes sociotécnicas por trás destas plataformas. Estas empresas fomentam a ideia de neutralidade algorítmica como uma estratégia para não assumirem a responsabilidade pela ação dos algoritmos, cada vez mais complexos, que desenvolvem. De acordo com o autor,

a cuidadosa articulação de um algoritmo como imparcial (mesmo quando essa caracterização é mais ofuscante do que explicativa) o certifica como um ator sociotécnico confiável, confere aos seus resultados relevância e credibilidade, e mantém a aparente neutralidade do provedor diante das milhões de avaliações que faz (GILLESPIE, 2014, p. 179).

Nas suas publicações oficiais, as empresas não costumam falar sobre os critérios usados para programar esses algoritmos, como se esses critérios fossem inerentes a esses atores e como se eles

fossem questões dadas. Como exemplo, podemos observar a descrição dos algoritmos do Google que, de acordo com a empresa,

utilizam mais de 200 sinais ou ‘pistas’ diferentes para adivinhar o que você realmente procura. Esses sinais incluem coisas como os termos em websites, a atualização do conteúdo, a região do usuário e o PageRank (GOOGLE, 2018, on-line).

Como podemos ver, as descrições são vagas e amplas e não deixam claro que os critérios levados em consideração pelos algoritmos foram pré-programados por alguém. Quando o Facebook afirma que, no Feed de Notícias, as publicações que o usuário vê “servem para manter você conectado com pessoas, locais e assuntos importantes” e que “as publicações que aparecem primeiro são influenciadas por suas conexões e atividades” (FACEBOOK, 2018), não há qualquer menção aos critérios de seleção que regem esses algoritmos e se parte do pressuposto de que o critério “importante” é um dado, sendo que se trata de um critério relativo. Tarleton Gillespie (2016) salienta que todas essas escolhas são feitas por humanos e, conseqüentemente, estão intrinsecamente carregada de valores.

Às vezes, elas são feitas no design do algoritmo, às vezes em torno dele. O resultado que vemos, (...) não é o output de ‘um algoritmo’ por si só, mas sim de um esforço que combina a atividade humana e a análise computacional, em conjunto, para produzi-lo. Então na verdade os algoritmos estão cheios das pessoas e das decisões que elas tomam (GILLESPIE, 2016, on-line).

Lucas Introna (2016) acrescenta a essa visão que os algoritmos também impõem importantes escolhas de valores

feitas de forma implícita ou explícita, por aqueles que os construíram e os implementam, o que pode ter im-

plicações significativas para aqueles que se baseiam neles em suas práticas diárias. Essas escolhas são feitas por alguém e, como escolhas, são parciais e têm valores que as sustentam (INTRONA, 2016, p.18).

Em matéria publicada no jornal *The Guardian*, John Naughton (2016, on-line) afirma que todo algoritmo faz escolhas com base em critérios especificados por seus designers e que, por sua vez, são expressões de valores humanos. “Os engenheiros podem pensar que eles [os algoritmos] são ‘neutros’, mas a longa experiência tem nos mostrado que eles [os engenheiros] são bebês na floresta da política, da economia e da ideologia”. Assim, apesar de afirmado por seus desenvolvedores e provavelmente defendido por seus engenheiros, longe de serem neutros e apolíticos, os algoritmos constroem e implementam regimes de poder e conhecimento, e o seu uso tem implicações normativas (KITCHIN, 2017).

## Algoritmos e sua lógica particular

Ao se trabalhar com algoritmos é preciso sempre ter em mente que eles são o resultado de complexas redes sociotécnicas, formadas por diversos atores humanos e não humanos, e que influenciam e participam da construção de sentidos dos usuários das plataformas em que eles trabalham.

De acordo com Jurno e d’Andréa,

a partir de cálculos pouco conhecidos, os algoritmos instauram uma lógica própria de administração dos fluxos de informações e reorganizam os conteúdos a fim de contemplar aquilo que detectam como sendo de interesse de cada um, instaurando um regime peculiar de visibilidade e invisibilidade (GILLESPIE, 2014) e uma nova lógica de conhecimento (BUCHER, 2012). (JURNO; D’ANDRÉA, 2017a, p. 480).

David Beer (2009) argumenta que, na era da seleção e filtragem algorítmica, nós não encon-

tramos as informações, mas elas nos encontram a partir dos critérios de seleção e ranqueamento de conteúdo. São os algoritmos do Facebook que decidem, por exemplo, sobre quais amigos e quais histórias nós vamos ter conhecimento (apesar de a empresa ressaltar que essa escolha é completamente baseada nas preferências do usuário e administrável por ele) e quais amigos vão ficar “sumidos” da nossa rotina. São os algoritmos do Google que decidem se vamos visualizar notícias sobre guerras ou fotos de praias paradisíacas ao digitarmos “Israel” na barra da ferramenta de busca. As informações e os posts que nos são exibidos têm a ver com os padrões de comportamento nos quais fomos enquadrados pelos algoritmos e o que vamos receber em resposta serão os conteúdos exibidos para aquele tipo de usuário. Assim, “consequentemente nossas experiências e visões do mundo serão moldadas pela triagem e filtragem dos algoritmos” (idem, 2009, p. 998).

Outra questão importante que deve ser levantada ao estudarmos os algoritmos são as questões políticas, econômicas e técnicas que estão por trás da seleção dos “melhores resultados” para os usuários. Melhores para quem? De acordo com quem? Em relação a quê? São perguntas simples, mas que estamos esquecendo de questionar sobre os algoritmos e seu funcionamento.

Ao discutir a implicação dessas questões políticas, econômicas e técnicas dos algoritmos de PageRanking do Google, Lucas Introna e Helen Nissenbaum (2000, p.173) criticam como a lógica de classificação das páginas por esses algoritmos só considera um link importante “se os outros que já são vistos como importantes indicam que você é importante”. Para eles, o design da ferramenta de busca é uma questão política que dá proeminência para os vencedores - se o seu link direciona e/ou é direcionado por páginas de confiança - então, ele é classificado como importante. Para entendermos melhor a situação que os autores criticam, basta pensar na diferença de alcance orgânico dos posts

de um jornal de bairro e dos posts de um colunista da Folha de S.Paulo no Facebook. A questão é que os algoritmos recompensam com visibilidade aqueles que já têm visibilidade e, assim, é cada vez mais difícil mudar essa relação.

Kate Crawford (2016, p. 77) também acredita que os algoritmos estão produzindo “vencedores em disputas de informação, muitas vezes com pouca visibilidade ou responsabilidade sobre a forma como essas disputas são concebidas”. Quais links merecem aparecer na primeira página do Google? Quais páginas merecem ter seu post inserido no Feed de Notícias dos usuários? Qual post merece ser o primeiro na ordem de exibição? Qual propaganda vai ser exibida mais vezes? Para compreendermos a relevância desses questionamentos, basta pensarmos em uma pesquisa no Google: quantas vezes chegamos a clicar em “próxima página” ou rolamos o scroll da tela para baixo, e quantas vezes acabamos clicamos nos primeiros links exibidos? Quando pensamos nos algoritmos apenas enquanto fórmulas matemáticas, perdemos de vista as questões políticas e econômicas que estão por trás do seu funcionamento e que geram uma nova lógica de conhecimento e (in)visibilidade.

Também não podemos perder de vista as questões sociais e de comportamento que emergem com os algoritmos. De acordo com Taina Bucher (2012), o algoritmo do Facebook tende a recompensar apenas o tipo ‘certo’ de compartilhamento, dando mais visibilidade a certos tipos de post às custas de outros. Assim, os usuários e as organizações são influenciados a produzirem conteúdos de determinadas formas, e a deixarem de produzir de outras, em busca do que gera mais engajamento e conseqüentemente mais visibilidade. Dessa forma, a autora acredita que os algoritmos não são apenas processos computacionais abstratos, mas entidades que têm também o poder de implementar realidades materiais ao moldar a vida social em vários níveis (BUCHER, 2016).

Por fim, vale ressaltar que essa não é uma

estrada de mão única: considerar a ação dos algoritmos não é adotar uma visão tecnocêntrica da realidade. Da mesma forma que os algoritmos são capazes e moldam a realidade das pessoas, as pessoas também são capazes e moldam os algoritmos de acordo com a sua realidade (basta lembrar do caso do botão “esconder o post” que mencionamos anteriormente nesse artigo). Para Bucher (2016, p. 41), “o poder social dos algoritmos - particularmente no contexto de aprendizagem de máquina - deriva da recorrente ‘relação de forças’ entre pessoas e algoritmos”. Os atores da rede se influenciam mutuamente.

## Considerações finais

Neste trabalho tentamos trazer algumas das questões que consideramos centrais no estudo de algoritmos, apresentando argumentos e visões de autores que têm se debruçado sobre o tema.

Contamos sobre como a palavra ‘algoritmo’ refere-se a uma fórmula matemática, a procedimentos codificados ou softwares que fornecem um resultado desejado a partir de um dado inserido (GILLESPIE, 2014). Falamos sobre como os algoritmos são empregados para automatizar e sistematizar problemas e processos, e funcionam com as mais diversas combinações de variáveis e dados. Vale ressaltar que aqui nos interessam os chamados “algoritmos de relevância pública” (GILLESPIE, 2014, p. 168) que funcionam prevendo comportamentos e preferências a partir de uma base de dados e de critérios previamente definidos. Eles nos interessam por que são o principal tipo de algoritmo usado nas plataformas de redes sociais e sites de busca.

Também falamos sobre como é preciso ter em mente que os algoritmos e suas respectivas ações fazem parte de redes sociotécnicas heterogêneas. Ou seja, eles agem e influenciam outros atores nas redes das quais fazem parte e também são influenciados e agenciados pela ação desses outros atores. Tentar considerar a ação de um algoritmo

de forma separada e isolada não só é errado, como impossível – estaríamos olhando para uma grande fórmula matemática inerte.

Mostramos como os algoritmos não são entidades estáveis, como seus desenvolvedores insistem em descrevê-los. De acordo com Rob Kitchin (2017), os efeitos dos algoritmos e do poder delegado a eles não podem ser previstos por que eles fazem parte de amplas redes de relações que medeiam e influenciam a sua ação; por que sua performance pode ter efeitos colaterais e consequências involuntárias; e por que eles podem conter erros ou bugs que modificam a sua ação. Além disso, os algoritmos passam por constantes alterações em seus critérios de funcionamento sem que os usuários percebam essas alterações na interface de acesso (JURNO; D'ANDRÉA, 2017a).

Também explicamos como os dados usados pelos algoritmos são previamente categorizados, organizados, limpos e valorados a partir de critérios criados por humanos e, conseqüentemente, estão intrinsecamente carregados de valores subjetivos. Dessa forma, longe de serem entidades neutras e apolíticas, os algoritmos constroem e implementam regimes de poder e conhecimento, e o seu uso tem implicações normativas (KITCHIN, 2017). De acordo com Jurno e d'Andréa (2017a), os algoritmos instauram uma lógica particular de administração e organização das informações, criando um regime peculiar de (in)visibilidade e uma nova lógica de conhecimento.

Assim, acreditamos que, para estudarmos essas entidades, é preciso levar em consideração as questões políticas, econômicas e sociais por trás das escolhas de resultados “mais relevantes” ou “conteúdos mais importantes”. Os algoritmos estão sempre inseridos em redes sociotécnicas compostas por diversos atores que agenciam e influenciam uns aos outros enquanto interagem. Dessa forma, acreditamos que a melhor maneira de estudá-los seria mapear os atores que fazem parte dessa rede e, seguindo os rastros deixados por eles nesse processo (LATOURE, 2012), tentar

entender as diversas relações de forças (e de poder) que estão ali presentes.

## Referências

- BEER, David. Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, n. 6, p. 985-1002, 2009, v.11.
- BRUNO, Fernanda. Rastros digitais sob a perspectiva da teoria ator-rede. *Revista FAMECOS: mídia, cultura e tecnologia*, Porto Alegre, n. 3, 2012, v.19.
- BUCHER, Taina. The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, v. 20, n. 1, p. 30-44, 2016.
- BUCHER, Taina. Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, n.7, p.1164-1180, 2012.
- CALLON, Michel. Technoeconomic networks and irreversibility. *The Sociological Review*, n. S1, p. 132-161, 1990, v.38.
- CHRISTIAN, Brian; GRIFFITHS, Tom. Algoritmos para viver: A ciência exata das decisões humanas. SÃO Companhia das Letras, 2017.
- CRAWFORD, Kate. Can an algorithm be agonistic? Ten scenes from life in calculated publics. *Science, Technology, & Human Values*, n. 1, p. 77-92, 2016, v.41.
- DIAKOPOULOS, Nicholas. Algorithmic-Accountability: the investigation of Black Boxes. *Tow Center for Digital Journalism*, 2014. Disponível em: < <https://towcenter.org/research/algorithmic-accountability-on-the-investigation-of-black-boxes-2/> >. Acesso em 20 mar. 2018.
- FACEBOOK. 2018. Central de ajuda. Disponível em: < <https://www.facebook.com/help/327131014036297/> >. Acesso em 20 mar. 2018.

FOER, Franklin. Facebook's war on free will. How technology is making our minds redundant. *The Guardian*, Londres, 19 set. 2017. Disponível em: <<https://www.theguardian.com/technology/2017/sep/19/facebooks-war-on-free-will>>. Acesso em: 15 mar. 2018.

GILLESPIE, Tarleton. Algorithms, clickworkers, and the befuddled fury around Facebook Trends. *Culture Digitally*, USA, 18 maio 2016. Disponível em: <<http://culturedigitally.org/2016/05/facebook-trends/>>. Acesso em: 15 mar. 2018.

GILLESPIE, Tarleton. Can an algorithm be wrong? Twitter Trends, the specter of censorship, and our faith in the algorithms around us. *Culture Digitally*. 19 out. 2011 Disponível em: <<http://culturedigitally.org/2011/10/can-an-algorithm-be-wrong/>>. Acesso em: 15 mar. 2018.

GILLESPIE, Tarleton. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven: Yale University Press: Virtual Books, 2018.

GILLESPIE, Tarleton. The Relevance of Algorithms. In: GILLESPIE, T. et al. (Orgs.) *Media technologies: Essays on communication, materiality, and society*. Cambridge, MA: MIT Press: Virtual Books, 2014. p. 167-194.

GITELMAN, Lisa; JACKSON, Virginia. Introdução. In: GITELMAN, Lisa (Ed.). *Raw data is an oxymoron*. Cambridge: MIT Press: Virtual Books, 2013.

GOOGLE. 2018. Algoritmos. In: *Google Por Dentro da Pesquisa*. [S.l.], 2018. Disponível em: <[https://www.google.com/intl/pt-BR\\_ALL/insidesearch/how-searchworks/algorithms.html](https://www.google.com/intl/pt-BR_ALL/insidesearch/how-searchworks/algorithms.html)>. Acesso em 15 mar. 2018.

GROSSMAN, Samantha. Facebook Is Keeping Track Of Every Post You Write and Don't Publish. The social network monitors what you say - and what you don't say. *Time*, USA, 16 dez. 2013 Disponível

em: <<http://newsfeed.time.com/2013/12/16/facebook-is-keeping-track-of-every-post-you-write-and-dont-publish/>>. Acesso em 15 mar. 2018.

INTRONA, Lucas D. Algorithms, governance, and governmentality: On governing academic writing. *Science, Technology, & Human Values*, n. 1, p. 17-49, 2016, v.31.

INTRONA, Lucas D.; NISSENBAUM, Helen. Shaping the Web: Why the politics of search engines matters. *The information society*, n. 3, p. 169-185, 2000, V.16.

JURNO, Amanda Chevtchouk; D'ANDRÉA, Carlos Frederico de Brito. (In) Visibilidade Algorítmica no "Feed de notícias" do Facebook//Algorithmic (In) Visibility in Facebook News Feed. *Contemporanea-Revista de Comunicação e Cultura*, n. 2, p. 463-484, 2017a, v.15.

JURNO, Amanda Chevtchouk; D'ANDRÉA, Carlos. Do algorithms have cosmopolitics? A discussion based on Facebook's nudity policy. In: ANNUAL CONFERENCE OF THE ASSOCIATION OF INTERNET RESEARCHERS, 2017b, Tartu, Estonia. Disponível em [https://aoir.org/wp-content/uploads/2015/10/AoIR2017\\_Conference\\_Schedule.pdf](https://aoir.org/wp-content/uploads/2015/10/AoIR2017_Conference_Schedule.pdf) Acesso em 10 de maio de 2018.

JURNO, Amanda Chevtchouk. *Agenciamentos coletivos e textualidades em rede no Facebook: uma exploração cartográfica*. Belo Horizonte, 2016. 137 f. Dissertação (Mestrado em Comunicação Social). Universidade Federal de Minas Gerais.

JURNO, Amanda. Objetividade e algoritmos: o Facebook e a controvérsia na lista Trending. In: ANAIS DO XXXIX- CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 2016. Disponível em: <<http://portalintercom.org.br/anais/nacional2016/resumos/R11-1456-1.pdf>>. Acesso em: 14 ago. 2017.

KITCHIN, Rob. Thinking critically about and rese-

arching algorithms. *Information, Communication & Society*, v. 20, n. 1, p. 14-29, 2017.

LATOURE, Bruno. *Ciência em ação: como seguir cientistas e engenheiros sociedade afora*. 2.ed. São Paulo: UNESP, 2011.

LATOURE, Bruno. *Reagregando o social: uma introdução à teoria do ator-rede*. Salvador: Edufba, 2012.

NAUGHTON, John. Here is the news - but only if Facebook thinks you need to know. *The Guardian*. 15 maio 2016 Disponível em: <<https://www.theguardian.com/commentisfree/2016/may/15/facebook-instant-articles-news-publishers-feeding-the-beast>>. Acesso em: 15 mar. 2018.

OREMUS, Will. Who Controls Your Facebook Feed? A small team of engineers in Menlo Park. A panel of anonymous power users around the world. And, increasingly, you. *Slate*. 3 jan. 2016 Disponível em: <[http://www.slate.com/articles/technology/cover\\_story/2016/01/how\\_facebook\\_s\\_news\\_feed\\_algorithm\\_works.html](http://www.slate.com/articles/technology/cover_story/2016/01/how_facebook_s_news_feed_algorithm_works.html)>. Acesso em 18 jan. 2016.

PASQUALE, Frank. 18 nov. 2009. *Assessing Algorithmic Authority*. *Madisonian: Essays on law, leadership, culture, and technology*. 18 nov. 2009. Disponível em: <<http://madisonian.net/2009/11/18/assessing-algorithmic-authority/>>. Acesso em: 15 mar. 2018.

VOLTOLINI, Ramon. Por onde você andou? Consulte seu histórico de localização junto à Google. *TecMundo*. 13 ago. 2014. Disponível em: <<https://www.tecmundo.com.br/google/60454-ultimos-passos-consulte-historico-localizacao-junto-google.htm>>. Acesso em 15 mar. 2018.

ZIEWTIZ, Malte. How to think about and algorithm: Notes from a not quite random walk. Discussion paper of the SYMPOSIUM KNOWLEDGE MACHINES BETWEEN FREEDOM AND CONTROL. Kulturfabrik Hainburg, Austria, October 6-7, 2011. Draft version: september 29, 2011. Disponível em <http://zwtz.org> Acesso em 10 de maio de 2018.

ZIMMER, Michael. Google: "Did you mean: 'He invented'?" Michael Zimmer. 9 maio 2007. Disponível em: <<http://www.michaelzimmer.org/2007/05/09/google-did-you-mean-he-invented/>>. Acesso em: 9 mar. 2018.

**Amanda Chevtchouk Jurno** – Doutoranda pelo Programa de Pós-Graduação em Comunicação da Universidade Federal de Minas Gerais (PPGCOM-UFMG). Faz parte do Centro de Convergência de Novas Mídias e do Núcleo de Pesquisa em Conexões Intermidiáticas. Mestre pelo PPGCOM e Bacharel em Comunicação Social - Jornalismo pela UFMG. **E-mail:** amandajurno@gmail.com

**Sílvia DalBen** – Mestranda pelo Programa de Pós-Graduação em Comunicação da Universidade Federal de Minas Gerais (PPGCOM-UFMG), onde estuda com bolsa de estudos da FAPEMIG. Faz parte do Centro de Convergência de Novas Mídias e do Núcleo de Pesquisa em Conexões Intermidiáticas. É Bacharel em Comunicação Social - Jornalismo pela UFMG. **E-mail:** silviadalben@ufmg.br

Recebido: 21 mar. 2018

Aprovado: 11 abr. 2018



# **Repensando *big data*, algoritmos e comunicação: para uma crítica da neutralidade instrumental**

## **Rethinking the big data, algorithms and communication: a critique of instrumental neutrality**

### **Tales Tomaz**

Centro Universitário Adventista de São Paulo.  
São Paulo, SP, Brasil

### **Guilherme Cavalcante Silva**

Universidade Estadual de Campinas  
Laboratório de Estudos Avançados em Jornalismo  
Campinas, SP, Brasil

## Resumo

Este artigo discute a presença, tanto em discursos sobre *big data* como nos estudos em comunicação, da ideia de neutralidade instrumental dos dados e algoritmos. O texto procura elucidar a noção etimológica e contextual de *data* e o modo como algumas das abordagens sobre *big data* tratam dados, algoritmos e informação. A partir daí, apresenta as dificuldades de compreender dados e algoritmos como meros instrumentos neutros para fins. Em seguida, o texto mostra como ideia semelhante já estava presente na compreensão cibernética da comunicação. Por fim, propõe-se uma crítica à neutralidade instrumental pressuposta na ontologia do *big data* e da comunicação, com destaque para a questão do apagamento da alteridade, seguindo a linha da fenomenologia de Martin Heidegger.

## Abstract

This paper discusses the presence both in big data discourses and in communication studies of the idea of instrumental neutrality in data and algorithms. The text seeks to elucidate the etymological and contextual notion of data and how some of the big data approaches deal with data, algorithms, and information. From there, it presents the difficulties of understanding data and algorithms as mere purpose-neutral instruments. Then the text shows how similar idea was already present in the cybernetic understanding of communication. Finally, we propose a critique of instrumental neutrality presupposed in the ontology of the big data and communication, with emphasis on the question of the erasure of alterity, following the argument of Martin Heidegger's phenomenology.

## Palavras-chave

Big data. Algoritmos. Comunicação.

## Keywords

Big data. Algorithms. Communication.

## Introdução

A administração da torrente de dados chamada de *big data* é um assunto que já transborda da área computacional para se tornar questão política, econômica e social de primeiro plano. Nesse contexto, este artigo propõe repensar *big data* e a mediação algorítmica para além de uma noção meramente técnica. A proposta surge de um incômodo com o regresso de elementos, em muitos desses discursos, que apontam para os dados como instrumentos neutros, mediadores objetivos da realidade. Fato teórico similar foi visto anteriormente, argumenta-se, em tratativas da comunicação que a tomam por troca de dados entre dois polos, processo neutro mediado por instrumentos neutros. O objetivo do texto é mostrar como o discurso acerca do *big data* e algoritmos que se baseia na sua neutralidade instrumental perde de vista a redução da realidade neles operada, semelhantemente ao que ocorre quando a comunicação é tratada como troca de informação.

Para alcançar o objetivo proposto, o artigo se inicia com uma investigação acerca da origem da ideia de *data* e sua relação com a matemática no sentido original dos gregos, seguindo pistas deixadas especialmente por Puschmann e Burgess. Depois, apresenta-se a relação do *big data* com os algoritmos e a computação, com destaque para um certo discurso de neutralidade instrumental que escamoteia a existência de toda uma lógica de conhecimento, como propõe Gillespie, ou ontologia, como sugere Capurro. Em seguida, o texto mostra como uma ideia semelhante já estava em operação na noção cibernética de comunicação, introduzida por Wiener e consolidada por Weaver e Shannon. Finalmente, o texto elabora críticas à neutralidade instrumental pressuposta na ontologia do *big data* e dos algoritmos, com destaque para a crítica ao apagamento da alteridade, que é feita a partir da fenomenologia de Heidegger e Merleau-Ponty.

## Origem dos data

Apesar do ar de novidade devido ao seu uso generalizado no contexto digital, o termo *data* não é recente. A própria expressão *big data* já era utilizada em meados do século 20, no contexto da emergência dos supercomputadores que centralizavam grande quantidade de procedimentos de cálculo e armazenamento. Nesse caso, *big data* era uma referência ao constante aumento – e o problema derivado disso – da quantidade de dados a serem armazenados e lidos em um computador e à necessidade de aprimoramento dos sistemas para comportar tanta informação.

Todavia, o discurso atual sobre *big data* se encaminha para direções que vão além de uma simples questão técnica. Boyd e Crawford (2012), por exemplo, propõem que o aspecto-chave do *big data* não é a consideração para com o tamanho dos dados, mas a avaliação *daquilo que pode ser transformado em dados*, mensurado, sondado, agregado e homogeneizado como um aglomerado. Dados, para essas pensadoras, são fenômenos culturais, políticos e econômicos tanto quanto técnicos. Para além de se resumir a questões técnicas, a abordagem atual do *big data* envolve a

crença generalizada de que maiores quantidades de dados oferecem uma forma superior de inteligência e conhecimento que pode gerar insights que eram previamente impossíveis, com uma certa aura de objetividade e exatidão. (BOYD; CRAWFORD, 2012, p.663-664).

De fato, uma característica marcante tanto em discursos negativos – que ressaltam as ameaças à privacidade, liberdade individual e livre concorrência – quanto em narrativas que afirmam o *big data* como um recurso ilimitado para ser manuseado potencialmente em todas as áreas possíveis é um entendimento implícito de dados como realidades transparentes, óbvias, uma for-

ma inquestionável de ter acesso às coisas (GITELMAN; JACKSON, 2013).

É interessante notar que etimologicamente *data* corresponde a essa compreensão, especialmente em seu uso no campo da matemática e das ciências exatas. Nos gregos, a expressão *data* se refere simplesmente a algo que é dado, um presente (Puschmann; Burgess, 2014).<sup>1</sup> Dados são, antes de tudo, aquilo que é dado de antemão em um contexto (GITELMAN; JACKSON, 2013). No contexto científico, o termo passou a designar aquilo que é dado previamente para análise. A ligação dos dados com a matemática fica mais clara quando se entende o sentido original desta última. Heidegger (2002, p.80) explica que *mathemata* (τὰ μαθηματα) era um modo de conhecer a realidade no qual as coisas são aprendidas de antemão, “aquilo acerca das coisas que já conhecemos verdadeiramente, de modo antecipado. Aquilo que, em consequência, não começamos por ir buscar as coisas, mas que, de certo modo, levamos conosco até elas”. Portanto, não é de se estranhar a presença dos *data* no contexto da matemática ainda antes do eclodir da era computacional, uma vez que o traço marcante da matemática é seu caráter de apreensão apriorística, isto é, nela as coisas são vistas a partir de um aspecto que, de alguma forma, já é *dado* de antemão. Nas palavras de Capurro (2006, p. 4), “o caráter paradigmático da ciência natural matemática não é a aplicação de matemática a processos naturais, mas a projeção *a priori* das entidades que ela descobre”. Essa projeção apriorística é exatamente o que está em jogo nos dados.

Entendidos como aquilo que já é apreendido de antemão, dados são objetos passíveis de mensuração, estudo, asseguramento e manipulação. E é precisamente no contexto do crescimento das ciências ditas duras e de um certo espírito positivista que os dados ganham realce como vetores do

conhecimento. Em resumo, o que podemos chamar de datificação, isto é, a conversão das coisas em dados, é um processo que encerra de antemão seu objeto como *algo mensurável*.

## Contribuição dos algoritmos para datificação

Os dados adquiriram a popularidade que têm hoje juntamente com o salto experimentado por outra categoria matemática, a dos *algoritmos*. Estes também se tornaram objeto destacado na atenção pública ao longo dos últimos anos, especialmente com a ascensão do Facebook e a preocupação de que seus códigos algorítmicos estejam alterando a percepção dos usuários sobre a realidade. Como ficará evidente a seguir, é crucial entender melhor também o que são algoritmos quando se está pensando *big data*.

A princípio, algoritmos podem parecer objetos exclusivamente de estudo técnico da computação, afinal, costumam ser definidos como “uma descrição formalizada e abstrata de um procedimento computacional” (DOURISH, 2016, p. 3). Mas o seu sentido é mais amplo, não se reduzindo a computadores ou outros artefatos tecnológicos. Algoritmos são uma tentativa de descrever o passo a passo necessário para a realização de um procedimento de forma tão precisa que possa ser executado sem qualquer interpretação de quem (ou o que) o executa. Dessa forma, obtém-se uma economia de tempo e esforço. O ideal do algoritmo é descrever o passo a passo tão eficazmente a ponto de o procedimento poder ser *automatizado*. Quando isso ocorre, há apenas que se definir o “antes [a escolha do procedimento] e o depois [o que o resultado significa], mas o procedimento em si já está resolvido” (TOMAZ, 2017, p. 53).

Se, por um lado, o sentido dos algoritmos vai além da computação, o fato é que, por outro lado, eles só foram alçados a esse papel de importância na vida humana com o surgimento do computador e, de forma ainda mais espetacular, com o desen-

<sup>1</sup> Puschmann e Burgess (2014) apontam para *δίδωμι* como a raiz grega de onde o termo se desenvolveu. Em latim, a palavra foi traduzida por *datum*, e *data* é sua versão plural.

volvimento do *big data*. Qual é, então, a relação da algoritmização com a computação e com *big data*? A questão é que a tentativa de estabelecer um passo a passo para a realização de um procedimento que não necessite de interpretação particular corresponde à tradução desse procedimento em linguagem formal. E computadores são justamente máquinas que funcionam sob essa linguagem. Não apenas isso. Como demonstrado por Turing, são máquinas capazes de executar qualquer algoritmo que siga rigorosamente a linguagem formal.<sup>2</sup> Não à toa, o computador é considerado fundamentalmente uma *máquina algorítmica* (GILLESPIE, 2014).

Ter à disposição uma máquina capaz de rodar qualquer algoritmo parece ter provido o incentivo necessário para a produção e a sofisticação de algoritmos. Com maiores possibilidades de execução, têm sido desenvolvidos algoritmos cada vez mais complexos, capazes de lidar de forma relativamente eficiente com problemas cuja resolução, até então, se atribuía exclusivamente à aplicação de excepcionalidades como o “espírito humano”, a “liberdade”, a “criatividade” etc. (SHROFF, 2014). Por sua vez, numa lógica de autorreforço, a sofisticação dos algoritmos significa que cada vez mais procedimentos do cotidiano podem ser automatizados, promovendo a aprovação do modo algorítmico de lidar com o real e incentivando ainda mais o recurso à mediação algorítmica. Isso parece ter provido também o incentivo necessário para a tradução de mais aspectos do real em dados, ou seja, para a mencionada datificação. Dessa forma, a linguagem dos algoritmos, a formalização, alcança seu ápice com a chegada da computação e a subsequente hiperprodução de dados (TOMAZ, 2017).

É apenas com este pano de fundo que os dados foram alçados à importância que têm hoje, sob a alcunha de *big data*. A questão é que, assim como ocorre com a datificação, a mediação algorítmica

2 Isso ocorre, em grande parte, graças aos desenvolvimentos da chamada máquina de Turing, que o cientista britânico concebeu como uma máquina universal capaz de executar todos os procedimentos computacionais. Para um resumo sobre essa ideia, ver Sipser (1997, p. 125-135) e Tomaz (2017, p. 51).

também aparece muitas vezes como uma forma objetiva, neutra de lidar com real (GILLESPIE, 2014). Por partir do caráter matemático da certeza, da exatidão do procedimento, os algoritmos passam uma impressão de credibilidade e neutralidade sem precedentes em comparação com a mediação, o juízo e a intervenção humanos. Isso pode ser visto, por exemplo, na maneira como os usuários de redes sociais, como o Facebook, tendem a confiar cada vez mais no seu feed de notícias para se manterem atualizados acerca do que ocorre em seu entorno em vez de se engajarem em uma busca ativa e consciente por diferentes meios de comunicação (ZUÑIGA et.al., 2017). Algoritmos e *big data* tendem a assumir um caráter instrumental de neutralidade.

De fato, Puschmann e Burgess (2014) observam que, na maior parte dos discursos, o *big data* aparece como uma certa força natural a ser controlada ou como uma ferramenta a ser utilizada, um meio para fins. A preocupação parece ser em como administrar e colocar tamanho potencial a serviço de finalidades úteis. Puschmann e Burgess (2014) arrolam alguns exemplos emblemáticos desse discurso. A revista Wired<sup>3</sup> de fevereiro de 2013 utiliza os termos “massa de dados [cujos valores devem ser extraídos]”, “pilha de *big data*” e “dados para além da compreensão humana” para descrever a potencialidade de dados disponíveis para a indústria de marketing potencializar sua comunicação com clientes e outras empresas. Um artigo do USA Today ressalta o *big data* como um oceano infinito para se nadar, onipresente, ali desde o momento em que buscamos saber a previsão do tempo até quando tentamos entrar em contato com alguém.<sup>4</sup> Já o Herald Cronicle aponta para um aprimoramento constante do processamento da torrente de infor-

3 “Smart Data: For Communications, Big Data Becomes Transformational”. Disponível em: <<https://www.wired.com/insights/2013/02/smart-data-for-communications-big-data-becomes-transformational/>> Acesso em: 21 fev. 2018.

4 “How has Big Data changed your life?”. Disponível em: <<https://www.usatoday.com/story/tech/2012/11/29/your-big-data-life-stories/1735223/>> Acesso em: 08 mar. 2018.

mação e circulação de dados na rede digital como um fator-chave para as companhias de tecnologia do Canadá conseguirem angariar fundos.<sup>5</sup>

Muitas vezes, *big data* e algoritmos são tomados como um “meio independente de verificar resultados” e como uma “matéria prima” (SCHROEDER, 2016, p. 12-13). O aumento na quantidade e disponibilidade de dados, bem como sua suposta uniformidade – que é à base da ideia de *big data* –, são vistos como irresistíveis para a teoria e prática comunicacional (JUNGHEER, 2015). Diversos estudos se voltam para as possibilidades do uso dos dados e seu impacto no ambiente digital, especialmente em redes sociais como Facebook, Instagram e Twitter (SCHROEDER, 2016), além das questões políticas e sociais envolvidas em seu uso e consideração por parte da mídia (NEUMAN et. al., 2014).

Portanto, o entendimento da neutralidade dos dados e dos algoritmos, que se alinha com e até mesmo atualiza o positivismo, vai além do universo das redes sociais ou das finanças e políticas globais e tende a se difundir por diversas facetas da vida humana (ILIADIS; RUSSO, 2016). Com *big data* e algoritmos eficientes, a existência parece ser mais controlável do que nunca, desde que saibamos fazer bom uso do poderoso instrumento em mãos.

Já é possível antecipar alguns problemas de uma noção exclusivamente instrumental do *big data* e dos algoritmos. Como coloca Krtilova (2015), os algoritmos não apenas servem como meios para lidar com uma realidade objetiva, estável, mas criam uma experiência em que tudo aparece *tão calculável, apreensível quanto possível*. Tudo tem de ser feito da maneira *mais exata possível*. A implicação é que está em jogo uma *forma de compreender o real*. Por isso, para Gillespie, a forma algorítmica de lidar com as coisas não é sim-

plesmente uma possibilidade à nossa disposição, mas é toda uma nova lógica de conhecimento:

O acesso algorítmico à informação, portanto, representa uma lógica particular de conhecimento, construída sobre premissas específicas sobre o que é o conhecimento e como se deveria identificar seus componentes mais relevantes. Que nós estejamos nos voltando a algoritmos para identificar o que precisamos saber é tão significativo quanto ter confiado em especialistas credenciados, no método científico, no senso comum ou na palavra de Deus. (GILLESPIE, 2014, p 168):

Nesse sentido, o algoritmo não é meramente uma ferramenta neutra, disponível à mão, indiferente às formas de uso de indivíduos e instituições. Seu uso encerra a necessidade de lidar formalmente com as coisas, de tratar procedimentalmente com o mundo. Por isso, há quem aponte algoritmos e *data* como elementos primordiais de uma espécie de ontologia digital (CAPURRO, 2006).

## A noção de comunicação

Por que essa discussão interessa especificamente à comunicação? Num sentido mais prático, como já visto *en passant*, um número crescente de processos comunicativos passa pela mediação algorítmica e se converte em *big data*. Só por esse motivo já se faz necessário considerar as implicações da perspectiva instrumental do *big data* para a comunicação. Entretanto, este texto quer explorar outro aspecto dessa relação, a saber, a ideia de que o discurso de neutralidade instrumental do *big data* e dos algoritmos é o mesmo – na realidade, uma atualização e radicalização – da *forma cibernética de pensar a comunicação*, uma forma que toma comunicação como uma relação neutra entre emissor e receptor.

<sup>5</sup> “Data Science pool touted”. Disponível em: <<http://thechronicleherald.ca/business/204136-data-science-pool-touted>> Acesso em: 08 mar. 2018.

O conceito paradigmático<sup>6</sup> de comunicação que, como se verá, exerceu atuação decisiva na formação da ontologia da neutralidade dos dados opera sob influência da cibernética de Norbert Wiener, colocada em movimento nos anos 40. Wiener estabeleceu que o objetivo principal da cibernética era o de “desenvolver uma linguagem e técnicas que nos capacitem, de fato, a haver-nos com o problema do controle e da comunicação em geral [...], sempre em luta contra a tendência [entrópica] da natureza de degradar o orgânico e destruir o significativo” (WIENERE, 1965, p. 17). O ponto em comum entre Wiener e outros ciberneticistas, como os engenheiros Claude Shannon e Warren Weaver – postuladores do primeiro modelo teórico a relacionar comunicação com troca de informação –, é que tal linguagem se encontra na informatização, sendo esta diretamente relacionada ao controle da entropia.<sup>7</sup>

Sob a influência da noção wieneriana de informação, controlável, como a chave para a redução da entropia, Shannon e Weaver imaginaram que o aspecto primordial da comunicação deveria ser a troca eficaz de informação, sem qualquer tipo de ruído (SHANNON, 1948). E, para contribuir para esse objetivo, eles elaboraram, em fins da década de 1940, o que ficou conhecido como a Teoria Matemática da Comunicação.

Nessa teoria, comunicação é entendida como um processo de troca de informação. Dessa forma, a comunicação pode ser *mensurada* a partir da efetividade na transmissão correta de informações de um pólo para outro, proporcionando o maior controle possível da situação por parte do emissor (SILVA, 2016). Entre as características principais deste teorema se encontram a linearidade e previsibilidade da comunicação e o seu caráter necessário de exatidão, onde o ruído e o imprevisível devem ser controlados e idealmente eliminados. A comunicação na cibernética é totalmente calculável, tendo melhor eficácia inclusive quando feita por máquinas. É então um processo de transmissão e interpretação de sinais, enfim, um sistema técnico. O ato comunicacional é visto como algo necessariamente replicável e mensurável – matemático, enfim. Algo que não é estranho conforme “a sociedade passou a ser vista como um conjunto de circuitos e canais, mantido pelas redes técnicas de informação” (RÜDIGER, 2004, p. 109).

Atualmente a influência do sistema de Shannon e Weaver pode parecer estar em baixa em um contexto acadêmico que rejeita cada vez mais a dicotomia emissor-receptor, vista como antiquada diante do novo cenário virtual (RÜDIGER, 2004). Todavia, como chamam a atenção Gunkel e Taylor (2014, p. 34-35),

6 Tratamos paradigma aqui em termos kuhnianos, isto é, como uma estrutura fonte de “métodos, áreas problemáticas e padrões de solução aceitos por qualquer comunidade científica amadurecida, em qualquer época que considerarmos” (KUHN, 1998, p.137-138).

7 Entropia, para Wiener, é uma medida de desorganização que afeta toda a estrutura natural. “Conforme aumenta a entropia, o universo, e todos os sistemas fechados do universo, tendem naturalmente a se deteriorar e a perder a nitidez, a passar de um estado de mínima a outro de máxima probabilidade; de um estado de organização e diferenciação [...] a um estado de caos e mesmice” (WIENER, 1965, p. 14).

enquanto grande parte dos estudos em Comunicação abandonou um modelo simples de transmissão e coloca a si mesma a tarefa intelectual de ressaltar as diversas complicações e distorções que ocorrem em um ‘simples’ ato de comunicação (por exemplo, o ‘barulho’ que pode interromper a comunicação e, mesmo se recebida, os diferentes modos que os receptores podem escolher para decodificar a mensagem etc.), grande

parte da disciplina ainda trabalha (ainda que implicitamente) com o modelo de Shannon e Weaver.

De fato, quando se vê o discurso de *big data* e algoritmos como instrumentos capazes de controlar melhor a vida humana e o planeta como um todo, percebe-se a continuidade e até mesmo radicalização da influência do *conceito cibernético de comunicação como controle*.

## **Instrumentalidade dos meios técnicos e apagamento da alteridade**

Atualmente, mesmo pensadores que se aproximam mais da cibernética já não aceitam mais esse discurso de neutralidade instrumental dos meios técnicos. Erick Felinto (2011, p. 241), por exemplo, embora aprecie o aspecto performativo da cibernética, afirma que “os meios não apenas nada têm de transparentes, senão também constituem nossa relação com o mundo e definem as condições transcendentais do pensamento”. Os Estudos Sociais da Ciência e Tecnologia – em alguns aspectos tributários da cibernética – também rejeitam a mera instrumentalidade dos objetos técnicos. Para Bruno Latour (2013), um dos mitos legados pela modernidade foi a ideia de que a ciência e a metodologia formal tinham condições de definir ontologicamente as coisas por estar habilitada a olhar objetivamente, *de fora*, para os fenômenos. Para Latour (2014, p. 7), “uma das coisas mais intrigantes na história do Ocidente não é que ainda existam pessoas que acreditam no animismo, mas a crença tola que muitos ainda possuem em um mundo ‘frio’ de meros objetos”. Está bastante claro, para ele, que o mundo do objeto é o mesmo mundo do sujeito.

Entretanto, a crítica à noção cibernética de comunicação ora reproduzida na ontologia digital do *big data* que queremos explorar neste texto é a da fenomenologia. Isso porque, diferentemente de outras críticas, ela ressalta a forma como esse tipo de ontologia promove um apagamento da al-

teridade. Isso fica bastante claro ao se examinar mais uma vez o sentido que Weaver (1977) dava para a comunicação. Para ele, a comunicação era o ato de influenciar, causar um efeito esperado, sobre o outro pólo, visto não como autônomo, mas como uma extensão do emissor. “A comunicação ou influencia a conduta ou não tem qualquer efeito perceptível e comprovável” (WEAVER, 1977, p. 26). O objetivo da comunicação seria equalizar os dois pólos o mais proximamente possível, tornar o fluxo de informações total e totalizante. De fato, uma era sem fronteiras. Mas o que é isso, no final das contas, senão o expurgo da alteridade, a recusa de um elemento imperscrutável de resistência no Outro?

Discutindo a atuação da deusa Atena na mitologia grega como metáfora para pensar a condição tecnológica que se anunciava então para a humanidade, no contexto de sua conferência “A proveniência da arte e a determinação do pensar”, Heidegger fornece pistas para se pensar os problemas ontológicos de uma proposta de controlar o Outro, superando as fronteiras.

Atena aparece como *skeptoméne*, a que medita. Aonde se dirige o olhar meditativo da deusa? À pedra-marco, à fronteira. Fronteira não é só onde algo termina. Fronteira designa aquilo mediante o qual algo é ajuntado no que lhe é próprio, para aparecer, de aí para fora, na sua plenitude e emergir em presença (HEIDEGGER, 2013, p. 123).

O olhar da deusa se firma na fronteira. E o que aparece então? Para Heidegger, aquilo que é. Em outras palavras, fronteira não é apenas algo que limita o avanço técnico-científico ou a eficácia da comunicação. Fronteira reúne algo, delimita-o e, dessa forma, o faz emergir como algo que é. Sem fronteiras, Eu e o Outro somos o mesmo. De fato, não há Outro. Tudo se torna uma extensão do mesmo. E é exatamente nessa pretensão que se firma a ciência cibernética e o

modelo técnico de comunicação. No paradigma cibernético da comunicação, “eu e Outro podem estar em qualquer e todo ente da cadeia de causalidades reproduzíveis [...], uma verdadeira explosão de Identidade e Alteridade em todas as direções” (TOMAZ, 2016, p. 15).

O pensamento cartesiano consolidou a compreensão de que todos os entes são *res extensa*, extensões de um sujeito que, aliás, surge com o próprio Descartes (OLIVEIRA, 1998). Descartes afirma que não encontrou nada senão o fato de ele [*res cogitans*] ser o que duvida [pensa] da certeza das coisas [objetos]. “Logo depois atentei que, enquanto queria pensar assim que tudo era falso, era necessariamente preciso que eu, que o pensava, fosse alguma coisa” (DESCARTES, 2002, p. 38). O conhecimento, para Descartes, repousa na certeza, e esta é colocada exclusivamente no âmbito do que pensa. Assim, a transcendência em Descartes é o próprio sujeito. Ao propor a indiferença absoluta a partir da transparência total da comunicação, a cibernética dá sequência a essa tradição filosófica. É o humano cercado de todos os lados por si mesmo. Um mundo sem Outro.

Ao reduzir comunicação à troca mensurável de informações, o paradigma técnico da comunicação crê na possibilidade de abstrair uma relação de alteridade infinitamente, isto é, reproduzi-la, esquecendo que “operar realidades com a implementação de operações técnicas (controlar processos, ‘fazendo as coisas funcionarem’) somente é possível se você abstrair do emaranhamento de práticas e relações com os outros, do enraizamento no mundo” (KRTILOVA, 2015, p. 33-34). Cada experiência, cada relação, é única, não replicável, não mensurável.

Quando se toma as coisas como fundamentalmente mensuráveis – em um certo discurso de *big data* e algoritmos e na primeira teoria da comunicação –, adota-se a fé implícita de que basta ampliar a quantidade de meios ou ferramentas disponíveis para tornar a comunicação plena e total, reproduzível em todas as partes. No entanto, a comunica-

ção não pode se reduzir a uma ferramenta ou um processo formal, pelo fato de, assim como o próprio humano, sempre se dar no mundo, junto a este, sujeita aos riscos e incertezas deste. Afinal, “o mundo está ali antes de qualquer análise que eu possa fazer dele, e seria artificial fazê-lo derivar de uma série de sínteses [...] quando ambos são justamente produtos da análise e não devem ser realizados antes dela” (MERLEAU-PONTY, 2011, p. 5).

## Considerações finais

Ao escavar os pressupostos de um certo entendimento que une *big data*, algoritmos e comunicação – que considera dados como realidade, algoritmos como instrumentos neutros e comunicação como troca de informação –, este texto buscou apontar a insuficiência epistemológica desse paradigma. Não há algo como dados puros, dos quais se deve deduzir objetivamente numa relação instrumentalmente mediada, operada por sujeitos autônomos. É a projeção de autonomia e neutralidade dos dados e algoritmos, como atores sociotécnicos confiáveis, que alimenta o seu impacto social (GILLESPIE, 2014).

Nesse ponto, conforme o texto demonstrou, os discursos instrumentais sobre *big data* e a ideia matemática de comunicação pecam, pois tomam dados e comunicação como procedimentos puramente neutros, à mão para uso e mensuração, ideia que já encontrava suas premissas na cibernética de Wiener. Eles incorporam certa “fantasia [...], aquela em que vida e técnica devem se fundir para fazer surgir um único organismo maquínico universal” (RÜDIGER, 2014, p. 73). Heidegger já via na cibernética o “ápice da tradição filosófica” ocidental (LEIDLMAIR, 1991, p. 195), o acabamento ou expressão máxima de uma determinada história em que as coisas se apresentam como *coisas à disposição*, à mão, passíveis de mensuração por meio de desenvolvimento tecnológico “eficaz”.

Quando tomam datificação e algoritmização como processos que trazem à luz a realidade dos fatos em si, os discursos sobre *big data* esquecem que esses são fenômenos históricos e, desde sempre, abstrações de um mundo cuja totalidade é irreduzível. Podem ter utilidade circunstancial, mas não caráter de *real*. Isso significa que é necessário pensá-los, bem como a comunicação que pode ocorrer por meio deles, para além do seu êxito ou de sua falha como um sistema. Implica que a comunicação é um fenômeno muito menos mensurável e calculável do que previamente pensado no paradigma emissor-receptor. Depende muito menos de quem manipula ou do manipulado. Uma abordagem mais contextual acerca do *big data*, reconhecendo suas circunscrições sociais, técnicas e históricas como fenômeno e distanciando-se da visão de comunicação como “troca de informações”, este texto argumenta, ajudaria uma melhor compreensão do *big data*.

## Referências

BOYD, Danah; CRAWFORD, Kate. Critical questions for Big Data: provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, n. 5, p. 662-679, jun. 2012, 15 v. DOI: 10.1080/1369118X.2012.678878.

CAPURRO, Rafael. Toward and Ontological Foundation of Information Ethics. *Ethics and Information Technology*, n. 4, p. 175-186, 2006, 8 v. DOI: 10.1007/s10676-006-9108-0.

DESCARTES, René. *Discurso do método*: para bem conduzir a própria razão e procurar a verdade nas ciências. São Paulo: Paulus, 2002.

DOURISH, Paul. Algorithms and their others: algorithmic culture in context. *Big Data & Society*, n. 1, p. 1-11, jul.-dez. 2016, 1 v. DOI: 10.1177/2053951716665128.

FELINTO, Erick. Da teoria da comunicação às teorias da mídia: ou, temperando a epistemologia com

uma dose de cibercultura. *Eco-Pós*, n. 1, p. 233-249, 2011, 14 v. Disponível em: <[https://revistas.ufrj.br/index.php/eco\\_pos/article/view/920](https://revistas.ufrj.br/index.php/eco_pos/article/view/920)>. Acesso em: 24 abr. 2018.

GILLESPIE, Tarleton. The Relevance of Algorithms. In: GILLESPIE, Tarleton; BOCZKOWSKI, Pablo; FOOT, Kirsten. (Orgs.). *Media Technologies: Essays on Communication, Materiality and Society*. Cambridge: MIT Press, 2014.

GITELMAN, Lisa; JACKSON, Virginia. Introduction. In: GITELMAN, Lisa (Ed.). *“Raw Data” is an oximoron*. Cambridge: The MIT Press, 2013.

GUNKEL, David. Communication and Artificial Intelligence: Opportunities and Challenges for the 21st Century. *Communication +1*, n.1, p. 1-27, 2012, 1 v. Disponível em: <<http://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1007&context=-cpo>>. Acesso em: 27 fev. 2018

GUNKEL, David; TAYLOR, Paul. *Heidegger and the media*. Malden: Polity Press, 2014.

HEIDEGGER, Martin. *Que é uma coisa?* Lisboa: Edições 70, 2002.

..... The Provenance of Art and the Destination of Thought. *Journal of the British Society for Phenomenology*, n. 2, p. 119-128, maio. 2013, 44 v. DOI: 10.1080/00071773.2913.11006794.

ILIADIS, Andrew; RUSSO, Federica. Critical Data Studies: an introduction. *Big Data & Society*, n. 7, p. 1-7, jan.-jun. 2016, 1 v. DOI: 10.1177/2053951716674238.

JUNGHERR, Andreas. *Analyzing political communication with digital trace data: the role of Twitter messages in social science research*. London: Springer, 2015.

KRTILOVA, Katarina. Media Matter: Materiality and Performativity in Media Theory. In: HERZOGENRATH, Bernd (Org.). *Media Matter: The Materiality of Media*,

- Matter as Medium. Londres: Bloomsbury Academic, 2015.
- KUHN, Thomas. *A estrutura das revoluções científicas*. 5 ed. São Paulo: Editora Perspectiva, 1998.
- LATOURETTE, Bruno. *Jamais fomos modernos: ensaio de antropologia simétrica*. 3 ed. São Paulo: Ed. 34, 2013.
- \_\_\_\_\_. Agency at the time of the Anthropocene. *New Literary History*, n. 1, p. 1-18, 2014, 45 v. DOI: 10.1353/nlh.2014.0003.
- LEIDLMAIR, Karl. *Künstliche Intelligenz und Heidegger: über der Zwiespalt von Natur und Geist*. Munique: Wilhelm Fink Verlag, 1991.
- MERLEAU-PONTY, Maurice. *Fenomenologia da percepção*. 4 ed. São Paulo: WWF Martins Fontes, 2011.
- NEUMAN, W. Russell; GUGGENHEIM, Lauren; MO JANG, S.; BAE, Soo Young. The Dynamics of Public Attention: Agenda-setting theory meets big data. *Journal of Communication*, n. 2, p. 193-214, 2014, 64 v. DOI: 10.1111/jcom.12088.
- \_\_\_\_\_, Liliana Souza de. A Ideia do Sujeito em Descartes e seu Significado para a Filosofia Moderna. *Revista Sociais e Humanas*, n. 1, p. 68-79, jan.-jun. de 1998, 11 v. Disponível em: <<https://periodicos.ufsm.br/sociaisehumanas/article/view/1049/660>>. Acesso em: 16 mar. 2018.
- PUSCHMANN, Cornelius; BURGESS, Jean. Metaphors of Big data. *International Journal of Communication*, n. 1, p. 1690-1709, 2014, 8 v. Disponível em: <<https://www.hiig.de/wp-content/uploads/2014/07/2169-11849-1-PB.pdf>>. Acesso em: 27 fev. 2018.
- RÜDIGER, Francisco. *Introdução à teoria da comunicação*. 2 ed. São Paulo: Edicon, 2004.
- \_\_\_\_\_. Cibernética (verbetes). In: MARCONDES FI-  
LHO, Ciro (Org.). *Dicionário da Comunicação*. 2 ed. São Paulo: Paulus, 2014.
- SCHROEDER, Ralph. Big Data and Communication Research. *Oxford Research Encyclopedia of Communication*. n. 1, p. 1-20, nov. 2016, 1 v. DOI: 10.1093/acrefore/9780190228613.013.276.
- SHANNON, Claude. A Mathematical Theory of Communication. *The Bell System Technical Journal*, n. 3, p. 379-423, jul. 1948, 27 v. Disponível em: <<http://math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf>>. Acesso em: 16 mar. 2018.
- SHROFF, Gautam. *The Intelligent Web: Search, Smart Algorithms, and Big Data*. Oxford: Oxford University Press, 2014.
- SILVA, Guilherme C. *Elementos para uma crítica da noção cibernética da comunicação: um estudo sobre as premissas cibernéticas da comunicação à luz do avanço da inteligência artificial*. 2016. Monografia (Bacharelado em Comunicação Social - Jornalismo). Faculdade de Comunicação Social, Centro Universitário Adventista de São Paulo - Unasp, 2016.
- \_\_\_\_\_, Michael. *Introduction to the theory of computation*. Boston: PWS, 1997.
- THOMAZ, Suzanne L.; NAFUS, Dawn; SHERMAN, Jamie. Algorithms as Fetish: faith and possibility in algorithmic work. *Big Data & Society*, n. 1, p. 1-11, jan.-jun. 2018, 1 v. DOI: 10.1177/2053951717751552.
- TOMAZ, Tales. Alteridade e Tecnologia: Implicações da Fenomenologia de Martin Heidegger. In: XXXIX CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 2016, São Paulo - SP. *Anais...* São Paulo: Intercom, 2016. p. 1-15.
- \_\_\_\_\_. *Máquina como Outro comunicativo: crítica da concepção cibernética à luz da fenomenologia de Heidegger*. 2017. 182 f. Tese (Doutorado - Progra-

ma de Pós-Graduação em Meios e Processos Audiovisuais) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2017.

WEAVER, Warren. A teoria matemática da comunicação. In: COHN, Gabriel. (Org.). *Comunicação e indústria cultural*. 3ª ed. São Paulo: Nacional, 1977.

WIENER, Norbert. *Cibernética e sociedade: o uso humano de seres humanos*. 2 ed. São Paulo: Cultrix, 1965.

ZUÑIGA, Homero Gil de; WEEKS, Brian; ARDÈVOL-ABREU, Alberto. Effects of the News-Finds-Me Perception in Communication: social media use implications for news seeking and learning about politics. *Journal of Computer-Mediated Communication*, n. 3, p. 105-123, maio. 2017, 22 v. DOI: 10.1111/jcc4.12185.

**Tales Tomaz** – Professor-assistente do Centro Universitário Adventista de São Paulo. Doutor em Meios e Processos Audiovisuais, pela ECA/USP. **E-mail:** tales.tomaz@unasp.edu.br

**Guilherme Cavalcante Silva** – Mestrando em Divulgação Científica e Cultural, no Laboratório de Estudos Avançados em Jornalismo (Labjor/IEL), Universidade Estadual de Campinas (Unicamp). Bolsista Capes. **E-mail:** guilhermecavalcantesilva@outlook.com

Recebido: 16 mar. 2018

Aprovado: 01 maio 2018

# Mediações algorítmicas: o poder de modulação dos algoritmos do Facebook

## Algorithmic mediations: the power of modulation of Facebook's algorithms

**Débora Franco Machado**

Universidade Federal do ABC  
Programa de pós-graduação em ciências humanas e sociais  
Santo André, SP, Brasil

## Resumo

O artigo apresenta uma pesquisa sobre as implicações do uso de processos algorítmicos pelas plataformas de mídias sociais. O objetivo é analisar as dinâmicas e motivações para a criação desses códigos, buscando compreender o que muda em uma comunicação cada vez mais mediada e modulada por agentes humanos e não humanos. Para isso, revisamos a literatura acerca do uso de algoritmos e os debates sobre performatividade e objetividade por autores como Pasquale, Gillespie e Silveira. Também delimitamos o conceito de modulação na sociedade de controle (Deleuze). As hipóteses que guiam a pesquisa nascem do debate emergente sobre o possível poder de modulação comportamental que esses códigos carregam. A metodologia utilizada para a presente pesquisa foi a análise de duas patentes registradas pela empresa Facebook Inc.

## Abstract

The article presents a research about the implications of the use of algorithms by social media platforms. The goal is to analyze the dynamics and motivations for the creation of those codes, seeking to understand what has changed in the communication process that it is getting more and more mediated and modulated by human and nonhuman agents. Therefore, we review the literature on the use of algorithms and debates on performativity and objectivity by authors like Pasquale, Gillespie and Silveira; and we also delimit the concept of modulation in the control society (Deleuze). The hypotheses that lead this research come from the emerging debate about the possible power of behavior modulation that these codes carry. The methodology used for the present research was an analysis of two patents registered by Facebook Inc.

## Palavras-chave

Algoritmos. Modulação. Facebook. Big data. patentes.

## Keywords

Algorithms. Modulation. Facebook. Big data. Patents.

## Introdução

A palavra algoritmo data ao século IX, mas seu significado tem interpretações diversas dependendo da área em que é utilizada (KITCHIN, 2017). Para as áreas técnicas, como a engenharia e a ciência da computação, o significado é simples: uma sequência de ações que pode ser usada por uma máquina para realizar uma tarefa. Para o público em geral a palavra define algo muito mais complexo, intangível e, apesar de algumas discussões no âmbito acadêmico e em alguns veículos de comunicação, ainda é um tema complexo e longe de discussões no dia a dia. Já para os cientistas sociais, a palavra parece definir um perigo emergente que, ao ser discutido, passa distante das equações matemáticas responsáveis por sua origem (GILLESPIE, 2014). Para se aprofundar nas implicações sociais decorrentes da inserção dos processos algorítmicos no cotidiano de grande parte da sociedade, é necessário entender sua complexidade e observar a forma como eles são entendidos dentro da comunidade tecnológica, das ciências sociais e da sociedade civil.

Nos últimos seis anos, o número de brasileiros com acesso à internet cresceu exponencialmente. Somente nas classes C e D, o número de usuários subiu de 18% para 50% (CETIC, 2015). Os brasileiros lideram a lista de tempo gasto nas redes sociais online e gastam em torno de 21,2 minutos por visita. Uma única plataforma, Facebook, consome 96,7% do tempo que os usuários passam nas redes sociais (COMSCORE, 2014). Segundo a empresa, oito em cada 10 brasileiros conectados estão no Facebook, somando um total de 99 milhões de usuários ativos mensais (COSSETTI, 2016).

Apesar do uso diário das mídias sociais ser uma prática comum na maioria dos países com acesso à internet, uma mesma tecnologia pode ter usos diferentes de acordo com a cultura e os processos históricos de cada sociedade. Um bom exemplo pode ser encontrado na pesquisa realizada em 2015 que mostra que 55% dos brasileiros acreditam que o

Facebook é a internet, enquanto apenas 5% dos americanos têm essa visão (QUARTZ, 2015). As práticas de zero-rating<sup>1</sup>, controversas por distorcerem os princípios de neutralidade da rede propostos pelo Marco Civil, incentivam usuários mobile a não saírem dessas plataformas enquanto estão online, uma vez que seu acesso é gratuito em certos planos de telefonia (RAMOS, 2015).

A funcionalidade dos algoritmos de seleção de conteúdo dessas plataformas ganhou relevância na mídia e na academia nos últimos anos, mas esses processos estão presentes em diferentes âmbitos da vida contemporânea que vão muito além das redes sociais online: 80% das transações em alguns mercados de capital americano, por exemplo, são feitas por algoritmos, que exercem o papel de decidir o fluxo do capital e o *credit score* de um cliente utilizando uma vasta base de dados e critérios pré-definidos (PASQUALE, 2015).

Para Mackenzie (2006), os algoritmos que permeiam a rede são performativos. Seus sistemas, ao serem aplicados, não alteram somente o cálculo do valor no mercado financeiro, mas também modelam as dinâmicas, práticas e comportamentos dos profissionais dessa área e das pessoas que serão afetadas por suas decisões – estejam elas cientes de como esse procedimento funciona ou não. Pasquale (2015) pontua que uma das problemáticas desses sistemas é a falta de transparência, visto que não é possível investigar, muito menos auditar, algo que se desconhece. Ele também nos atenta a um fato contraditório: se por um lado as decisões tomadas por empresas e instituições financeiras são escondidas por trás de acordos de não divulgação, por outro, tudo que o usuário faz na rede está sendo monitorado. O autor escolheu o mercado financeiro como seu principal objeto de estudo, mas nos lembra de que “o que as financeiras fazem com o dinheiro, as

1 O Zero Rating é uma prática realizada pelas operadoras e algumas empresas de tecnologia que consiste em não cobrar o tráfego de dados móveis a alguns serviços online, como apps de mídias sociais e mensageiros instantâneos.

empresas de internet líderes no mercado fazem com a atenção. Elas a direcionam a certas ideias, bem e serviços e para longe de outras" (*idem*, 2015, p.6, tradução nossa).

Para Gillespie (2014) "mais do que meras ferramentas, os algoritmos são estabilizadores de confiança" (*idem*, 2014, p.13, tradução nossa). E, para isso, é preciso que sejam vistos como imparciais, objetivos e livres de subjetividade, erro ou tentativa de influência. Contudo, ele refuta a afirmação de que se trata da mesma objetividade exigida aos jornalistas. No caso dos algoritmos, não se trata de uma objetividade imposta por normas institucionais ou éticas, mas se supõe uma neutralidade mecânica, intrínseca da máquina. Mesmo os termos usados para demonstrar relevância, como *top stories*, *trends*, ou "melhores" normalmente são avaliados como neutros, como se a escolha da nomenclatura fosse um processo não sujeito a questões de interpretação. Essa objetividade algorítmica é fortemente difundida, não apenas pelas empresas proprietárias destes algoritmos, como também pela mídia. Deste modo, os algoritmos são, ao mesmo tempo, defendidos como ferramentas que realizam uma avaliação neutra para aqueles que são críticos aos seus resultados, e vendidos como uma ferramenta de promoção seletiva a anunciantes em potencial.

Nick Seaver (2013) aponta que os algoritmos costumam ser entendidos como preocupações estritamente racionais, que juntam as certezas da matemática com a objetividade da tecnologia. Contudo, o mesmo nunca é puramente abstrato e matemático. Apesar da tentativa dos programadores de manter um grau de objetividade, se distanciando de qualquer tipo de influência – inclusive cultural e refletindo contextos locais – o processo de tradução da tarefa ou conhecimento para um sistema algorítmico não se mantém imune a essas interferências (*Ibid.*). Para Kitchin, os algoritmos são criados para propósitos que, na maioria das vezes estão longe de serem neutros: "para criar valor e capital, para impulsionar um comportamento e estruturar prefe-

rências de uma certa forma; e para identificar, selecionar e classificar pessoas" (KITCHIN, 2017, p. 18, tradução nossa).

Seaver enfatiza a necessidade de se estudar não apenas os impactos socioculturais dos processos algorítmicos como também suas especificidades técnicas e critica o uso generalista do termo empregado em frases como o "algoritmo do Google" ou "algoritmo do Facebook". Ele nota que muitas vezes seria mais apropriado utilizar o termo "sistema algorítmico" – sistemas dinâmicos de códigos e pessoas – visto que na maior parte dos casos o código em si é apenas uma pequena parte de um sistema dinâmico muito mais amplo, que envolve códigos e pessoas, um ou mais bancos de dados, negociações e experimentações (SEAVER, 2013).

Para Bernhard Rieder (2017, p.101), se o software se tornou uma técnica de poder é "desconcertante que a análise crítica de objetos, procedimentos e práticas técnicas concretas seja extremamente rara".

## Os algoritmos, as mídias sociais e o conceito de modulação

Atualmente, o foco das pesquisas da área de tecnologia e sociedade voltou-se para o uso de algoritmos que controlam o fluxo de informação dentro das plataformas de mídias sociais, visto que elas já são a principal fonte de informação de grande parte da população conectada (PARISER, 2012). A mais utilizada pelos brasileiros, o Facebook, armazena em torno de 300 *petabytes* de dados dos usuários (JOLER; PETROVSKI, 2016), essenciais para o trabalho de filtragem de informação que seus softwares exercem. A personalização de conteúdo recebido por cada usuário é vital não apenas para permitir que eles recebam as postagens mais relevantes (de acordo com os critérios de relevância da empresa), mas também para atingir os objetivos do mercado publicitário, responsável por 92% da re-

ceita da empresa em 2014 (SILVEIRA, 2017). Para uma análise tão detalhada do perfil de cada usuário a empresa especializa-se em produzir novos softwares que possibilitem um monitoramento intenso do comportamento, dos interesses e da comunicação de quem a utiliza.

Para Foucault (1973), o poder está diretamente ligado à produção de verdade e ao saber. Alain Desrosières (2002), ao analisar a história do pensamento estatístico, que data dos séculos XVII e XVIII, também se dedicou a estudar como uma forma de saber pode se transformar em poder. Ao descrever o início da busca por cálculos probabilísticos, o autor mostra que a questão que essa nova ciência estava tentando solucionar veio de uma tentativa antiga de encontrar algo que possibilitasse especular sobre o que até então estava apenas na mão de Deus, era sagrado: o acaso. “Quer permitindo que ele decida casos difíceis, ou integrando a avaliação de um futuro incerto no presente” (DESROSIERES, 2002, p.46, tradução nossa). Assim, o autor compara o papel do estatístico com o de um juiz.

O ato de classificar, valorizar e avaliar fenômenos naturais a nossa volta não é algo novo. No entanto, uma característica que marca a nossa época é o fato de realizarmos essas tarefas utilizando “ferramentas tecnológicas e virtuais compostas de algoritmos capazes de sintetizar, processar e divulgar dados em uma velocidade e quantidade jamais testemunhadas em nossa história” (LEITE; SARTORE, 2017, p.13).

Uma das características da sociedade informacional é que ela se constitui com tecnologias que comunicam e controlam ao mesmo tempo (SILVEIRA, 2017). Por isso, o controle é um tema tão importante ao estudarmos os processos algorítmicos utilizados nas plataformas de mídias sociais. Para Deleuze, passamos da era das Sociedades Disciplinares, conforme apresentadas por Foucault, e estamos vivendo em uma Sociedade de Controle (DELEUZE, 1992). O medo da punição é substituído por dispositivos de modulação de condutas onde o controle caminha junto à sensação de conforto,

resolve problemas, “melhora a experiência do usuário”, “não geram medo, mas afeto” (SILVEIRA, 2017, p.83). Esse tipo de modulação se tornou fundamental para o marketing.

Depois da captura e armazenamento de dados para processamento e mineração, as empresas formam amostras de perfis similares que servem aos dispositivos de modulação. O que eles fazem? A partir dos gostos, do temperamento, das necessidades, das possibilidades financeiras, do nível educacional, entre outras sínteses, as empresas oferecem caminhos, soluções, definições, produtos e serviços para suas amostras, ou seja, para um conjunto potencial de consumidores que tiveram seus dados tratados e analisados. O sucesso da modulação depende da análise precisa das pessoas que serão moduladas (SILVEIRA, 2017, p.84.).

Uma das características da modulação é a possibilidade de criar um espaço para o individual, dar a sensação de liberdade para o indivíduo enquanto o mantém em um ambiente restrito (HUI, 2015). Foucault (1998), ao estudar o poder disciplinar, enxerga a liberdade como condição de relações de poder. As dinâmicas de uso propostas pelas plataformas de mídias sociais como o Facebook parecem elevar ao extremo o paradoxo da liberdade controlada já analisado por alguns autores ao estudar a modernidade. Elas oferecem ambientes onde o usuário é incentivado ao compartilhamento, mas só recebe a informação que uma série de algoritmos decidiu ser mais relevante para ele. É incentivado a se expressar, mas seguindo regras de conduta, ou escolhendo dentre seis emoções que representem o que está sentindo.

E é esse tipo de controle que consegue ao mesmo tempo restringir e passar a sensação de liberdade que Deleuze chama de modulação.

Os confinamentos são moldes, distintas moldagens, mas os controles são uma modulação, como uma moldagem auto-deformante que mudasse

continuamente, a cada instante, ou como uma peixeira cujas malhas mudassem de um ponto a outro (DELEUZE, 1992, p.221)

Yuk Hui (2015, p.95) caracteriza esse novo tipo de controle, entendido em termos de modulação, pela possibilidade de "criação de um espaço para o indivíduo, como se ele ou ela tivesse a liberdade de se entrelaçar e criar, enquanto sua produção, bem como seus fins, seguem a lógica das forças intangíveis". Para Lazzarato (2006), a modulação como exercício de poder também se ocupa dos corpos, mas é principalmente a dimensão incorporeal que está em jogo. O autor aponta que, ao contrário das sociedades disciplinares foucaultiana, as sociedades de controle "se investem da memória mental, mais do que da memória corporal" (*idem*, 2006, p.84).

## **A análise de patentes como estudo de processos algorítmicos em mídias sociais**

Uma das maiores mudanças que a sociedade em rede (CASTELLS, 2013) proporcionou foi a possibilidade de todo leitor ser também um produtor de conteúdo. Isso transformou a situação de escassez da mídia de massa em uma abundância de dados e conexões, possível somente em uma rede distribuída como a internet. Com isso, a quantidade de dados armazenados na rede, assim como tecnologias para analisá-los, cresceu exponencialmente, dando origem ao termo Big Data. Boyd e Crawford (2012, p.9, tradução nossa) apontam que trabalhar com "Big Data oferece às disciplinas humanistas uma nova maneira de reivindicar o status de ciência quantitativa e método objetivo. Isso torna muitos espaços sociais mais quantificáveis".

Trabalhar com dados ou mesmo analisar redes sociais digitais não é algo novo dentro das ciências sociais. A novidade da pesquisa usando Big Data refere-se ao que alguns pesquisadores chamam de "os quatro Vs": velocidade, variedade, vinculação e

validade da informação (SILVEIRA, 2015). Uma das questões que preocupa pesquisadores do tema é o fato das plataformas de mídias sociais funcionarem como uma caixa preta, em que certas informações são impossíveis de serem coletadas caso a pesquisa não seja realizada em parceria com a empresa que administra a plataforma.

A análise de patentes mostra-se como uma alternativa interessante para entender os processos algorítmicos que ocorrem nas plataformas de mídias sociais. Apesar de muito utilizada nas áreas de inovação como uma forma de mapear tecnologias e analisar tendências, a análise de patentes não é uma metodologia frequente na área das ciências sociais. Portanto, os métodos criados para guiar esse tipo de análise foram desenvolvidos com base na necessidade dos pesquisadores das áreas de inovação, de universidades ou empresas, para analisar grandes bases de dados agrupando informações relacionadas. Essa análise é feita, principalmente, por meio de técnicas de mineração de texto e ferramentas de visualização com o auxílio de softwares específicos, com o objetivo de criar redes, mapas e *clusters* de patentes que tratam de tecnologias similares, ou como forma de hierarquizar a relevância das invenções mais recentes (ABBAS et. al., 2014). Tais técnicas mostram-se pouco eficazes para realizar uma pesquisa como essa, que possui como objetivo a identificação de possibilidades de orientação de comportamento nas invenções selecionadas, características difíceis de serem mapeadas por meio da identificação de palavras-chave ou de uma análise semântica automatizada.

Ao ser questionado pelo website, The Verge, sobre algumas patentes registradas pela Facebook Inc. que chamaram a atenção da mídia, um representante da empresa afirmou que a mesma, frequentemente, registra patentes para tecnologias que nunca são implementadas, e por isso "patentes não deveriam ser tomadas como um indicador de futuros planos" (NEWTON, 2017, online, tradução nossa). Contudo, há uma extensa literatura que

aponta a análise de patentes como um método eficiente para identificar padrões, entender em que direção à equipe de desenvolvimento de produto da empresa está caminhando e acompanhar o tipo de conhecimento que a empresa analisada adquiriu no decorrer dos anos.

Daim et. al.(2006) compara o registro de patentes a artigos publicados em revistas acadêmicas. Eles acreditam que mesmo que poucas patentes venham a se tornar algo com valor comercial, a maioria é tecnicamente significativa, pois elas servem como ponto de partida para o desenvolvimento de novas tecnologias. Breitzman e Moge (2002) mostram que a análise de patentes muitas vezes revela a competência tecnológica central de uma empresa e é, também, uma forma eficaz de projetar em que áreas uma empresa ou tecnologia se concentrará em um futuro próximo.

## As patentes da Facebook inc. e a modulação de comportamento

As primeiras patentes registradas pela Facebook Inc. datam de 2008 e, desde então, a empresa pede anualmente o registro de centenas de patentes para a *United States Patent and Trademark Office*. O texto da maioria das patentes costuma seguir um padrão: primeiro explica-se porque ela foi criada, qual a necessidade dessa invenção e depois se detalha o que ela faz.

Em 2015, a plataforma Facebook lançou a ferramenta *reactions*: botões que representam seis emoções que o usuário pode utilizar para demonstrar o que sentiu ao ler uma postagem em seu *News Feed*, como já fazia com o botão curtir, porém sem expressar de forma clara suas emoções, como tristeza, amor, raiva, entre outros. Desde então, a empresa vem investindo cada vez mais em tecnologias capazes de analisar uma grande quantidade de dados relacionados às emoções e sentimentos dos usuários da rede.

Para a presente pesquisa, a metodologia utilizada foi a análise de patentes. A base de pedidos de patente selecionada para busca e recuperação dos documentos foi o banco de patentes da *United States Patent and Trademark Office*<sup>2</sup>. Foi feita uma busca por patentes da empresa Facebook Inc. que tenham sido publicadas entre um de janeiro de 2017 e 31 de janeiro de 2017 e contenham o termo “emotion” em seu título ou descrição<sup>3</sup>. A busca gerou 39 resultados e optou-se por analisar apenas as duas primeiras patentes da lista. Para a análise das patentes foi feita uma leitura aprofundada dos textos e ilustrações com o objetivo de relacionar as funcionalidades das invenções com o conceito de modulação apresentado neste artigo. Listamos abaixo:

### Techniques for emotion detection and content delivery

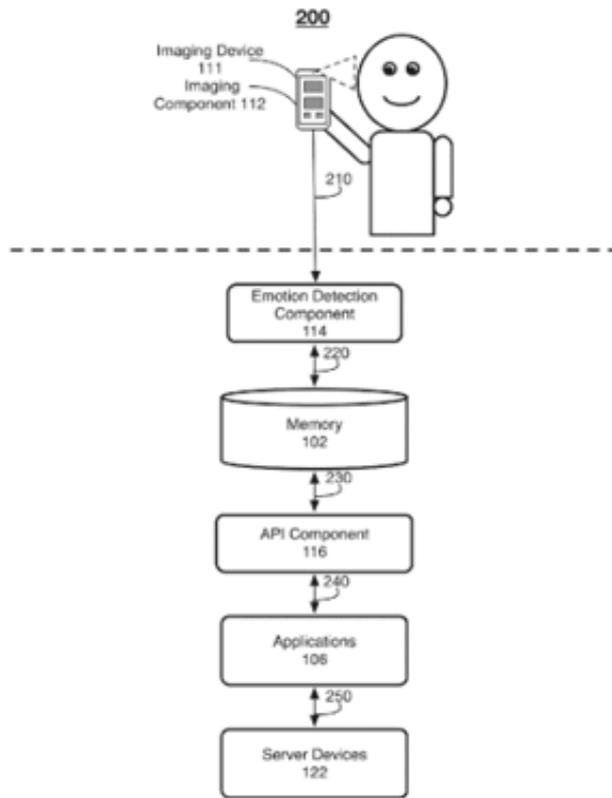
Em português “Técnicas para detecção de emoção e entrega de conteúdo”. A patente mostra uma funcionalidade capaz de detectar qual emoção um usuário está sentindo ao usar uma rede social e assim direcionar a ele postagens identificadas como adequadas para aquele tipo de sentimento. O texto cita um exemplo.

Um usuário pode ver uma série de histórias postadas em uma rede social usando uma aplicação nativa em um celular. À medida que um usuário percorre o conteúdo o aplicativo pode solicitar informações sobre o tipo de emoção. Quando um usuário tiver escolhido uma foto específica para visualizar, um tipo de emoção atual pode estar associado à foto. Em um exemplo, uma foto de um cachorrinho pode estar associada a um tipo de emoção feliz. Em ou-

<sup>2</sup> A busca foi feita por meio do site *freepatentsonline*, que possui a mesma base de dados do *United States Patent and Trademark Office*, porém conta com uma melhor usabilidade e mecanismo para busca.

<sup>3</sup> A busca foi feita usando o atalho AN/“Facebook Inc”“emotion” ISD/1/1/2017->12/31/2017.

tro exemplo, uma foto de uma ex-namorada pode estar associada a um tipo de emoção triste (NAVEH, 2014, p.5, tradução nossa).



**Figura 1 - Ilustração do fluxo de dados do sistema.**  
Fonte: Barak Naveh

O texto aponta que fotos é apenas um tipo de conteúdo que pode ser mostrado, mas esse pode variar entre postagens de texto, vídeo, áudio, games ou mesmo anúncios publicitários. A descrição segue com outros exemplos relevantes e mostra que “ao usar informações associando emoções com conteúdo, um sistema de entrega de conteúdo pode fornecer mais informações associadas a emoções felizes do que emoções tristes, por exemplo,” (NAVEH, 2014, p 5, tradução nossa), ou mesmo “se um usuário for identificado como entediado, um con-

teúdo engraçado poderá ser mostrado a ele” (*Ibid.*, tradução nossa).

Segundo a patente, o sistema também pode ser útil para o mercado publicitário, pois “o estado emocional do usuário poderá ser utilizado para apresentar propagandas” (NAVEH, 2014, p.6, tradução nossa), uma vez que a ferramenta consegue identificar quando o usuário está olhando para a tela do celular ou não. Assim, é possível solicitar que anúncios apareçam apenas quando o usuário estiver atento, ou mesmo que quando o usuário demonstrar tristeza, anúncios inspiradores apareçam, e quando estiver entediado, anúncios interativos apareçam.

Ao analisar o texto da patente é possível observar a intenção de modulação de comportamento embutida no sistema. Para exemplificar a sua funcionalidade o autor propõe que, ao identificar que um usuário está triste, a mudança de comportamento desejada seria deixá-lo feliz. Sentimentos como tristeza e tédio são vistos como indesejados e a alteração dos mesmos é vista como parâmetro de sucesso. A hipótese que pode ser levantada, levando em consideração que a rentabilidade da empresa depende da permanência do usuário atento à plataforma, é a de que o sentimento desejado é aquele que mantém o usuário conectado.

Para Silveira (2017, p.85), “a eficiência da modulação parece depender da análise correta e em profundidade do padrão de comportamento, as vontades e as necessidades dos indivíduos que, como bem apontou Deleuze, ‘tornaram-se individuais’”. Identificando qual emoção um usuário está sentindo, qual tipo de conteúdo ele costuma consumir nesse estado emocional, e consequentemente, registrando o que ele sente após consumir certos tipos de conteúdo, é possível ter uma base de dados valiosa para processos algorítmicos com o poder de direcionar o usuário a sentir ou deixar de sentir emoções definidas pela plataforma.

## Augmenting text messages with emotion information

A patente, em português “Aumentando mensagem de texto com informação de emoção”, trata de uma funcionalidade que permite que um sistema identifique emoções relacionadas a mensagens digitadas pelo usuário e que formate esse texto de acordo com a informação adquirida. O sistema prevê a emoção que o usuário está sentindo e define uma formatação de texto capaz de expressar melhor, visualmente, esse sentimento.

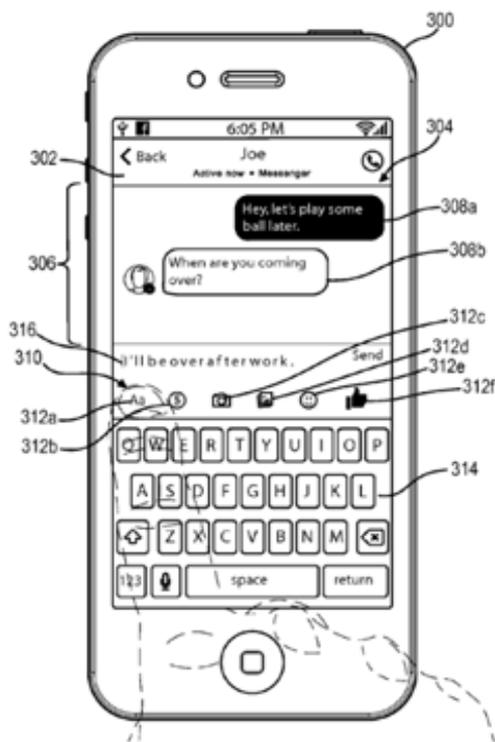


Figura 2 - Ilustração da interface do sistema.

Fonte: Aran Donohue

Para prever o que o usuário está sentindo ao digitar a mensagem, a ferramenta coleta diversas informações que vão além do conteúdo redigido por ele, por exemplo, a velocidade da

digitação, a pressão do toque ao digitar, padrões de interação do usuário com o aplicativo ou mesmo a sua localização no momento da digitação.

Por exemplo, se o remetente digitar lentamente, o sistema de mensagens pode determinar que o remetente sente que a mensagem não é urgente. Em formas alternativas, o sistema de mensagens pode prever outras emoções com base na velocidade de digitação associada à entrada de informação pelo teclado, como se o remetente está interessado na conversa se está com pressa (DONOHUE, 2015, p.7, tradução nossa).

Dentre as alterações que o sistema pode fazer a patente cita modificar o formato do balão de fala, o plano de fundo da aplicação de mensagens ou mesmo “fornecer informações adicionais junto à mensagem para explicitamente ou implicitamente indicar a emoção prevista” (DONOHUE, 2015, p.8, tradução nossa). Para demonstrar que o usuário não está com pressa e por isso está digitando lentamente, o sistema pode, por exemplo, espaçar as letras das palavras digitadas.

Com base na localização do usuário o sistema pode prever que ele esteja cansado se ele enviar a mensagem de dentro de uma casa noturna entre às 12am e às 6am, e a partir dessa informação formatar a mensagem de acordo. Já com base na pressão que o usuário utiliza para digitar, o sistema pode prever que o usuário está ansioso ou com raiva (ao pressionar forte) ou se ele está calmo (ao pressionar fraco). Usando o microfone do celular, é possível prever que o usuário está dançando em uma festa com música alta. Usando a câmera do celular o sistema pode identificar expressões faciais do usuário. Além disso, ele pode mostrar essa formatação antes do usuário enviar a mensagem ou apenas depois que a mensagem já foi enviada e recebida.

Por fim, é possível observar a intenção da empresa em adicionar mais uma camada de intervenção automatizada na comunicação do usuário, diminuindo sua autonomia e decidindo por ele quais informações são mais relevantes e merecem mais destaque. A classificação de conteúdos por relevância já se mostrou como um objetivo projetado pela Facebook Inc. desde suas primeiras patentes, como a denominada *Generating a feed of stories personalized for members of a social network* requerida em 2006, e é uma prática que acompanha a evolução do Feed de Notícias do Facebook (ARAÚJO, 2017).

Para Gillespie, “relevância” é um termo fluido e aberto a interpretação. O autor mostra um entendimento do termo – e dos mecanismos que o utilizam – que vai à contramão do discurso da objetividade propostos pelas empresas de tecnologia e plataformas de mídias sociais.

Como não há uma métrica independente para o que realmente são os resultados de pesquisa mais relevantes para qualquer consulta, os engenheiros devem decidir quais resultados parecem “corretos” e ajustar seu algoritmo para atingir esse resultado, ou fazer alterações com base em evidências de seus usuários, tratando cliques rápidos e sem pesquisas na sequência como uma aproximação, não exatamente de relevância, mas de satisfação (GILLESPIE, 2014. p. 10, tradução nossa).

Para a empresa, sistemas como esse são importantes, pois permitem “que o destinatário da mensagem interprete com mais precisão a mensagem à luz das emoções do remetente” (DONOHUE, 2015, p.5, tradução nossa), contudo, o texto da patente não deixa claro como essa precisão é avaliada.

## Considerações finais

A interação por meio de plataformas de mídias sociais há tempos ocupa uma parcela significativa de tempo no dia a dia nos brasileiros. Devido à crescente quantidade de informação armazenada por essas plataformas e ao modelo de negócio que as mantém no topo das empresas mais valiosas no mundo<sup>4</sup> – obtendo valor a partir dos dados dos usuários e do tempo que eles se mantêm conectados e em constante interação – seu funcionamento depende de agentes que automatizem a entrega e a organização da informação que circulam dentro delas. Não há dúvidas de que o sucesso e a eficácia dessas plataformas dependem, em boa parte, de seus processos algorítmicos. Portanto, é importantíssimo que pesquisadores na área da comunicação e das ciências sociais se dediquem a estudar e/ou aprofundar as consequências dessa nova e efêmera dinâmica social. Para isso, é preciso entender os algoritmos como atores performativos e o usuário como um agente ativo, não passivo, nessa interação.

A presente pesquisa mostra como a análise de patentes é uma metodologia relevante para entender à lógica que acompanha a criação desses algoritmos, acompanhar o conhecimento tecnológico adquirido por essas empresas e identificar padrões nas aplicações dos mesmos. Além de ser uma alternativa para tentar conhecer uma parcela maior desses mecanismos que são, em sua maior parte, opacos e fechados, e impossibilitam que pesquisadores tenham um conhecimento mais aprofundado sobre o seu funcionamento. Contudo, a metodologia também possui algumas restrições: apesar da descrição das invenções ser detalhada, a forma com que elas podem ser aplicadas pela empresa varia e muitas vezes são genéricas.

<sup>4</sup> Fonte: Brand Finance Global 500 2018. Disponível em: <http://brandfinance.com/knowledge-centre/reports/brand-finance-global-500-2018/>. Acesso em 25 abr. 2018.

A bibliografia estudada e as patentes analisadas apontam para um poder de modulação de comportamento presente nessas plataformas que merece uma análise crítica. Esse artigo se propõe apenas a uma pequena contribuição para um debate emergente, interdisciplinar e essencial sobre os algoritmos que operam algumas das plataformas mais utilizadas no Brasil e no mundo.

## Referências

ABBAS, Assad; KHAN, Samee; ZHANG, Limin. A literature review on the state-of-the-art in patent analysis. *World Patent Information*, North Dakota, p.1, nov. 2014. Disponível em: <<http://dx.doi.org/10.1016/j.wpi.2013.12.006>>. Acesso em: 20 abr. 2018

ARAÚJO, Willian Fernandes. *As narrativas sobre os algoritmos do Facebook: uma análise dos 10 anos do feed de notícias*. 2017. Tese (Doutorado em Comunicação e Informação) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2017.

BOYD, Danah; CRAWFORD, Kate. Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. *Information, Communication, & Society* p. 662-679, 2012. Disponível em: <[https://people.cs.kuleuven.be/~bettina.berendt/teaching/ViennaDH15/boyd\\_crawford\\_2012.pdf](https://people.cs.kuleuven.be/~bettina.berendt/teaching/ViennaDH15/boyd_crawford_2012.pdf)>. Acesso em: 20 abr. 2018.

BREITZMAN, Anthony; MOGEE, Mary Ellen. The many applications of patent analysis. *Journal of Information Science*, Nova York, v. 28, n. 3, p. 187-205, 2002. Disponível em: <http://journals.sagepub.com/doi/abs/10.1177/016555150202800302>>. Acesso em: 20 abr. 2018.

CASTELLS, Manuel. *A Sociedade em Rede: Economia, Sociedade e Cultura*. 6. ed. São Paulo: Paz e Terra, 2013, 1 v.

CETIC. *TIC Domicílios*. 2015. Disponível em: <[http://data.cetic.br/cetic/explore?idPesquisa=TIC\\_DOM](http://data.cetic.br/cetic/explore?idPesquisa=TIC_DOM)>. Acesso em: 20 abr. 2018.

COMSCORE. *The State of Social Media in Brazil and the Metrics that Really Matter*. Disponível em: <<http://www.comscore.com/Insights/Presentations-and-Whitepapers/2014/The-State-of-Social-Media-in-Brazil-and-the-Metrics-that-Really-Matter>>. Acesso em: 20 abr. 2018.

COSSETTI, Melissa. *Facebook revela dados do Brasil na CPBR9 e WhatsApp 'vira ZapZap'*. *Techtudo*. 2016. Disponível em: <<http://www.techtudo.com.br/noticias/noticia/2016/01/facebook-revela-dados-do-brasil-na-cpbr9-e-whatsapp-vira-zapzap.html>>. Acesso em: 20 abr. 2018.

DAIM, Tugrul; RUEDA, Guillermo; Martin, Hilary; GERDSRI, Pisek. Forecasting emerging technologies: Use of bibliometrics and patent analysis. *Technological Forecasting & Social Change*, Oxford, v. 73, n. 8, p. 981-1012, out. 2006.

DELEUZE, Gilles. Post-Scriptum sobre as Sociedades de Controle. In: *Conversações*. São Paulo: Editora 34, 1992.

DESROISIÈRES, Alain. *The politics of large numbers: a history of statistical reasoning*. Cambridge: Harvard University Press, 2002.

DONOHUE, Aran. *Augmenting Text Messages with Emotion Information*. US 2017/0147202, 25 maio 2017.

FOUCAULT, Michel. *A verdade e as formas jurídicas*. Rio de Janeiro: NAU Editora, 1973.

\_\_\_\_\_. *A história da sexualidade 2: o uso dos prazeres*. Rio de Janeiro: Graal, 1998.

GILLESPIE, Tarleton. *The Relevance of Algorithm. Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA: MIT Press. Virtual Books 2014. Disponível em: <ht-

- [tps://www.academia.edu/2257984/The\\_Relevance\\_of\\_Algorithms](https://www.academia.edu/2257984/The_Relevance_of_Algorithms)>. Acesso em: 20 abr. 2018.
- HUI, Yuk. *Modulation After Control. New formations: a journal of culture/theory/politics*, Londres, v. 84-85, p.74-91, 2015.
- JOLER, Vladan; PETROVSKI, Andrej. *Immaterial Labour and Data Harvesting*. Share Lab. 2016. Disponível em: <<https://labs.rs/en/facebook-algorithmic-factory-immaterial-labour-and-data-harvesting/>>. Acesso em: 20 abr. 2018.
- KITCHIN, Rob. Thinking Critically About and Researching Algorithms. *The Programmable City Working Paper 5*. 2017. Disponível em: <<https://ssrn.com/abstract=2515786>>. Acesso em: 20 abr. 2018.
- LAZZARATO, Maurizio. *As revoluções do capitalismo*. Rio de Janeiro: Civilização Brasileira, 2006.
- LESSIG, Lawrence. *Code version 2.0*. Nova York: Basic Books, 2006. Kindle Edition.
- MACKENZIE, Donald. *An engine, not a camera: How financial models shape markets*. Cambridge: MIT Press, 2006.
- NAVEH, Barak. *Techniques for emotion detection and content delivery*. US 2015/0242679, 27 ago. 2015.
- NEWTON, Casey. *A boredom detector and 6 other wild Facebook patents*. The Verge. Disponível em: <https://www.theverge.com/2017/6/12/15772342/facebook-patents-boredom-detector-dating-delivery-app>>. Acesso em: 20 abr. 2018.
- PARISER, Eli. *O Filtro Invisível: o que a internet está escondendo de você*. Rio de Janeiro: Jorge Zahar, 2012. Kindle Edition.
- PASQUALE, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press, 2015.
- QUARTZ. *Millions of Facebook users have no idea they're using the internet*. 2015. Disponível em: <<http://qz.com/333313/milliiions-of-facebook-users-have-no-idea-theyre-using-the-internet/>>. Acesso em: 20 abr. 2018.
- RAMOS, Pedro Henrique. Zero-rating: uma introdução ao debate. *PolITICs* Disponível em: <<https://politics.org.br/edicoes/zero-rating-uma-introdu%C3%A7%C3%A3o-ao-debate>>. Acesso em: 20 abr. 2018.
- RIEDER, Bernhard. Scrutinizing an algorithmic technique: the Bayes classifier as interested reading of reality. *Information, Communication & Society*, n. 1, p. 100-117, 2017, 20 v.
- SARTORE, Marina; LEITE, Elaine. Desconstruindo os dispositivos dos mercados: aportes da Sociologia Econômica. *Revista TOMO*, 2017. Disponível em: <<https://seer.ufs.br/index.php/tomo/article/viewFile/6708/5475>>. Acesso em: 20 abr. 2018.
- SEAVER, Nick. *Knowing Algorithms*. *Media in Transition 8*, Cambridge, MA. April 2013. Disponível em: <<https://nick-seaver.squarespace.com/s/seaverMiT8.pdf>>. Acesso em: 20 abr. 2018.
- SILVEIRA, Sergio Amadeu. Governo dos Algoritmos. *Revista de Políticas Públicas*, v. 21, n. 1, p. 267-282, 2017. Disponível em: <<http://www.periodicoseletronicos.ufma.br/index.php/rppublica/article/view/6123>>. Acesso em: 20 abr. 2018.
- ..... *Tudo Sobre Tod@s: Redes digitais, privacidade e dados pessoais*. São Paulo: Edições SESC, 2017. Kindle Edition.

..... *Ser Cientista Social Hoje*, no século XXI. In: *O que é ser Cientista Social*. / Adriana Capuano de Oliveira e Marilda Aparecida de Menezes (Orgs.). São Bernardo do Campo: EDUFABC, 2015.

**Débora Franco Machado** – Mestranda em Ciências Humanas e Sociais na UFABC e pesquisadora no Laboratório de Tecnologias Livres da UFABC. **E-mail:** [deborafmachado@gmail.com](mailto:deborafmachado@gmail.com)

Recebido: 14 mar. 2018

Aprovado: 01 maio 2018



# Topologias de espaços híbridos na era da Internet das Coisas

## Topologies of hybrid spaces in the Era of Internet of Things

**Fernanda da Costa Portugal Duarte**

Universidade Federal de Minas Gerais  
Departamento de Comunicação Social  
Belo Horizonte, BR, Brasil

## Resumo

Este artigo discute como a mineração de dados produz representações espaciais multi-escalares na contemporaneidade: da escala íntima dos dados fisiológicos individuais à dimensão global dos modelos preditivos da *big data*. Para compreender os arranjos entre essas escalas, caracterizamos a mineração de dados como ação epistêmica da materialidade digital e, portanto geradora de formas de saber algorítmicas. A partir daí, revisitamos o conceito de espaços híbridos (SOUZA E SILVA, 2006) e analisamos como diversas práticas de coleta, análise e reapropriação de dados constituem topologias fundadas sobre práticas de anotação espacial (*data annotation*), visualização de dados (criação de *data doubles*) e modelos preditivos.

## Abstract

This paper discusses how data mining produces multiscale spatial representations in the contemporaneity: from private scale of individual physiological data to the global dimension of predictive models of big data. In order to understand the arrangements among these scales, we characterize data mining as an epistemic action from digital materiality; therefore, it is generator of algorithmic modes of knowledge. Based on this, we revisited the concept of hybrid spaces (SOUZA E SILVA, 2006) and analyzed how several practices of collect, analysis and reappropriation of data constitute topologies grounded on practices of data annotation, data visualization (creation of data doubles) and predictive models.

## Palavras-chave

Espaços híbridos. Internet das coisas. Big data. Episteme digital.

## Keywords

Hybrid spaces. Internet of Things. Bid data. Digital episteme.

## Introdução

Sensores, tecnologias de comunicação sem fio e recursos de geolocalização criaram uma arquitetura de rede de dados móvel e ubíqua que é nomeada por Kevin Ashton como Internet das Coisas, ou IoT (ASHTON, 2009). Desenvolvimentos recentes na computação pervasiva, tais como microchips de geolocalização e outros dispositivos de biotecnologia embutidas em corpos humanos e não humanos adotam práticas de mapeamento que recolhem informações sobre a mobilidade desses corpos no espaço geográfico. Além disso, informam também sobre a performance interna do corpo físico que habita esse espaço. Sensores para monitoramento das funções fisiológicas, tais como implantes que monitoram o nível de glicose em pacientes diabéticos, ou sonares para facilitar a mobilidade de deficientes visuais são exemplos de práticas tecnológicas que atrelam seu funcionamento ao mapeamento fisiológico do corpo. Estas práticas nos dizem sobre formas emergentes de mapeamento e formas de conhecimento espacial que desafiam as noções de lugar, como mera coordenada geográfica, e o corpo conectado, como um nó em uma rede.

A inclusão de redes de sensores no espaço urbano, casas, objetos e nossos corpos biológicos cria espaços híbridos (SOUZA E SILVA, 2006) ou territórios informacionais (LEMOS, 2007) multiescalares. Espaços híbridos são definidos por Adriana de Souza e Silva (2006, p.262) como “espaços móveis, criado pelo movimento constante de usuários que carregam dispositivos móveis continuamente conectados à Internet e a outros usuários”. André Lemos (2007) discute a emergência de territórios informacionais também a partir da hibridização entre experiências espaciais físicas e telemáticas. A construção deste terceiro lugar híbrido não existe *a priori* ou de forma desarticulada de suas dimensões informacionais ou geográficas. Os territórios informacionais estão imbricados nos

pertencimentos simbólicos, afetivos, sociais, culturais e políticos das localidades geográficas ao mesmo tempo em que atuam como linhas de força produtivas destes pertencimentos.

Estudos que posteriormente foram reconhecidos como característicos de uma “virada espacial” nas teorias sociais apresentaram uma abordagem renovada sobre o conceito de espaço como uma construção social relevante para as diferentes histórias dos sujeitos e para a produção de fenômenos culturais (WARF; ARIAS, 2008, p.1). A “virada espacial” reproblematicizou a noção de espaço, que até então tinha sido desfavorecida por visões historicistas da cultura que privilegiavam sua dimensão temporal. Mais recentemente, estudos no campo de mobilidades (Creswell, 2002, Dourish, 2006, Massey, 1992) reposicionam as noções de espaço e lugar como práticas sociais corporificadas (ou *embodied social practices*) produzidas pela articulação entre cultura e tecnologia. Nesta ótica, a noção de espaço, em sua significação simbólica, assume uma constituição processual, seja ele, por exemplo, uma construção física (espaços arquitetônicos), ou atravessado por tecnologias midiáticas (espaços híbridos produzidos por tecnologias de geolocalização).

Neste sentido, quando confrontados com o paradigma da IoT, somos compelidos a descentralizar tecnologias de geolocalização como ponto de partida para a construção de espaços híbridos e considerá-la apenas um dos diversos parâmetros recolhidos através da mineração de *big data*. Neste artigo, argumentamos que a produção de espaços híbridos é animada por uma ação epistêmica própria da materialidade digital. Ocorre por meio de um continuum de topologias articuladas em múltiplas escalas de dados—de uma pequena escala de dados pessoais encadeado à larga escala de dados de muitos usuários armazenados em servidores.

## A ecologia da internet das coisas

Apesar do termo Internet das Coisas ser comumente adotado pela indústria para denominar pesquisa e desenvolvimento de “objetos inteligentes”, sensores e atuadores conectados à internet, entendemos a IoT como uma ecologia midiática animada pelo imaginário social da computação ubíqua, moldada pelas apropriações e subversões de tecnologias digitais assim como variadas formas de ativismo digital.

Historicamente, a IoT emerge de iniciativas coletivas e concomitantes: pesquisas de desenvolvimento de produto em computação pervasiva e ubíqua por empresas como Xerox e IBM; pesquisas experimentais por agências governamentais e laboratórios acadêmicos, como DARPA e MIT; e iniciativas populares como o movimento *hacker* e suas diversas manifestações em *hacklabs*, *hackerspaces* e *makerspaces*.

O imaginário cultural de uma ecologia de rede que congrega humanos e máquinas de forma simbiótica é descrito por Mark Weiser (1991), em *The computer of the 21st century*, como uma arquitetura de rede acessível por múltiplos painéis de acesso de tamanhos e capacidades variadas instaladas no espaço físico do usuário. Weiser, Gold & Brown (1999) descrevem este novo paradigma da interface humano-computador como mais focada na usabilidade humana e na construção de ambientes em rede cuja capacidade de processamento era compartilhada entre diversos pontos da rede (ao invés de restrita à um terminal computacional tipo *desktop*). O que eles chamam de computação calma implica em uma rede virtualmente disponível de forma ubíqua, pervasiva e naturalmente imbricada em todas as atividades humanas, de forma que o aparelho computacional e sua infraestrutura desapareça da vista imediata e apenas o serviço prestado por ela seja perceptível. Para o alcance destes objetivos, o design das interfaces humano-computador deveria seguir os

seguintes critérios: primeiro, a infraestrutura de coleta e processamento de dados digitais deve ser embarcada em objetos de forma que sua interação possa ser tátil e não limitada às telas; para isso o *hardware* deve ser de tamanho mínimo e não perceptível pelo usuário. Segundo, estes objetos físicos devem ter suas utilidades expandidas na medida em que são dotados de capacidades computacionais; suas interfaces devem ser de fácil manipulação pelo usuário. Terceiro, a ubiquidade dos serviços computacionais depende da conectividade de rede compartilhada entre aparelhos e usuários. Em resumo, a visão de computação calma, se realizaria quando a sensação de transparência entre as relações sociais e as interações entre máquinas fosse estabelecida.

Ainda que esta visão tenha animado o imaginário cultural sobre o futuro da comunicação e computação em rede, a manifestação atual desta rede na forma da IoT, diverge da transparência idealizada por Weiser (1991). A IoT se forma em um cenário tecnológico em meio às diversas formas de regulação e literação digital, onde o *affordance* entre a infraestrutura digital e seus usuários prevalecem sobre modelos teóricos homogênicos, onde as especificidades culturais de apropriação tecnológica conformam diversas formas de uso e adoção dessas interfaces, em oposição ao design das interfaces propostas por Weiser, Gold & Brown (1999).

A emergência de formas criativas e subversivas de apropriação tecnológica por *hackers* e *makers* - diversas dos usos funcionais previstos pela pesquisa experimental acadêmica ou de desenvolvimento de produtos - é evidência de que a IoT conforma-se em uma ecologia múltipla. Desde a década de 1960, a cultura *hacker* articula variadas práticas computacionais—que beiram o limiar da legalidade e criminalidade—como forma de exercício da liberdade de informação e expressão política de apropriação tecnológica. Baseado nos princípios de descentralização da informação das instituições, coletividade e meritocracia, *ha-*

ckers formaram comunidades online e offline para compartilhar conhecimento técnico, ferramentas e colaborar no desenvolvimento de projetos (Levy, 2001). De forma similar com que *hackers* nas décadas de 1980 e 1990 se organizaram para experimentar programação em sistemas *open software*, como Linux; nos anos 2000, espaços físicos chamados *hackerspaces* e *makerspaces* surgiram para explorar o potencial do *open hardware*, impressão 3D e o uso de sensores (Magias, 2000). Os *hackerspaces* e *makerspaces* atuais compartilham, em muitos casos, dos princípios geridos pela cultura *hacker*, e são atores que compõem a ecologia da IoT.

A ecologia da IoT comporta-se, portanto como uma rede de atores humanos e não humanos (Latour, 2002), cujos comportamentos são afetados contiguamente, pelas agências que compartilham. Tal abordagem é rica para os estudos de comunicação, porque descentraliza os estudos de mídia dos aparelhos tecnológicos e os redireciona para as relações que emergem entre as tecnologias e os sujeitos. Ao adotarmos esta direção, podemos efetuar um deslocamento daquilo que entendemos por mídia, do círculo antropocêntrico para o campo dos afetos: entre corpos e corpos, corpos e discursos, signos e signos, máquinas corporais ou incorporais. Tal gesto torna possível trabalhar a comunicação tanto no campo de sua discursividade, enunciabilidade e materialidade, em uma visão ecológica, no sentido dos afetos produzidos nas relações entre humanos e não humanos, entre todo tipo de máquinas.

Definimos a IoT a partir das suas capacidades em agenciar múltiplas topologias de dados, em escalas individuais e globais - sustentados pela contínua e automática coleta e processamento de metadados - pela mineração contínua de *big data*, com o intuito de promover análises preditivas. No entanto, problematizamos essa noção ao focar não em sua eficiência e precisão na identificação de padrões entre dados, mas ao

discutir os processos de formação e reprodução dessa episteme digital, suas consequências políticas e sociais.

## Mineração de *big data* e a Episteme digital

Entendemos que a ecologia midiática que caracteriza a IoT é sustentada pela lógica material de mineração de *big data*. *Big data* é um fluxo dinâmico de dados digitais não esquematizados, flexíveis, refinados, de alta resolução, que são passíveis de serem relacionados entre si para geração de correlações indutivas. A *big data* não diz respeito apenas a um grande volume de dados, mas também aos mecanismos de coleta e estratégias de processamento destes dados, pela alta velocidade de coleta e automação do processamento de dados em tempo real, e a grande variedade de fontes de dados não estruturados. Sensores espalhados no espaço físico medem dados ambientais, como temperatura e umidade; dados demográficos medem fluxo de trânsito e densidade populacional, e dados pessoais indicam a geolocalização de indivíduos específicos e dados fisiológicos. A mineração de *big data* procura capturar dados como estes, de usuários, populações ou sistemas, e manter sua natureza desestruturada, não-esquemática e relacional para combinar diferentes conjuntos de dados para criação customizável de análise de dados (Boyd & Crawford, 2012). Devido à complexidade destes dados, eles apenas são passíveis de serem processados por algoritmos ou formas semelhantes de "inteligência acionável". Ou seja, o cérebro humano é incapaz de promover análises diante desses grandes grupos de dados e suas análises devem ser necessariamente terceirizadas a algoritmos. A *big data* é como um corpo vivo pulsante de dados numéricos discretos, de escalabilidade modular, que são constantemente rearticulados, reorganizados e correlacionados com outros corpos de dados.

A aspiração da *big data* é alcançar o poder preditivo de chegar às correlações algorítmicas que não haviam sido imaginadas. Neste aspecto, Mark Andrejevic e Kelly Gates (2014) discutem que o potencial de usabilidade de qualquer dado digital é apenas especulativo, uma vez que ele só é aferido em conjunto com outros grupos de dados. Dados agrupados hoje podem apenas se tornar relevantes no futuro, quando novos grupos de dados são coletados e um padrão ou tendência são identificados. E por este motivo justifica-se coletar o maior volume de dados possível, a todo tempo, tendo em vista a maior completude de acesso a dados analógicos transcodificados; mesmo que inicialmente alguns dados pareçam irrelevantes.

A pervasividade das tecnologias digitais e a emergência de aplicações de IoT contribuem para a naturalização de parâmetros quantificados que caracterizam dados digitais como uma medida objetiva da realidade. A alta resolução de dados digitais, aliada às correlações indutivas trazidas pela análise algorítmica da *big data*, confere aos dados digitais uma aura de “verdade objetiva”, capaz de efetivamente aferir a realidade através de agregação e correlação de grande volume de dados. Por consequência, a análise algorítmica da *big data* pode associar a agregação de dados informacionais com a construção inequívoca de conhecimento sobre a realidade. O mito que associa dados algoritmos à verdade objetiva falha em reconhecer que a representatividade numérica do digital existe apenas enquanto inserido em um regime de representação. A mineração de dados é ação de um processo de construção de conhecimento, e, portanto, uma articulação de um sistema epistêmico; mas não um fim em si mesmo. A correlação de dados é usada para formar hipóteses através de uma abordagem indutiva; “mais do que validar uma teoria através da análise dos dados relevantes, as novas formas de análise de dados procura obter *insights* originários do próprio corpo de dados” (Kitchin, 2014, p.

2)<sup>1</sup>. Baseada na correlação de dados, é possível identificar o que está acontecendo no grupo de dados; no entanto, a natureza estatística das correlações não inclui a informação contextual que explica os comportamentos descritos. Enquanto algoritmos podem executar correlação de dados, as interpretações epistemológicas dos resultados e as formas de conhecimento que surgem desses resultados são mais complexas dos que os aspectos descritivos que estes dados podem gerar.

A partir do reconhecimento da mineração de *big data* como a lógica material que molda a ação epistêmica das práticas espaciais da IoT, propomos uma expansão do entendimento da dinâmica constituidora de espaços híbridos (de Souza e Silva, 2006) de forma a reconhecer a contribuição das dinâmicas de coleta, processamento e análise de *big data* como práticas de representação espacial. Argumentamos que a IoT opera em escalas de topologias integradas (da micro escala das nanotecnologias embarcadas em corpos aos modelos de cidades inteligentes), na medida em que a partir da coleta de dados pessoais, transforma o indivíduo (e a performance fisiológica e social de seu corpo biológico) em um espaço a ser mapeado, e a partir da correlação em nível populacional, desenha projeções normativas e preditivas a respeito das formas de ação sobre estes espaços.

## Topologias de espaços híbridos

Adriana de Souza e Silva (2006) explica que a noção de espaço híbrido emerge dos usos das interfaces móveis, entendidas como interfaces que permitem a conectividade à internet móvel, de forma contínua enquanto explora-se o espaço geográfico. Na medida em que interfaces móveis são investidas de tecnologias de geo-localização, elas passam a mediar às relações espaciais físicas agora também

1 “rather than testing a theory by analyzing relevant data, new data analytics seek to gain insights ‘born from the data’”. Tradução livre.

atravessadas por camadas de dados digitais. Um exemplo dessa mediação é o jogo locativo *BotFighters*. *BotFighters* é um MMORPG<sup>2</sup> criado para ser jogado em um dispositivo móvel e num ambiente urbano. Segue o gênero de jogo de tiro em primeira pessoa no qual os jogadores assumem um avatar de um robô, cujo objetivo é eliminar os outros jogadores. Para alcançar este objetivo, jogadores têm que se mover pelo espaço urbano e “atirar” em outros jogadores ou participar de batalhas coletivas. A precisão dos tiros varia de acordo com a escolha da arma e a proximidade geográfica entre atirador e alvo. Neste exemplo, a dinâmica de jogo é fundada sobre a interação entre espaços físicos e informacionais e apenas toma forma a partir da participação coletiva entre jogadores. Do mesmo modo, o *smartphone* é caracterizado como agente tecnológico móvel que é a porta de entrada que viabiliza a mediação deste espaço híbrido.

Diante da riqueza da diversidade de agenciamentos sociotécnicos que se articulam como a IoT, somos agora demandados a descentralizar o papel aferido às tecnologias de geolocalização como vetores viabilizadoras das experiências espaciais-informacionais e expandir o entendimento das formação de espaços híbridos para abarcar a ação epistêmica da *big data*. Dispositivos móveis de tela, como *smartphones*, são uma das ferramentas de visualização e manipulação de dados georeferenciados, mas não são mais os únicos métodos para acesso e representação de dados espaciais. A mineração de *big data* é ação material que organiza simbolicamente os comportamentos lógicos das tecnologias sensoriais digitais, que estão agora dispersas em espaços físicos públicos—por exemplo, no uso de acelerômetros para monitorar os níveis dos rios e de sensores infravermelhos que mensuram o volume de trânsito nas estradas; embarcados em espaços físicos domésticos; em câmeras de segurança; em sistemas automáticos de controle de temperatura

e luminosidade; e, em corpos biológicos, como no uso de computação vestível para monitoramento de atividade física. Por consequência, a construção de espaços híbridos na IoT não é mais amarrada à relação direta entre geolocalização e sua correspondente representação digital. Enquanto que a noção de espaço híbrido, discutida sob os parâmetros da mídia móvel, aponta para o impacto da mobilidade na representação do espaço físico/geográfico/estático; a discussão sobre as topologias possíveis de espaços híbridos, sob os parâmetros da mineração de *big data*, não referencia, necessariamente, um espaço físico pré-existente. As topologias de espaços híbridos são constituídas como representações autorreferentes de corpos de dados; elas são (des)(re)construídas, (re)agrupadas de forma a atender demandas e contextos específicos. Pensar nestas topologias, no entanto, não implica em desassociar essas representações de seus referentes físicos, na medida em que a mineração de *big data* ocorre a partir da transcodificação digital da realidade material tangível. Mas, implica em reconhecer que essa transcodificação técnico-cultural é múltipla e excede a coleta de dados de geo-localização.

A produção de espaços híbridos, portanto, ocorre através da articulação de diversas topologias de dados, materializadas em escalas diversas - da micro escala de dados pessoais, quando, por exemplo, adotamos sensores vestíveis como o *FitBit* para monitorar a performance do nosso corpo biológico - articulada em escalas globais destes dados pessoais para análise de tendências que constituem parâmetros normativos de, por exemplo, performance atlética. Desta forma, sugerimos um *framework* de análise que demonstra essas articulações, conforme resumido na tabela 1. O objetivo desta análise é problematizar como a ação epistêmica da mineração de *big data* sustenta a construção de espaços híbridos na era da IoT. Por esses motivos, os parâmetros para esta análise partem da investigação das formas de coleta, análise e visualização de dados e como elas são performadas por agentes humanos e não humanos.

2 MMORPG significa massively multiplayer online role-playing game, jogo de simulação online com vários jogadores.

Além disso, considera a heterogeneidade das redes, a multiplicidade de dispositivos IoT. Ainda que essa análise apresente exemplos de topologias discretamente descritas, esclarecemos que elas não são autônomas ou autossuficientes uma vez que o que legitima essas topologias como arranjos sociotécnicos é a interrelação co-constitutiva entre forças sociais, materiais e tecnologias. Deste modo, a primeira topologia a ser discutida não é considerada “mais primitiva” do que as que seguem. O uso da computação ubíqua-pervasiva e de estratégias de mineração de *big data* não sobrescrevem outras formas e epistemologias espaciais.

### Topologia 1: espaços híbridos como anotação de dados

Esta primeira topologia emerge do gesto intencional da coleta de dados para uma finalidade específica e ou limitada. Nesta topologia, a coleta de dados apenas acontece quando propositadamente iniciada ou permitida pelo usuário, como por exemplo, em na divulgação de dados autorreportados. Esta dinâmica de coleta de dados está presente, por exemplo, no projeto Wiki Mapas (2010), criado pela ONG “Rede Jovem” e realizado na favela Pavão-Pavãozinho no Rio de Janeiro (DUARTE; SOUZA E SILVA, 2014). O projeto consiste no mapeamento colaborativo das ruas,

pontos de interesse, serviços e recursos disponíveis aos moradores através da inserção de dados em um aplicativo móvel que utiliza a base de dados *Google Maps*. Através da construção coletiva da representação digital da favela, a dinâmica de anotação de dados reforça o senso de comunidade. Ao inserir dados e acessar o aplicativo Wiki Mapas, o projeto dá visibilidade às localidades geográficas que eram “invisíveis”, ou seja, que não haviam sido mapeadas. A importância de projetos como este é a sustentabilidade de práticas colaborativas de criação de espaços híbridos na medida em que tecnologias móveis são apropriadas para moldar a geografia social dos lugares. As práticas vividas das tecnologias de locatividade, neste caso, incluem não só a anotação de dados específicos para cada lugar geográfico, mas também o compartilhamento de dados e criação de redes sociais a partir desses serviços.

Nesta topologia, *smartphones* e dispositivos móveis de tela são as interfaces técnicas mais comumente utilizadas para a coleta e processamentos dos dados. Ao mapear espaços geográficos e associá-los aos dados informacionais (por exemplo, quando um usuário faz *check-in* em uma localidade geográfica através do aplicativo *Foursquare*), usuários estabelecem uma correlação entre suas experiências vividas e o registro destas como dados que conferem a este espaço geográ-

**Tabela 1: Topologias de espaços híbridos de acordo com a materialidade das práticas espaciais, infraestrutura e ações epistêmicas.**

	Dados como anotações <i>Data annotation</i>	Dados como duplos do real <i>Data doubles</i>	Dados como modelos preditivos <i>Predictive models</i>
Modo de coleta e processamento de dados	Acionável	Acionável/ Automático	Automático/ Sensciente
Pervasividade	Móvel	Móvel/ Embarcado	Embarcado/ Invasivo
Interface	Visual/Tela	Visual/Tela/ Tátil	Biométrica/ sensores
Ação epistêmica	Traçar o espaço físico	Correlação de dados	Modelos preditivos de big data
Plataforma	Dispositivo externo (carregado pelo usuário)	Dispositivo vestível (atachado em corpos)	Integrante de corpos

fico, uma outra camada informacional. É nesse sentido que de Souza e Silva (2006) argumenta que usuários reinventam suas práticas espaciais como uma rearticulação de espaços híbridos, na medida em que as fronteiras entre espaços físicos e digitais são anuviadas.

## Topologia 2: espaços híbridos como duplos do real

Quando usuários e desenvolvedores de aplicativos começam a se apropriar de sensores presentes em dispositivos móveis para a coleta de dados (como por exemplo, acelerômetro e leitor de impressão digital), testemunhamos uma transição para formas automatizadas de coleta de dados. Acelerômetros em *smartphones* são utilizados, por exemplo, para contar o número de passos dados e a intensidade da atividade física, na medida em que a variação captada pelo sensor dos valores dos eixos X, Y e X - que equivalem à representação geométrica do espaço tridimensional - é transcodificada como o movimento do corpo que carrega aquele telefone.

Paralelamente, o movimento que este corpo percorre no espaço geográfico, é traçado pelo GPS ativado neste telefone. Na tela do aparelho, o usuário pode visualizar a rota traçada, sua velocidade média e uma estimativa das calorias consumidas (estimada pela correlação entre a intensidade do exercício e sua duração). Neste caso, o arranjo tecnológico formado entre o GPS, a computação em nuvem, o *smartphone*, e o corpo biológico do usuário são organizados de forma que o agenciamento para construção dessa performance espacial é distribuído. Nesta topologia, a coleta e análise automática de dados marcam como a produção de espaços híbridos nem sempre seguem parâmetros iniciados ou controlados pelo usuário. Outros dispositivos de computação vestível com interfaces hápticas, como sensores de frequência cardíaca, ou de atividade física;

como o *Nike FitBit*, por exemplo, são construídos para que possam coletar dados de forma contínua e ininterrupta, independente do acionamento do usuário. Neste caso, o *output* dos dados é analisado pelo software do aplicativo e disponibiliza ao usuário uma representação gráfica dos resultados que pode, por exemplo, demonstrar sua variação cardíaca em um recorte temporal ou correlacioná-lo a outras variáveis como qualidade do sono, tipo de atividade física realizada, queima de calorias etc.

Em um estudo de caso sobre usuários do fórum *Quantified Self*<sup>3</sup>, que reúne discussões sobre as melhores práticas para monitoramento digital da performance fisiológica e *lifelogging*, nota-se a valorização da coleta contínua e ininterrupta de dados. Os usuários do *Quantified Self* estão investidos na adoção de tecnologias digitais como ferramentas que indicam parâmetros para aperfeiçoamento fisiológico, cognitivo e emocional. Usuários, deste fórum, relatam o desejo de relacionar o maior número de variáveis possíveis para o estabelecimento de correlações entre elas, como por exemplo, a correlação entre a concentração de oxigênio no quarto e a qualidade do sono; a ingestão de cafeína e a leitura de ondas cerebrais que identificam o estado de alerta. Práticas conforme as dos usuários do *Quantified Self* demonstram como os parâmetros para tomada de decisão são delegados à inteligência algorítmica e como o ideal humano é revisto como um método quantificado de subjetivação.

Outro aspecto relativo à dimensão quantitativa de validação dos dados digitais está expresso na forma como sensores vestíveis são comercializados de forma a reforçar um ideal da representação exaustiva que circunda a aura da *big data*. O grande volume de dados próprio da *big data* reitera a ilusão da sua capacidade de fornecer uma re-

<sup>3</sup> [www.quantifiedself.com](http://www.quantifiedself.com). O slogan que define esta plataforma é "auto conhecimento através de números" (self knowledge through numbers). Acesso em 18 set. 2017.

apresentação exaustiva da realidade. No caso dos usuários do *Quantified Self*, suas práticas visam criar uma representação precisa da sua fisiologia. Através da correlação de dados e sua análise pela inteligência algorítmica, padrões são identificados e informam os usuários sobre correlações entre variáveis. Espaços híbridos, nessa topologia, são construídos na medida em que usuários criam “duplos de dados” de si mesmos. Eles se constituem na interrelação entre os corpos biológicos, o sempre-crescente banco de dados gerado por sensores vestíveis e as visualizações gráficas geradas pelos algoritmos, que posicionam a performance destes corpos diante de outros corpos (quando, por exemplo, os dados coletados pelo *Nike FitBit* não só indicam a sua performance em uma corrida, mas estabelece um comparativo da sua performance com as de outros usuários).

É preciso esclarecer que a construção de espaços híbridos através do mapeamento/coleta automática de dados não se restringe à coleta de dados pessoais da fisiologia e pode ser estendida às escalas maiores, em nível de populações. Múltiplas fontes de dados pessoais, como fisiologia, geolocalização, fluxo de trânsito, etc., são automaticamente coletadas, correlacionadas com outros bancos de dados e terceirizadas para a geração de visualização de dados. O projeto *LIVE Singapore*<sup>4</sup>, desenvolvido por Carlo Ratti no *MIT Senseable Media Lab*, disponibiliza uma plataforma online para apresentação de gráficos gerados com dados em tempo real que informam sobre os usos dos espaços públicos da cidade (Ratti, Claudel, KloECKI, 2014). O objetivo desse projeto é informar aos gestores e aos cidadãos a melhor tomada de decisão e oferecer novos *insights* para planejamento urbano e serviços. As fontes de dados para essas visualizações são variadas: temperatura, índice pluviométrico, consome de energia, fluxo de trânsito, uso de telefonia celular, etc.

4 <http://senseable.mit.edu/livesingapore/>. Acesso ago. 2016.

O projeto *LIVE Singapore!* se apresenta como “um ecossistema e uma caixa de ferramentas de dados coletados e exibidos em tempo real que descrevem a dinâmica urbana”<sup>5</sup>. Nesse exemplo, espaços híbridos são compartilhados entre e construídos por usuários que acessam o mapa e outros os quais navegam o espaço físico da cidade, que por contribuírem para a geração de dados afetam a representação visual de seu mapa. A partir do que a visualização de dados aponta sobre o fluxo do trânsito, pessoas na cidade tomam decisões a respeito de quais rotas tomarem e quais evitarem. A criação de espaços híbridos como ferramenta para tomada de decisão marca a transição para a terceira topologia. Nesta terceira categoria, a coleta automática de dados para fins específicos (coleta da frequência cardíaca para monitoramento da saúde), é expandida para a coleta sentiente, ininterrupta, de todas as fontes de dados possíveis. Mesmo quando o propósito da coleta de dados não está claro, a correlação e análise de dados apontam para a relevância da construção de espaços híbridos como modelos conceituais preditivos.

### Topologia 3: espaços híbridos como modelos conceituais preditivos

Na medida em que redes de sensores e a dinâmica da *big data* estão largamente distribuídas por espaços físicos e corpos, a mineração de dados se torna mais sentiente e a análise desses dados passa por processos mais sofisticados de correlação. Nessa topologia, uma maior variedade de sensores e maior heterogeneidade de tipos de rede de comunicação produzem o que Roger Clarke chama de uma vigilância de dados em massa (*mass data-veillance*) (Clarke, 2003), através da agregação de dados de diferentes fontes: *cookies* de navegadores de internet, etiquetas RFID em passaportes e

5 “ecosystem and a toolbox for real-time data that describe urban dynamics.” Tradução livre.

carteiras de identidade, chips biotecnológicos implantáveis, etc. Traços digitais são coletados não só para descrever atividades no espaço físico em tempo real (como descrito pela topologia anterior), mas para criar modelos preditivos baseado em padrões identificados historicamente. A discussão trazida por Mark Andrejevic (2014) e Jeremy Parker (2013) sobre o regime de vigilância por drones exemplifica como a mineração de *big data* gera ação preemptiva por governos:

Drones são a frente de batalha tanto para observação quanto para resposta. Eles não só coletam dados mas estão, cada vez mais, recebendo a função de analisar os dados. Por fim, a cadeia de comando que partia da decisão de quem é o inimigo (a decisão política), passava pela localização do inimigo (o processo de observação), até a localização do inimigo (o processo da infantaria) está se tornando um procedimento singularmente determinado por parâmetros digitais. Não é só dizer que drones conseguem localizar o inimigo com mais eficiência; mais do que isso, eles podem coletar, processar quaisquer dados necessários para determinar algoritmicamente o que é uma ameaça em qualquer situação/sujeito e agir de acordo com seu julgamento.(PACKER, 2013, p.299)<sup>6</sup>

Nesta topologia, a mineração e análise de *big data* são levadas ao limite da representação da realidade física. Espaços híbridos são acionados através de dados coletados por drones e não visam ape-

nas representar a realidade física presente. Baseado no acúmulo de dados coletados e na identificação de tendências neste corpus, espaços híbridos são construídos para representar realidades futuras plausíveis e informar a tomada de decisão que nos levará para a realidade que queremos criar.

Outro exemplo de aplicação de mapeamento a partir de mineração de dados é o mapeamento genético. A primeira sequência genética foi mapeada em 2003. Doze anos depois, cerca de três bilhões de pares de genes foram identificados (MAYER – SCHONBERGER, CUKIER, 2013O.). A correlação entre os dados do genoma permitiu a associação de síndromes a genes específicos e a entender como o pareamento de determinados genes levam à produção de determinados traços físicos. A partir da história familiar e a análise do genoma pessoal é possível, por exemplo, calcular a probabilidade de manifestação de uma doença transmitida geneticamente, como o câncer. A análise preditiva é um ponto de destaque nessa topologia, na medida em que a correlação de dados pode ser condensada em informação estatística e, por exemplo, contribuir para a decisão de se fazer um tratamento médico preventivo. Apesar desta topologia ser constituída a partir do acúmulo de dados históricos, do passado e do presente, ela é orientada ao futuro. Portanto, os espaços híbridos nessa topologia constituem-se como modelos algorítmicos da realidade física.

A mineração de dados também impacta a construção de espaços híbridos em espaços urbanos. Novos modelos de cidades inteligentes são implementados como um banco de ensaio sempre em andamento, onde a mineração de dados no presente determina a reorganização do espaço em um futuro em constante atualização. Orit Halpern (2013) descreve a implementação de Songdo, um projeto de cidade inteligente na Coreia do Sul elaborado pela Cisco onde toda a infraestrutura da cidade atua como sensor de coleta de dados do ambiente que ajustam-se automaticamente às necessidades da sua população. O projeto da cidade prevê um am-

6 Drones are the experimental forefront of both observation and response. They not only collect the data but are increasingly being given the task of processing the data. Finally, the chain of command that led from deciding who was the enemy (the political decision), to locating the enemy (the observational process), to executing the enemy (the soldiering process) is becoming a single digitallydetermined procedure. It is not simply that drones can locate real pre-existent enemies more accurately; rather they can collect and process the necessary data to determine algorithmically the threat potential of any giver situation/subject and act accordingly. Tradução livre.

biente urbano embarcado por sensores capazes de coletar e mapear *inputs* de dados de humanos e não humanos; de movimentos dos olhos a movimentos do trânsito, toda a gestão do espaço é planejada de forma a se adequar instantaneamente para seu melhor funcionamento: da entrega de publicidade customizada ao usuário, à reprogramação dos semáforos de trânsito. Neste sentido, Songdo é um modelo de espaço urbano que é sempre provisional, está sempre em versão beta. É um espaço híbrido experimental, um espaço físico moldado por inteligência algorítmica que está em constante reformulação.

Os avanços em mapeamento genéticos, implantes biotecnológicos de sensores e chips também visam transformar os corpos biológicos em espaços híbridos mutáveis, manipuláveis, plataformas de metadados que possibilitam atualizações eternas de versões betas de nós mesmos. Esses avanços demonstram como os arranjos tecnológicos entre sensores e técnicas de mineração de dados marcam uma virada para uma epistemologia espacial que não está investida apenas na criação de representações que descrevem espaços físicos em sua atualidade concreta. Tanto no caso das cidades inteligentes quanto no mapeamento fisiológico dos nossos corpos, e da ação dos drones, o esforço representacional está voltado para a construção de modelos espaciais preditivos baseados em correlação de dados. Enquanto que nas topologias anteriores os espaços híbridos estão atrelados a seus referentes físicos, nesta topologia eles tendem à autorreferência, construídos como territórios de dados que emergem da performance algorítmica.

## Considerações finais

Argumentamos que a ação epistêmica digital que anima a mineração de dados pode ser apropriada como ferramenta analítica para análise das práticas de representação espacial na era da IoT. Estas topologias articulam-se na medida em que a mineração de dados assume maior automação

dos processos de coleta, processamento e análise dos dados, e na medida em que as redes de comunicação sem fio tornam-se mais heterogêneas.

A primeira topologia descrita é caracterizada pela coleta de dados autor-reportados. Como no exemplo do Wiki Mapa, a coleta de dados acontece com finalidade específica e, geralmente, pré-conhecida. A segunda topologia é marcada pela crescente pervasividade e automação da mineração de dados. O uso de dispositivos de computação vestível, que discretamente coletam dados sobre a performance física do corpo, são exemplos de usos de interface háptica que não dependem de dispositivos de tela, como tende a ser o caso da primeira topologia. Espaços híbridos são constituídos na interrelação entre os espaços físicos dos corpos/geográficos e a visualização de dados gerados pela correlação algorítmica.

Finalmente, a terceira topologia leva a representação espacial ao limite da autorreferência. O processamento de dados remessa à construção de modelos espaciais preditivos. O objetivo desta topologia é a construção de espaços híbridos sempre sentientes às realidades físicas, de modo que estejam em contínuo melhoramento de seus próprios modelos conceituais. Em todas as topologias, a mineração de dados ocorre em níveis multi-escalares e materializam formas de saber algorítmicas através da ação epistêmica característica da materialidade digital.

## Referências

ANDREJEVIC, Mark. Big Data, Big Questions| The Big Data Divide. *International Journal of Communication*, 8, 2014, p. 1673-1689.

ANDREJEVIC, M. & GATES, K. Big Data Surveillance: Introduction. *Surveillance & Society* 12.2, 2014, p.185-196.

ASHTON, Kevin. That "Internet of Things" thing. *RFID Journal*. Disponível em <<http://www.rfid->

- journal.com/articles/view?4986, 22 de junho de 2009>. Acessado em 31 out. 2016.
- BOYD, Danah e CRAWFORD, Kate. Critical questions for big data. *Information, Communication and Society* 15 (5), 2012, p. 662-679.
- CLARKE, Roger. *Dataveillance, 15 years on*. Disponível em <<http://www.rogerclarke.com/DV/DVNZ03.html>>, 31 de março de 2003>. Acessado em 15 maio 2015.
- CRESWELL, Timothy. Introduction: Theorizing place. In G. Vearstraete and T. Cresswell (eds.) *Mobilizing Place, Placing Mobility: The Politics of Representation in a Globalized World*. Amsterdam: Rodopi, 2002, p.11-32.
- DOURISH, Paul. *Place and space: Ten years on*. Paper presented at the Anniversary conference on Computer Supported Cooperative Work, Alberta, Canada, 2006.
- DUARTE, Fernanda, & SOUZA E SILVA, Adriana de. Arte, mov, Mobilefest and the emergence of a mobile culture in Brazil In G. Goggin, & Hjorth, L. (Eds.). *The Routledge companion to mobile media*. New York: Routledge, 2014, p. 206-215.
- HALPERN, Orit. et al. Test-bed Urbanism. *Public Culture* 25:2, 2013, p 272-306.
- KITCHIN, R. Big data, new epistemologies and paradigm shifts. *Big Data & Society*, april-june, 2014, p. 1-12.
- LATOUR, Bruno. Morality and Technology. The end of means. *Theory, Culture & Society*, n. 5-6, 2002, p. 247-260, 19 vol.
- LEMOS, André. *Mídia locativa e territórios informacionais*, 2007. Disponível em: <[http://www.facom.ufba.br/ciberpesquisa/andrememos/midia\\_locativa.pdf](http://www.facom.ufba.br/ciberpesquisa/andrememos/midia_locativa.pdf)>. Acesso em 28 jun. 2017.
- LEVY, S. *Hackers: Heroes of the Computer Revolution*. New York: Penguin Books, 2001.
- MAXIGAS. Hacklabs and hackerspaces. *Journal of Peer Production*. Available at: <<http://peerproduction.net/issues/issue-2/peer-reviewed-papers/hacklabs-andhackerspaces/>>. Acesso em: jan. 2015.
- MASSEY, D. Power-geometry and a progressive sense of place. In: Bird, J, Curtis, B, Putnam, T, Robertson, G and Tickner, L eds. *Mapping the Futures: Local cultures, global change*. New York: Routledge, 1992, p. 59-69.
- MAYER-SCHONBERGER, V.; & CUKIER, K. N. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston, MA, USA: Eamon Dolan/Houghton Mifflin Harcourt, 2013.
- PACKER, Jeremy. Epistemology Not Ideology OR Why We Need New Germans. *Communication and Critical/Cultural Studies*, n. 2-3, p. 295-300, June-September, 2013, 10 vol.
- RATTI, Carlo; CLAUDEL, Matthew; LIVE Singapore The Urban Data Collider. *Transfers*. n 3, Asia Issue 2: Special Section on Rethinking Mobility History Asia, p. 117-121, 2014, 4 vol.
- SOUZA E SILVA, Adriana de. From cyber to hybrid: Mobile technologies as interfaces of hybrid spaces. *Space and Culture*, n. 3, p. 261-278, 2006.
- WARF, B., & ARIAS, S. (eds.). *Spatial Turn: Interdisciplinary Perspectives*. Florence, KY, USA: Routledge, 2008
- WEISER, Mark. The Computer of the 21st Century. *Scientific American*, 265, 3, 1991, p.66-75.
- WEISER, M., GOLD, R., & BROWN, J. S. The origins of Ubiquitous computing research at PARC in the late 1980's. *IBM Systems Journal*, 38 (4), 1999, p. 693-696.

**Débora Franco Machado** – Mestranda em Ciências Humanas e Sociais na UFABC e pesquisadora no Laboratório de Tecnologias Livres da UFABC. **E-mail:** [deborafmachado@gmail.com](mailto:deborafmachado@gmail.com)

Recebido: 16 mar. 2018

Aprovado: 11 abr. 2018

# O uso de algoritmos na mídia programática

## Use of algorithmic in the programmatic media

**Stefanie Carlan da Silveira**

Universidade Federal de Santa Catarina  
Florianópolis, SC, Brasil

**João Gabriel Danesi Morisso**

Universidade de São Paulo  
São Paulo, SP, Brasil

## Resumo

A publicidade digital encontra espaços de operação na internet desde o seu surgimento. Inicialmente com banners e formatos que remetiam à mídia tradicional, os processos evoluíram ao longo do tempo em busca de audiências segmentadas, monitoramentos mais efetivos e retornos sobre investimento mais metrificados. Atualmente, a mídia programática é responsável pela automatização dos investimentos publicitários em diversos sites da rede. Apesar de útil, a ferramenta traz conflitos e prejuízos tanto para os consumidores quanto para os investidores, promovendo faturamento para grupos responsáveis por extremismo ou pirataria. Este artigo promove uma discussão aberta sobre esse cenário e propõe um debate acerca da automatização da publicidade digital, com exemplos e reflexões teórico-empíricas.

## Abstract

Digital advertising has found operating spaces on the internet since its beginning. Initially with banners and formats that refer to traditional media, the processes have evolved over time in search of segmented audiences, more effective monitoring, and more metrified returns on investment. Currently, the programmatic media is responsible for automating the advertising investments in several sites of the network. Although useful, the tool brings conflicts and damages to both consumers and investors, promoting billing for groups responsible for extremism or piracy. This article promotes an open discussion about this scenario and proposes a debate about the automation of digital advertising, with examples and theoretical-empirical reflections.

## Palavras-chave

Publicidade digital. Mídia programática. Automatização. Algoritmos.

## Keywords

Digital advertising. Programmatic media. Automation. Algorithms.

## Introdução

A publicidade e a veiculação de mídia no ambiente digital estão em constante desenvolvimento e cada vez mais o ecossistema de empresas que atuam neste segmento aumenta na mesma velocidade. Naquilo que costumamos chamar de *web 1.0*, cada *website* era responsável por comercializar e publicar seus próprios anúncios em *banners* digitais. Este movimento se dava em função de uma transposição da lógica dos modelos de publicidade atrelados à mídia impressa ou tradicional. Ou seja, acreditava-se que expor a marca num formato que se assemelha a um “*outdoor* digital” seria suficiente e eficaz para a conquista da atenção dos consumidores.

A *web* – e seus resultados sobre os processos produtivos e criativos –, se configurou como uma das principais transformações da atualidade. Na publicidade ela vai chegar entre o fim dos anos 1990 e o começo dos anos 2000, ainda sofrendo todas suas limitações (CARVALHO et al, 2015, p. 8). A partir dos anos 2000, a transformação se acentuou com o desenvolvimento da tecnologia e dos processos interativos, especialmente os potencializados pelas plataformas de redes sociais online. Esse desenvolvimento gerou novas ferramentas e novos espaços de mídia, o que dá início à pulverização da verba publicitária antes focada em espaços que mimetizavam formatos anteriores.

Os modelos iniciais de publicidade na *web* foram se mostrando pouco eficazes e prejudicando sites que se tornavam visualmente poluídos, com anúncios que nem sempre correspondiam aos interesses dos consumidores ou traziam retorno efetivo sobre investimento para os clientes. A evolução da publicidade digital acabou trazendo novos formatos que eram capazes de superar esses problemas oferecendo audiência segmentada e opções mais personalizadas e eficientes. Alguns exemplos são: anúncios em redes sociais,

*links* patrocinados e formatos *mobile*. Com o desenvolvimento tecnológico dos mecanismos de troca de informação entre *websites* e empresas é possível hoje trabalhar com *targets* (alvos) específicos de audiência para que anunciantes entreguem suas mensagens direcionadas não apenas pelas clássicas divisões demográficas, mas também se utilizando do comportamento do usuário e informações em tempo real, como a temperatura da cidade onde o usuário está e qual hora do dia ele está navegando pela rede.

Este formato mais recente utilizado para definição dos elementos que serão exibidos ao consumidor é chamado de mídia programática. Além de permitir a utilização de diferentes e novos critérios na hora de estabelecer qual anúncio é mais relevante, esse tipo de ferramenta também permite a exibição dos materiais em uma quantidade maior de sites e agiliza o processo, facilitando o trabalho de publicitários e profissionais de mídia. “A propaganda online é gerenciada automaticamente: alguns softwares/plataformas navegam no inventário de espaços publicitários online e otimizam o planejamento, a efetividade e o investimento” (TESTORI, 2014, p. 11).<sup>1</sup>

Apesar dos benefícios, este sistema não surge isento de problemas e questões passíveis de análise. Em função da automatização dos processos, alguns anúncios acabam sendo exibidos em sites ou redes que apresentam conteúdos que não estão de acordo com a empresa que está anunciando ou ainda, representam grupos extremistas, de pirataria ou material pornográfico. Este aspecto da automatização da publicidade é o que pretendemos discutir neste artigo.

Como objetivo central, buscamos trazer aqui uma discussão aberta sobre o *status* atual da publicidade digital e os problemas que a mediação

<sup>1</sup> Trecho original: “Web advertising space is managed automatically; some software/platforms surf the inventory of web adv spaces to optimize the media planning/effectiveness and investment.” Tradução dos autores.

majoritariamente técnica pode acarretar neste caso. Para isso, inicialmente, explicamos do que se trata e como funciona a mídia programática. Em seguida, destrinchamos um pouco mais alguns aspectos relacionados à automatização dos processos e o uso de algoritmos. Por último, trazemos exemplos de casos recentes enfrentados por empresas e sites que acabaram demandando a atenção e até mesmo a ação de profissionais com relação a eles. Assim, esperamos contribuir para este debate buscando um olhar mais atento às ações e soluções propostas e implementadas pelo mercado.

## Mídia programática

É no cruzamento de dados sobre o usuário para a definição de quais anúncios serão exibidos e onde eles serão apresentados que encontramos a primeira base pela qual funcionam os anúncios segmentados via mídia programática. De uma forma geral, é possível definir mídia programática como a compra e venda de espaços publicitários online de maneira automatizada e dentro disso, existem diversas maneiras de efetuar essa transação.

Busch (2014) esclarece que há diversos termos sendo utilizados conjuntamente à “mídia programática”, entre eles display de dados, compra programática e comércio automatizado, no entanto, todos eles se referem ao uso massivo de dados, tecnologia e inteligência artificial no marketing com o objetivo de melhorar sua eficiência em tempo real. “A publicidade programática descreve a veiculação automatizada de anúncios digitais em tempo real com base em oportunidades de impressões de anúncios individuais” (BARDOWICKS E BUSH, 2013, *apud* BUSCH, 2014, p. 8).

Segundo Rosa (2016), o termo se opõe ao que conhecíamos tradicionalmente no mercado publicitário como venda direta, onde o departamento de mídia da agência era responsável pela

definição de todas as etapas da venda, envolvendo o que era escolhido pelo anunciante, à divisão do investimento, os formatos, onde eles seriam inseridos e como apareceriam para o usuário, além da mensuração dos resultados posteriores. Ainda de acordo com a autora, as plataformas e as empresas de análise que hoje atuam no mercado podem ser divididos em três níveis: os que assessoram anunciantes, os que atuam junto aos veículos e os que centralizam o processo, operando com os dois lados de forma conjunta.

Esse processo é possível graças a uma interligação de plataformas e de encadeamento de fluxos de informação com algoritmos de análises e classificação complexos que são executados em frações de milésimos de segundo. Junto à base tecnológica há também uma série de predefinições que são estabelecidas por anunciantes e veículos, as quais servem de parâmetro para as interações entre as plataformas que fazem parte da programática. Cada plataforma torna-se um agente no mercado de compra e venda de espaços publicitários, que podem ser separados por sua função: os que atuam assessorando anunciantes, os que atuam junto aos veículos e os que atuam mediando o processo de transação, funcionando para ambos os lados.

Os agentes que operam junto aos anunciantes (*Demand Side*) os auxiliam a definir os *targets* que derivam das estratégias do planejamento para que a compra de mídia seja a mais adequada possível. O trabalho é feito levando em consideração tanto os dados psico-demográficos dos usuários quanto às taxas de conversão necessárias para obter bons resultados nas suas campanhas. Além disso, esses agentes também atuam na definição do anunciante sobre o quanto ele está disposto a investir em cada exibição dos anúncios de sua campanha.

Junto aos veículos (*Supply Side*), estão os agentes que monitoram as audiências dos portais e coletam informações dos visitantes a fim de en-

riquecer a decisão de compra de espaço nestes locais, que se baseia em uma análise de adequação ao *target*. Uma vez mapeado a audiência do inventário, ele os classifica por preços mínimos de exibição em cada veículo.

Por fim, os agentes que atuam como mediadores (*Ad Exchange*) do processo de transação. Eles operam como uma plataforma na qual, de um lado, os anunciantes delimitam para quem e por quanto tempo a campanha será visível, além de qual investimento máximo será disponibilizado para aquele momento. Do outro lado desta plataforma, os veículos definem quais espaços estão disponíveis e qual o investimento mínimo que deve ser feito para que eles possam ser comprados do inventário.

O processo então se repete para cada posição de mídia que deve ser preenchida em cada página visualizada na internet – dentro, claro, dos veículos que optam por estar inseridos nos inventários de mídia programática. Por definição, o IAB explica a mídia programática como o “uso de sistemas automatizados e processos de compra e venda de inventário em negociações por meio de leilões em tempo real” (IAB BRASIL, 2015, p. 1) sem levar em consideração outras situações possíveis.

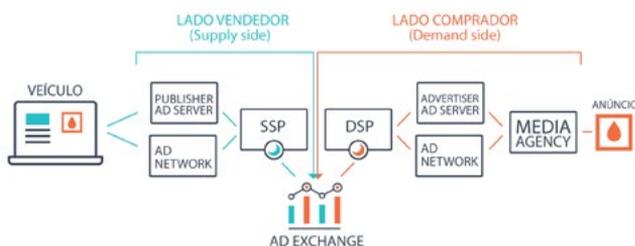


Figura 1 - Funcionamento da mídia programática.

Fonte: Autores do trabalho

Segundo Rosa (2016), alguns dos formatos mais utilizados pela mídia programática são: exibição de mídia online em formatos *display*<sup>2</sup>, anúncios em redes

<sup>2</sup> Formato Display é o formato de exibição padronizado pelo número de pixels exato, sem extrapolar a dimensão designada.

sociais<sup>3</sup>, *links* patrocinados<sup>4</sup>, vídeos *pre-roll*<sup>5</sup> e formatos *mobile*<sup>6</sup>. Segundo dados do Marketer (2016c), inclusive, “mais de dois terços do total das mídias display já é comercializado através de plataformas programáticas” (*idem*, p. 41). Existem diversas modalidades desse processo como reserva de espaços, lances pré-pagos e espaços *premium*. Mesmo com uma amplitude de modalidades de compra e venda que existem na mídia programática, vamos abordar o RTB (*Real-Time-Bidding*), leilão em tempo real de espaços disponíveis no *ad server* do SSP (*Supply Side Platform* - lado vendedor) que acontece no *Ad Exchange* baseado em informações detalhadas sobre as audiências.

A operação do RTB se faz mais clara quando utilizamos um exemplo: dois usuários visitam o mesmo *site*, porém um deles é um senhor aposentado que tem um histórico de visitas e comportamento relacionado a turismo internacional e cruzeiro, já o segundo usuário é uma jovem que foi identificada como praticante ou simpatizante de assuntos como *kickboxing* e pilates. Neste cenário, o RTB promete que o leilão em tempo real do espaço publicitário disponível neste *website* seja diferente para cada usuário e cada um deles possa ser exposto a um anúncio que esteja mais relacionado com seus interesses pessoais. As informações de cada um dos visitantes são recebidas no servidor de *Ad Exchange* junto aos critérios do espaço disponível no SSP (lado vendedor) para que os lances sejam direcionados e o lance vencedor no DSP (*Demand Side Platform* - lado comprador) entregue o anúncio correto para o *Ad Server*

<sup>3</sup> Anúncios em rede social podem possuir formatos diversificados que são relacionados à plataforma em que são inseridos, podendo ser posts patrocinados, vídeos, links e formatos interativos.

<sup>4</sup> São utilizados principalmente por sistemas de busca e priorizam resultados pagos colocando-os acima dos resultados orgânicos.

<sup>5</sup> São chamados *pre-roll* os vídeos que antecedem o conteúdo que o usuário quer assistir, uma espécie de trailer que pode ter algumas regras como: pular depois de cinco segundos assistidos ou obrigar o usuário a assistir de maneira integral.

<sup>6</sup> São formatos que utilizam as potencialidades do dispositivo móvel, como tela sensível ao toque e sensores próprios do aparelho.

do SSP, que por sua vez publica anúncios diferentes para cada visitante.

De acordo com a IAB (*Interactive Advertising Bureau*) (2015b), o caminho entre o usuário carregar uma página que tenha em seu código uma solicitação de um anúncio para uma *Ad network* ou *Ad Exchange* e o anúncio aparecer na página já carregada deve ser de até 10 milissegundos. Nesse sistema quase que instantâneo, dados são a base que solidifica as decisões que são tomadas por máquinas e dispositivos automatizados munidos de regras e critérios para que as operações aconteçam.

O marketing digital e a publicidade programática, sendo novos meios de comunicação são grandes forças que estão trabalhando no conflito e consequentemente utilizando indiretamente informações que são de cunho privado. Mesmo que essas informações ou rastros que são deixados pelos usuários não sejam compartilhados, organizações, através dos portadores desses dados podem ter vantagem para suas campanhas e anúncios. Google e Facebook, por exemplo são empresas portadoras dessa rede de banco de dados. A partir da solicitação das organizações, em forma de leilão (a empresa que pagar mais tem seu anúncio com mais intensidade direcionado para o consumidor), os anúncios são encaminhados para os usuários a partir do perfil de cliente que as mesmas organizações traçaram e a partir dos rastros e informações que são pesquisadas pelos usuários, esses clientes são selecionados e as campanhas e anúncios encaminhados para que cheguem aos consumidores (ZANETTI, 2017, p. 62).

Segundo Busch (2014), é possível elencar pelo menos cinco características da mídia programática. A primeira é a granularidade, que diz respeito à possibilidade de se levar em consideração as oportunidades de impressões de forma individualizada, em plataformas e ambientes específicos de publicidade. A segunda é a troca em tempo real, que indica a possibilidade de decisão sobre uma impressão no exato momento em que

o seu espaço é disponibilizado. A terceira é a informação em tempo real, sobre o acesso imediato à coleta de dados sobre a transação. A quarta é a criação em tempo real, indicando o posicionamento dos anúncios imediatamente após o vencimento do leilão. A última característica é a automação que diz respeito à reserva e a publicação automatizada.

Vale lembrar que todos estes processos somente são possíveis em função da quantidade de dados disponíveis que usuários deixam registrados durante sua navegação. Esses dados nos remetem ao que atualmente chamamos de *Big Data*. O desenvolvimento tecnológico fez com que muitos dados e informações fossem gerados numa velocidade e volume muito maiores do que ocorria antes do surgimento dos computadores. O nascimento de máquinas pessoais que operam também como sensores, como os *smartphones*, por exemplo, fez com que se pudessem armazenar mais dados e contabilizá-los mais rapidamente. Com a Internet, a quantidade de informações armazenadas e também em circulação tomou proporções ainda maiores. Segundo Tascón e Coullaut (2016, p. 51), estas "bases de dados gigantes às quais podemos acessar e trabalhar com" confirmam o chamado *Big Data*. Na publicidade, a imensidão do *Big Data* garante que plataformas que capturam informações dos usuários a todo o momento possam negociar essas informações para anunciantes, que por sua vez, buscam a todo instante atingir de forma mais assertiva os consumidores que interessam ao seu negócio. "Analisando o consumo de informação e como este se engendra, se podem realizar modelos preditivos que melhorem os produtos e a experiência dos leitores" (TASCÓN; COULLAUT, 2016, p. 51-52)<sup>7</sup>.

## Algoritmos e regras automatizadas

As regras e critérios, a partir dos quais o sistema

<sup>7</sup> Tradução dos autores, trecho original: "Analizando el consumo de información y cómo este se produce, se pueden realizar modelos predictivos que mejoren los productos y la experiencia de los lectores".

de mídia programática opera, funcionam através dos algoritmos. Sendo assim, cabe-nos aqui explicar um pouco mais sobre este conceito e seu funcionamento. A definição clássica indica que um algoritmo pode ser definido como um conjunto de regras finitas organizadas com o objetivo de resolver um problema específico ou executar alguma tarefa. No caso da tecnologia digital, dos computadores e da internet, os algoritmos são executados por máquinas, pois este conjunto de regras que os define foi escrito utilizando alguma linguagem de programação (CORMEN, 2014).

Além de serem escritos a partir de linguagens de programação, os algoritmos mais evoluídos também estão baseados no que chamamos de *machine learning*. Silveira (2017) explica que, no caso de um *software*, o programador é responsável por descrever no código todas as operações que serão executadas pelo algoritmo, mas no *machine learning*, as funções que serão executadas são definidas pelo próprio algoritmo. Nesse caso, como explica a autora, os algoritmos

têm capacidade de aprendizado à medida que novos *inputs* vão sendo colocados no sistema. Assim, conforme os gostos e interações vão mudando, o algoritmo vai aprendendo e se adaptando para sempre exibir os conteúdos que consideram mais adequados para o usuário" (SILVEIRA, 2017, p. 174).

Dentre os algoritmos presentes nas interações entre as plataformas de mídia programática estão os de classificação. Segundo Cardoso (2012), esses algoritmos são utilizados para catalogar, ou seja, têm, por exemplo, a capacidade de construir perfis de audiência baseados em dados coletados em múltiplas plataformas.

O objetivo dos algoritmos de classificação é prever o valor de uma determinada variável (dado o alvo ou objetivo). Se a variável-alvo for categórica, é um problema de classificação e, se for numérica, é um problema de regressão. Para cada registro no conjunto de dados, é determinado o valor da clas-

se atributo. Assim, é construído o modelo baseado num conjunto de treino, sendo usado para prever novos dados. Estes algoritmos são construídos com um processo de duas fases. A primeira fase é a construção do modelo, em que, para cada amostra (ou registro, ou objeto, ou instância, ou exemplo), se assume que tem um valor para o atributo-alvo. O conjunto de amostras usado para a construção do modelo designa-se de conjunto de treino (CARDOSO, 2012, p. 5-6).

De forma menos técnica, o conjunto de dados que treina o algoritmo é considerado um modelo de categoria que é aplicado em novos e diferentes dados. A partir desses novos *inputs*, o algoritmo trabalha por semelhança e as informações recebem o valor atribuído (classe atributo) tornando possível assim a classificação de novos dados com base nos modelos determinados pelo conjunto inicial.

Os dados passíveis de classificação de audiência em grupos determinados pelos conjuntos modelo variam de: informações comportamentais do usuário; seu histórico de navegação, informações geográficas e de localização, demográficas; como sexo, idade, rendimento, informações temporais; hora do dia, clima e temperatura, informações tecnográficas; como dispositivo, tamanho, sistema operacional. Outros dados utilizados para modelar as regras dos anunciantes são: datas de início e fim das exposições, número máximo de impressões, dias da semana e horários, intervalos de tempo, geolocalização e clima (DOMINGOS, 2017).

Uma vez que os agrupamentos sejam feitos e construídos, os conjuntos de treino que servem de modelo para categorização de novos dados, as regras e critérios estão delineadas para que os algoritmos sejam executados em modo não supervisionado, ou seja, automatizados sem intervenção humana. Nesse caso, o agente automatizado perfeito supostamente eliminaria grande parte burocracia na compra e venda de anúncios, parecendo improvável que as empresas pudessem criar agentes que prejudicassem a base de seus

negócios. No entanto, a automatização e à falta de intervenção humana podem fazer com que os locais disponíveis para exibição desses anúncios nem sempre sejam averiguados e qualificados de acordo com seu conteúdo e não somente a audiência que é capaz de gerar.

## Problemas da automatização

A automatização é um processo que como o próprio nome diz, acontece sem intervenção humana e, portanto é passível de erros, comportamentos inadequados, padrões discordantes, aberrações, contaminantes, anomalias e uma série de defeitos na operação como um todo. Dentre os erros possíveis, este artigo limita-se a discutir apenas os que se relacionam com os exemplos do objeto de estudo, a publicidade digital, mais especificamente, a mídia programática.

Observando o fluxo entre as plataformas agentes da mídia programática os *Ad Servers* de ambos os lados (DSP e SSP) conversam com o *Ad Exchange* e com as *Ad Networks* de maneira automatizada a partir de comandos, regras e códigos definidos com pouca transparência. Os estrategistas digitais e responsáveis por marketing e propaganda online normalmente não possuem uma visão clara sobre como um anúncio vai parar em determinado local. Existem análises de *tags*<sup>8</sup> e *blacklists*<sup>9</sup> nos servidores de *Ad Performance* para evitar problemas e fraudes, mas estas nem sempre são efetivas. A automatização traz consigo muitos benefícios, como anúncios direcionados por audiência, inventário multiplataforma entre dispositivos, visibilidade em tempo real, segmentação um-a-um, custo por performance, entre outros.

8 Dentro das linhas de código de programação presentes nos sites existem metadados, que são dados visíveis apenas para as outras máquinas e que categorizam algumas intenções e operações específicas, são chamadas tags de código.

9 São listas com sites maliciosos, também chamada de lista negra, onde são mapeados e expostos inventários que devem ser recusados.

Apesar dos benefícios, o sistema deixa brechas para que ocorram padrões discordantes que não se encaixam no esperado, agindo de maneira errática com valores e ações que fogem da considerada normalidade. Estes erros ou padrões discordantes também são denominados anomalias, que podem acontecer acidentalmente ou de maneira intencional, como no caso de fraudes.

Fraude ainda é um problema. Toda indústria conversa sobre a luta contra fraudes mas isso não parece ter eliminado o problema. O ecossistema da programática é especialmente suscetível a práticas fraudulentas como a chamada "*url masking*" onde um veículo lista o seu *website* no *Ad Exchange* como se fosse outro, normalmente com melhor reputação. O anunciante não faz ideia onde aquele anúncio está sendo veiculado (KANTROWITZ, 2015 s/p.).<sup>10</sup>

Lopes (2012) categoriza as anomalias em três tipos: pontuais, contextuais e coletivas. Esta divisão permite que seja possível nos aprofundarmos nas características responsáveis pelos comportamentos erráticos dos códigos. Quando uma operação é discordante de todo resto do sistema ela deve ser considerada uma anomalia pontual. Quando os dados entre si são discordantes do resto, a anomalia é coletiva. Porém o tipo de anomalia, que encontramos em nosso objeto de estudo, é do tipo contextual:

Se um dado objeto for anômalo relativamente ao seu contexto e não relativamente a todo o universo de objetos, este é denominado uma anomalia contextual. Os problemas que atentam a este tipo de anomalias requerem a presença, no seu

10 Escrito por Ale Kantrowitz para a revista Ad Age. Trecho original: "Fraud is still a problem. All the industry talk about fighting fraud doesn't seem to have eliminated it. The programmatic ecosystem is especially susceptible to a fraudulent practice called "URL masking," where a publisher lists its website in the ad exchanges as another, usually more reputable, website entirely. The buyer has no idea where the ads are actually running." Acesso em :16 jan. 2018. Tradução dos autores.

conjunto de dados, de atributos capazes de introduzir a noção de contexto, como por exemplo tempo, posição geográfica ou temperatura. Desta forma, os métodos de detecção tentam identificar anomalias recorrendo a atributos denominados comportamentais tendo em consideração os de contexto (LOPES, 2012, p. 20).

É possível perceber anomalias contextuais levando em consideração a semântica nesse engendramento de fatores. Uma palavra como “bomba” pode estar inserida em um contexto terrorista e violento ou relacionada a uma *éclair* também conhecida como bomba de chocolate. O contexto faz toda diferença nessa interpretação, porém nem sempre os algoritmos conseguem ter a mesma clareza que um humano teria ao tomar decisões na compra de espaços e publicação de anúncios online.

Um dos exemplos que podemos citar aconteceu em uma notícia sobre um derramamento de petróleo no Golfo do México, publicada no portal *Marketwatch*. Nesse caso, a automação de mídia inseriu um anúncio da Deloitte<sup>11</sup>, uma consultoria internacional que no título do anúncio relaciona seus serviços com o sucesso de uma extração de petróleo com alta *performance*. Neste caso, a palavra-chave petróleo tinha dois contextos, um negativo relacionado ao derramamento e outro positivo relacionado à alta *performance* de uma extração, o que provocou uma má inserção devido ao conflito contextual.

Outro exemplo de anomalia ocorreu no portal YahooNews em uma notícia sobre o primeiro ataque cardíaco desencadeado pelo consumo de café. Junto ao texto, a automação da mídia inseriu um anúncio<sup>12</sup> de café da marca Folger's, que em seu título trazia a frase “aproveite uma segunda xícara de

café”. Aqui, a relação entre os malefícios do consumo de café e o anúncio oferecendo uma segunda xícara também produzem uma contradição contextual.

Mais recentemente, anomalias contextuais envolvendo o Google por meio da plataforma YouTube também ocorreram. Anunciantes tiveram seus materiais exibidos em vídeos ligados a grupos extremistas.

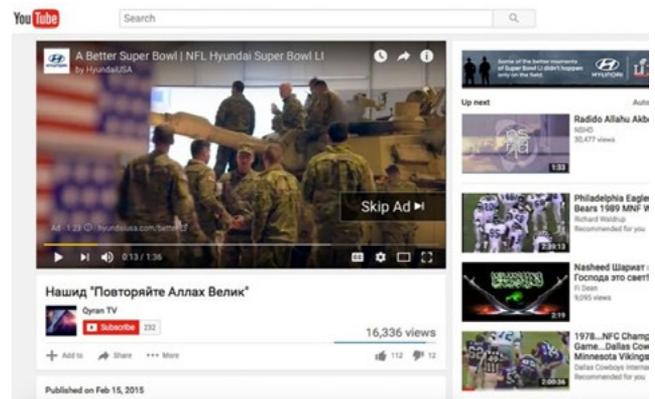


Figura 2 - Anúncio da Hyundai no YouTube junto a vídeo Jihadista. Fonte: Reprodução

O Google está sob ataque após vários anúncios serem veiculados ao lado de conteúdos ofensivos no Reino Unido – incluindo vídeos neonazistas e jihadistas – no YouTube e em outros websites que a sua rede serve anúncios. A revelação primeiramente reportada em fevereiro pelo jornal Times of London, levou a agência Havas a retirar todos anúncios da rede do Google incluindo Youtube. Outras empresas como Johnson & Johnson, JPMorgan, AT&T e Verizon entre outras, também suspenderam ou retiraram os anúncios junto ao Google (CASTILHO, 2017, s/p).<sup>13</sup>

<sup>13</sup> Trecho original: “Google is under fire after several ads were seen running next to offensive content in the U.K. — including neo-Nazi and jihadist videos — on YouTube and other websites it serves ads on. The revelation, first reported in early February by The Times of London, led media buying agency Havas to pull all its YouTube and Google Digital Network ads in the U.K. Johnson & Johnson, JPMorgan, AT&T and Verizon, among others, have suspended or pulled advertising with Google as well.” Acesso em: 20 dez., 2017. Tradução dos autores.

<sup>11</sup> Fonte: <https://static.businessinsider.com/image/4f91ac4feab8ea866e000009-400/image.jpg>. Acesso em 11 de set. 2017

<sup>12</sup> Fonte: <https://static.businessinsider.com/image/4f91ad00ecad04fd35000007-400/image.jpg>. Acesso em 11 de set. 2017

Este caso provocou uma série de debates entre anunciantes e plataformas de mídia programática em relação aos conteúdos relacionados aos espaços disponíveis nos inventários de mídia. Os anunciantes exigiam uma maior transparência e segurança em relação aos locais disponibilizados.

Esses problemas têm relação com os critérios e regras que permitem a entrada desse tipo de espaço nos inventários. A lógica de critérios mal desenhados para seleção de espaços para publicação de mídia automatizada também ocorreu num portal de notícias do Jihad que trazia um anúncio da IBM cujo tema era a empatia com os clientes para oferecer produtos na hora mais adequada.



Figura 3 - Anúncio da IBM publicado no portal Jihadista Arrahmah. Fonte: Reprodução

A partir dos exemplos exibidos, observamos que o cenário de automação de compra e venda da mídia online, apesar dos benefícios, ainda enfrenta muitas dificuldades tecnológicas. Os algoritmos responsáveis pelas trocas e coletas de informação que definem decisões e negociações entre os *players* do ecossistema programático necessitam evoluir ainda mais. Diante deste cenário, estrategistas digitais, profissionais de marketing e pesquisadores devem estar atentos ao poder que está sendo depositado em sistemas de automação de compra e venda de mídia. O ponto de discussão aqui se assemelha a outros emba-

tes que envolvem a curadoria técnica quando utilizada como única alternativa, ou seja, estamos prontos para confiar somente nos algoritmos?

## Considerações finais

Observamos neste trabalho, o caminho evolutivo da mídia online e como funciona a operação da compra e venda de espaços de publicidade na atualidade. A nova realidade proporcionada pelos métodos de negociação de mídia é sem dúvida cheia de benefícios tanto para os veículos quanto para os anunciantes, no entanto, junto às possibilidades trazidas pela tecnologia, há a complexidade desse novo ecossistema que precisa ser debatida. As plataformas de mídia programática interagem em velocidades sobre-humanas, dificultando assim a interferência manual, fazendo com que apenas os códigos consigam modelar a ação das máquinas.

Como deuses, estes modelos matemáticos são opacos, sua operação é invisível para todos exceto os cardeais em suas áreas: matemáticos e cientistas computacionais. Seus vereditos, mesmo quando errados ou nocivos, são imunes a disputas ou apelos" (ONEIL, 2016, p. 14-15).<sup>14</sup>

O problema ocorre quando há situações em que as máquinas tomam decisões inadequadas baseadas em dados desviantes e produzem anomalias e aberrações na operação. A velocidade com que a transação deve ocorrer é tão grande que não há maneira possível de supervisão humana durante o processo. Resta aos profissionais envolvidos analisar os resultados e exigir regras mais claras, maior transparência de códigos e algoritmos. A partir da coleta de resultados desviantes é possível ajustar os blocos responsáveis

14 Trecho original: "Like gods, these mathematical models were opaque, their workings invisible to all but the highest priests in their domain: mathematicians and computer scientists. Their verdicts, even when wrong or harmful, were beyond dispute or appeal." Tradução dos autores.

pelo comportamento indevido da transação de mídia online. Mapear os tipos de erros e anomalias parece ser um dos caminhos para aprimorar a operação dos códigos.

A utilização de relatórios padrão de exibição e controle de espaços do inventário de mídia baseado também no conteúdo do veículo indica outra forma de proteger anunciantes para não prejudicar sua marca associando-se a temas indevidos. Veículos considerados mais nobres podem adotar este tipo de medida para qualificar os espaços disponíveis em suas publicações, colocando-os numa posição elevada no índice de segurança da marca, dado que atesta a qualidade e a segurança do ambiente de mídia disponibilizado nos inventários de custo mais elevado.

Enquanto anunciantes estão se beneficiando da programática, veículos *premium* devem agir para garantir que eles não sejam rebaixados a meros fornecedores de espaços publicitários e não tenham que perder largos pedaços de sua verba” (MUNSTERMANN, 2016, p.36).<sup>15</sup>

Além é claro de melhorar as práticas, responsabilizando não apenas os códigos, mas as empresas e plataformas por trás dos algoritmos que são utilizados durante os leilões.

## Referências

BUSCH, Oliver. The programmatic advertising principle. In: BUSCH, Oliver (Org.). *Programmatic Advertising*. London: Springer, 2014.

CARDOSO, José Pedro Vieira. *Publicidade baseada em Informação de Contexto*. 2012. Dissertação (Mestrado em Engenharia Informática e Computa-

<sup>15</sup> Trecho original: "While advertisers are benefiting from programmatic advertising, premium publishers must take action to ensure that they are not degraded to mere suppliers of ad space and do not have to lose large chunks of their contribution margins" Tradução dos autores.

ção) - Faculdade de Engenharia da Universidade do Porto, Portugal, 72 p.

CARVALHO, Cristiane Mafacioli; CHRISTOFOLI, Márcia Pillon; BOMBARDELLI, Rita de Cássia Breier. *Novos Modelos e Novos Negócios na Prática do Mercado Publicitário*. Trabalho apresentado no GP Publicidade e Propaganda, XV Encontro dos Grupos de Pesquisas em Comunicação, evento componente do XXXVIII Congresso Brasileiro de Ciências da Comunicação. In: *Anais...*, 2015, Rio de Janeiro.

CORMEN, Thomas. *Desmistificando Algoritmos*. Rio de Janeiro: Elsevier, 2014.

CASTILLO, Michele. *Here's how it's possible for Google to sell ads next to offensive content and not know about it*. Disponível em:

<<https://www.cnbc.com/2017/03/24/google-ad-scandal-how-companies-buy-youtube-and-google-display-ads.html>>. Acesso em: 20 dez. 2017.

DOMINGOS, Pedro. *O Algoritmo Mestre*. São Paulo: Novatec, 2017.

IAB Brasil. *Manual de mídia programática*. São Paulo: IAB Brasil, 2015a.

\_\_\_\_\_. *Mídia programática e o desafio do consumidor multiconectado*. 2015b. Disponível em: <<https://iabbrasil.com.br/midia-programatica-e-o-desafio-do-consumidor-multiconectado/>>. Acesso em: 27 dez. 2017.

KANTROWITZ, Alex. *10 Things you need to know about programmatic buying*. Disponível em: <<http://adage.com/article/print-edition/10-things-programmatic-buying/298811/>> Acesso em: 16 jan. 2018.

LOPES, Eduardo Luís da Silva. *Detecção de anomalias em modelos de publicidade pay-per-click*. 2012. Dissertação (Mestrado em Engenharia Infor-

mática) - Universidade do Minho, Portugal, 106 p.

MUNSTERMANN, Holm; WURTENBERGER, Peter. *Programmatic disruption for premium publishers*. In: BUSCH, Oliver (Org.). *Programmatic Advertising: The successful transformation to automated, data-driven marketing in real-time*. Suíça: Springer, 2016.

O'NEIL, Cathy. *Weapons of math destruction: how big data increases inequality and threatens democracy*. Nova Iorque: Crown, 2016.

ROSA, Luccas Martins da. *Impactos do advento da comercialização programática de mídia digital no mercado publicitário brasileiro*. 2016. (Trabalho de Conclusão de Curso) - Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul, Porto Alegre, 148 p.

SILVEIRA, Stefanie C. da. *Os algoritmos no fluxo de conteúdos: uma visão a partir da TAR*. In: SAAD, E.; SILVEIRA, S. C. (Org.). *Tendências em Comunicação Digital* - São Paulo: ECA/USP [e-book], 2017, 2 v.

TASCÓN, M.; COULLAUT, A. *Big data y el internet de las cosas*. Madrid: Catarata, 2016.

TESTORI, Matteo. *The Applications of Video Analytics in Media Planning, Trade and Shopper Marke-*

*ting* In: DISTANTE C.; BATTIATO. S; CAVALARO. A (Org.) *Video Analytics for Audience Measurement*. Suécia: Springer, 2014.

ZANETTI, Guilherme Orsato. *Conflitos entre publicidade e privacidade na internet*. 2017. (Trabalho de Conclusão de Curso) - Universidade de Passo Fundo, Passo Fundo, 88 p.

**Stefanie Carlan da Silveira** – Professora adjunta da Universidade Federal de Santa Catarina (UFSC), doutora em Ciências da Comunicação pela Universidade de São Paulo (USP). **E-mail:** me@ssilveira.com

**João Gabriel Danesi Morisso** – Acadêmico do Curso de Especialização Lato Sensu em Gestão da Comunicação Digital da Universidade de São Paulo (USP).

Recebido: 16 mar. 2018

Aprovado: 07 maio 2018

# **Um jornalismo para chamar de meu? Algoritmos e o fenômeno da customização de notícias**

## **Is there a journalism to call my own? Algorithm and the phenomenon of news customization**

### **Raquel Dornelas**

Universidade do Estado do Rio de Janeiro  
Programa de posgraduação em Comunicação Social  
Rio de Janeiro, RJ, Brasil

## Resumo

Este artigo tem como objetivo apresentar uma breve revisão teórica sobre o fenômeno da criação de noticiários hiperindividualizados, possibilitados por serviços e plataformas disponíveis na Internet. Percebendo que a prática tem ampla influência dos algoritmos, acionam-se discussões sobre o conceito de “algocracia” e “bolha informacional”, a fim de fornecer insumos teóricos sobre o processo mencionado. Sem desconsiderar o fato de que o jornalismo tradicional massivo ainda exerce ampla influência na sociedade e procura sobreviver se adaptando à lógica digital, conclui-se que é urgente tornar mais transparentes questões que estão por trás da oferta das novas ferramentas de personalização: a privacidade de dados dos usuários, os riscos de usos ilegais e a “financeirização” das informações.

## Abstract

This paper aims to present a brief theoretical review about the phenomenon in which hyper-individualized newspapers are created due to services and platforms available online, especially social networks. Realizing that the practice has wide influence of the algorithms, the paper discusses the concepts of algocracy and informational bubble, in order to provide theoretical overviews on the aforementioned process. Without disregarding the fact that traditional journalism still has a large influence on society and seeks to survive by adapting to digital logic, we claim that it is urgent to make public what are behind the offer of these new personalization tools: data privacy, risks of illegal uses and the financialization of information.

## Palavras-chave

Algoritmos. Algocracia. Bolha informacional. Redes sociais.

## Keywords

Algorithms. Algocracy. Informational bubble. Social networks.

## Introdução

Quando Robert Park (1996) escreveu, na primeira metade do século XX, que a notícia é uma forma de conhecimento, a imprensa tradicional exercia um papel diferente do que vemos na atualidade. Situado entre o saber científico e o senso comum, o conhecimento gerado pelos jornais impressos alcançaria uma dimensão pública. Naquele contexto, o noticiário cumpria, na perspectiva do autor, um papel crucial de orientar a população, de formular uma agenda que chamasse a atenção das pessoas e de informar sobre o que acontecia no mundo. Além disso, a interpretação das notícias teria ainda outra importante função: a constituição da opinião pública.

De lá para cá, muita coisa mudou. A crise de confiabilidade nas instituições (inclusive a imprensa) e o surgimento de outras vias informativas tangenciais estremeceram um pouco a centralidade do jornalismo tradicional.

Uma das principais transformações nesse cenário foi o surgimento das tecnologias digitais, mais especificamente a Internet. É sabido que a maior parte da informação noticiosa que temos acesso ainda continua nas mãos, principalmente, de conglomerados de informação massiva. A lógica anterior de concentração midiática não mudou completamente. Mas a massificação dos ambientes on-line complexificou o processo de produção, disseminação e consumo de notícias.

Para Alex Primo (2011), o que presenciamos hoje é uma mudança estrutural do sistema midiático e um deslocamento do papel das instituições noticiosas – tão centrais no pensamento de Park. Mais de duas décadas depois da chegada da Internet comercial ao Brasil, ocorrida em 1995, muitas questões ainda merecem ser levantadas. Ainda há uma considerável lacuna entre a preparação técnica dos repórteres e as possibilidades digitais<sup>1</sup>, a rela-

1 Conforme o estudo “The state of technology in global newsrooms”. Disponível em [www.icfj.org/resources/first-ever-global-survey-news-tech-reveals-perilous-digital-skills-gap](http://www.icfj.org/resources/first-ever-global-survey-news-tech-reveals-perilous-digital-skills-gap). Acesso em 01 jan. 2018.

ção produtor/consumidor de notícias está cada vez mais híbrida e as empresas de comunicação trabalham hoje em conjunto com a dinâmica das chamadas redes sociais<sup>2</sup> – sistemas pertencentes, prioritariamente, a instituições não jornalísticas. Além disso, a notícia que lemos hoje também é elaborada e consumida em rede.

Merecem ser destacadas também a fragmentação da audiência; a atuação incisiva dos algoritmos; e a possibilidade de customização da informação recebida. Para entender a atuação de tais elementos no cenário de hoje, este trabalho pretende caminhar à luz de discussões que passem pela lógica dos algoritmos, pelas imbricações entre redações tradicionais e produção de notícia digital, pela privacidade de dados na web. Longe de se adotar uma perspectiva entusiasta ou pessimista das plataformas digitais, o que se pretende é realizar uma discussão que não cometa o deslize de incorporar um fetiche à técnica, sem inseri-la em tensionamentos próprios da tessitura social e do consumo de notícia.

No primeiro tópico, será apresentada a influência dos algoritmos nesse novo cenário. A seguir, o trabalho evidencia discussões sobre a interação dos sujeitos com os códigos de programação das redes sociais e serviços digitais noticiosos. Por fim, apresentam-se possíveis consequências do fenômeno da hipersegmentação jornalística. Mais do que chegar a resultados a partir de um objeto específico, pretende-se mapear reflexões que possam servir de ponto de partida para futuras pesquisas que orbitam em torno da temática.

## Os algoritmos e a bolha informacional

Já é realidade o fato de que hoje nos informamos bastante pela Internet, incluindo as redes sociais. Segundo o relatório “Reuters Institute Digital

2 Existem outros termos para se referir a essas plataformas. Para este trabalho, utilizaremos “redes sociais”, apesar de estarmos cientes que a genealogia do conceito vai além dos dispositivos hospedados na internet. Para mais, recomendamos a leitura de Recuero (2011).

News Report 2017<sup>3</sup>, a Internet é o segundo meio pelo qual o brasileiro obtém dados sobre o que ocorre no país, perdendo apenas para a televisão. Mais da metade (66%) se informa no ambiente on-line. Quanto ao tempo de conexão, a população permanece conectada, em média, mais de 4h40 por dia, durante a semana, e mais de 4h30 diariamente aos sábados e também aos domingos (BRASIL, 2017).

Entre as plataformas de comunicação disponíveis na web, as redes sociais ganham destaque e o Facebook assume uma posição privilegiada. Criada em 2004 nos Estados Unidos, é considerada a maior rede social do mundo, com mais de dois bilhões de usuários. Ocupa também o primeiro lugar entre as plataformas da mesma natureza no Brasil, com cerca de 111 milhões de contas ativas por mês.<sup>5</sup>

O Facebook é também a primeira rede social que os brasileiros procuram para se informar: 57% dos que possuem uma conta a utilizam para ler notícias. Ou seja, mais da metade dos 111 milhões de perfis estão logados para obter conteúdo jornalístico. Apesar de recentes levantamentos (FAAP, 2017; BRASIL, 2017)<sup>4</sup> apontarem que o Facebook tem perdido audiência para outros serviços, incluindo empresas do mesmo grupo (como Instagram e WhatsApp), os dados ainda comprovam a importância da rede pertencente ao norte-americano Mark Zuckerberg.

Segundo anúncio do próprio Facebook, realizado em janeiro de 2018, a empresa tem privilegiado a exibição de conteúdo de amigos em detrimento ao de páginas (pertencentes a figuras públicas, instituições, causas e veículos de comunicação). Com as alterações, a necessidade de impulsionamento (as páginas necessitam pagar para aparecer na *timeline*) torna o acesso às notícias ainda mais escasso e segmentado. A Folha de S. Paulo<sup>5</sup> já comunicou que não

irá mais publicar na plataforma. A empresa paulista criticou a não valorização de conteúdo realizado por jornalistas profissionais e admitiu que a decisão foi influenciada por questões financeiras exigidas pelo Facebook. No comunicado, a Folha cita o projeto *Instant Articles*, que sugeria aos veículos transferir todo o conteúdo para a rede de Zuckerberg, sem ganhar qualquer remuneração por isso, com o argumento de maior rapidez no carregamento das páginas.

A Folha de S. Paulo lamentou também que a quase obrigatoriedade de pagamento para exibir publicações na “vitrine” do Facebook resultará na hipersegmentação do conteúdo. Tal fato poderia intensificar a formação de nichos ideológicos, segundo a opinião do veículo. “Isso reforça a tendência do usuário a consumir cada vez mais conteúdo como o qual tem afinidade, favorecendo a criação de bolhas de opiniões e convicções, a propagação de ‘fake news’ [...]” (FOLHA DE S. PAULO, 2018, p. 8).

Neste ponto, é preciso lembrar que a exibição de informação em plataformas, como Facebook, Twitter e Instagram, obedece, principalmente, à lógica dos algoritmos. Trata-se de códigos de programação, constantemente atualizados, cuja promessa é entregar um conteúdo mais relevante para o usuário, dentro do que a plataforma julga ser interessante para eles. Portanto, a maior parte do que vemos em tais ambientes é ditada pelos algoritmos – o que ocorre também nos serviços de buscas.

Costa (2014) esclarece que o algoritmo das redes sociais pode barrar informações que a plataforma julga estar fora de sua política editorial, mesmo que os jornais produtores do conteúdo original as considerem éticas e importantes. A lógica de funcionamento dos algoritmos que influenciam na disponibilidade das informações, evidentemente, não é transparente. As redes sociais e os aplicativos para dispositivos móveis são propriedades de empresas privadas que objetivam o lucro e não têm a obrigação de disponibilizar seu código publicamente. O segredo tecnológico de cada companhia garante o diferencial competitivo, a particularidade

3 Disponível em <http://po.st/lfJFXh>. Acesso em 01 jan. 2018.

4 Disponível em [http://faap.br/nimd/pdf/ms360faap/MS-360FAAP\\_2018-%2001\\_Q4.pdf](http://faap.br/nimd/pdf/ms360faap/MS-360FAAP_2018-%2001_Q4.pdf). Acesso em 08 jan. 2018.

5 Disponível em: [www1.folha.uol.com.br/poder/2018/02/foha-deixa-de-publicar-conteudo-nofacebook.shtml](http://www1.folha.uol.com.br/poder/2018/02/foha-deixa-de-publicar-conteudo-nofacebook.shtml). Acesso em 08 jan. 2018.

dos serviços prestados e, portanto, sua sustentabilidade no mercado. Apesar disso, existem pressões de governos, teóricos e ativistas para demarcar regulamentações quanto ao modo de ação dos algoritmos e ao fato de como eles podem chegar a influenciar tomadas de decisões ou consolidar tendências entre os usuários.

Danaher (2016) e O'Neil (2016) são alguns dos críticos à opacidade dos códigos informacionais. Os autores tentam alertar para o fato de que estamos vivendo em uma verdadeira "algorocracia": um cenário no qual uma infinidade de dados produzidos em massa é coletada, armazenada, rearranjada por códigos de programação, acabando, em algum grau, organizando nossas vidas e o que vemos na Internet. Assim como Costa, Danaher (2016) destaca que os algoritmos produzem impactos no nosso mundo social, econômico, cultural e político, sob uma perspectiva pouco transparente e inacessível à maior parte da população. Já para O'Neil (2016), a grande questão é que tais códigos são programados pelos próprios humanos e estão, conseqüentemente, enraizados em preconceitos sociais. Tecnologias, assim como os seus desenvolvedores, são falhas e podem cometer injustiças quando avaliam profissionais, medem a periculosidade de criminosos e decidem se uma informação é útil ou verdadeira, utilizando apenas métricas matemáticas e o sistema binário da tecnologia da informação (O'NEIL, 2016).

María Flores Dorda (2017), apesar de versar sobre a temática do consumo musical, ajuda-nos a pensar como os algoritmos de recomendação oferecem respostas individualizadas para cada usuário também no campo noticioso, construindo um mundo informacional virtual etiquetado com a identidade de cada um, "contendo aspectos como sus gustos, intereses, datos personales o aficiones" (*idem*, 2017, p. 9). Recorrendo a Eli Pariser (2011), Dorda também destaca que tal cenário pode ocasionar o "efeito bolha": tamanha individualização de conteúdo criaria um universo próprio para cada usuário, em vez de contribuir para um mosaico de troca informacional em rede.

Evidentemente, não é intenção deste trabalho negligenciar o poder de agência dos internautas frente à tecnologia. Para além dos mecanismos algorítmicos, cada um de nós também ativa funções que ajudam a dar forma ao "efeito bolha". Temos, naturalmente, a tendência de nos conectar virtualmente a pessoas que pensam de forma parecida conosco e de conversar sobre tópicos de igual concordância. Além disso, podemos facilmente deixar de seguir uma conta que forneça conteúdo não agradável ou desfazer uma amizade virtual com usuários dos quais discordamos. No entanto, somada as nossas intervenções às influências dos algoritmos, aumentam as chances de passarmos a conviver em nichos cada vez mais homogêneos na Internet, potencializando a distância entre os que pensam diferente de nós.

Essas bolhas tendem a isolar os atores dentro de grupos onde apenas alguns tipos de informação circulam, criando uma percepção falsa de [esfera pública] (onde "todos" falam) e de opinião pública (onde a "maioria" concorda). Ao mesmo tempo, pesquisas têm demonstrado que a mídia social é hoje um dos principais canais informativos do grande público. Com isso, o silenciamento do contraditório pode ter efeitos no posicionamento político e nas próprias instituições democráticas. Esses elementos são particularmente importantes em contextos políticos de crise, como o do Brasil atualmente. (RECUERO; ZAGO; SOARES, 2017, p. 2)

Porém, é preciso relativizar e lembrar que a aproximação de pessoas com visões de mundo semelhantes e a segmentação de conteúdo informativo são fenômenos que sempre existiram. Basta pensarmos na TV a cabo, com inúmeras opções de assuntos que vão desde programas para apaixonados por carros até canais de leilões de joias online. Cada qual com sua audiência fiel. O jornalismo também sempre possuiu, em sua natureza, o cará-

ter de personalização de conteúdo. Podemos citar tanto os impressos que já nascem segmentados, quanto as revistas de moda, de artesanato ou de dicas de alimentação, abrangendo ainda os cadernos ou seções separadas, com informações bem específicas: esportes, cultura, economia e política.

No entanto, o cenário que se problematiza aqui é outro. A segmentação clássica que o jornalismo oferecia, até então, possuía certo limite. A revista sobre carros, por exemplo, era direcionada tanto para quem estava interessado em saber sobre os últimos lançamentos quanto para um profissional da área da mecânica que pretendia entender sobre motores. O jornal que trazia o suplemento feminino permitia às leitoras também terem contato com o mesmo conteúdo de política que o restante da família iria ler. A customização da informação ocorria até certo ponto, mas não alcançava o nível da hiperpersonalização que se percebe hoje.

De acordo com algumas perspectivas, a atuação dos algoritmos nas redes sociais e até mesmo nos portais jornalísticos, tem proporcionado novas camadas da individualização da notícia, tornando o conteúdo tão personalizado que poderia fragmentar a troca de informações e, consequentemente, o debate coletivo.

Alguns veículos no Brasil e no mundo estão produzindo conteúdo específico para cada rede social. Ou direcionando as reportagens de acordo com o perfil dos consumidores. Fazendo paralelo com os jornais de papel, é como se as mulheres recebessem em casa apenas o suplemento feminino, as crianças o infantil e os homens o caderno de esportes. Reforçando estereótipos, preconceitos e bolhas digitais. Nesse ritmo, no futuro, podemos pensar em enviar certo tipo de notícia para os eleitores do Bolsonaro e um outro, completamente diferente, para os que acreditam no Lula. É o velho e amarelado papel social do jornalista sendo trocado pela audiência das redes sociais no papel. (VIEIRA, 2017, n.p.)

Outras abordagens caminham por uma via diferente, acreditando na maior possibilidade de diálogo no ambiente on-line permeado por algoritmos – falaremos disso adiante. Além disso, não podemos nos esquecer de que a ação dos usuários, como já foi dito, é central para o fenômeno da personalização. Tanto em *websites* quanto nas redes sociais, o internauta não é um ente passivo. É o que será tratado na seção a seguir.

### **Usuário e algoritmo: uma ação conjunta**

Ações deliberadas dos usuários podem incrementar o fenômeno da hiperindividualização do conteúdo. Já em 1995, Negroponte falava sobre o jornal do futuro, o *Daily Me*: edições feitas de forma totalmente customizadas para cada leitor. Publicado no mesmo ano em que a Internet comercial chegava ao Brasil, o livro previa um deslocamento da comunicação de massa para a personalizada, costurado pela ubiquidade da tecnologia digital na vida cotidiana.

What if a newspaper company were willing to put its entire staff at your beck and call for one edition? It would mix headline news with “less important” stories relating to acquaintances, people you will see tomorrow, and places you are about to go to or have just come from. It would report on companies you know. In fact, under these conditions, you might be willing to pay the Boston Globe a lot more for ten pages than for a hundred pages, if you could be confident that it was delivering you the right subset of information. You would consume every bit (so to speak). (NEGROPONTE, 1995, p. 153)

Alinhada a essa discussão, um vocábulo pouco usual no campo jornalístico tem sido agora amplamente usado: *bricolagem*. O termo se refere ao trabalho manual e é encontrado no âmbito das belas artes, arquitetura e até no artesanato. A palavra também já apareceu em discussões sobre jornalismo, mas com outra conotação. Segundo Charron,

Damian-Gaillard e Travancas, o trabalho da imprensa é realizado

[...] mais ou menos livremente através da ajuda dos colegas, dos conhecimentos práticos, das linhas do métier, do know-how. A organização empírica do trabalho jornalístico inclui uma parte de bricolagem e de artesanato. (CHARRON; DAMIAN-GAILLARD; TRAVANCAS, 2014, p.15)

Mas, recentemente, o termo tem sido acionado para se referir ao esforço do cidadão comum (não o jornalista profissional) em construir um mosaico de informações de seu interesse, a partir do que ele encontra on-line. Poderíamos dizer que o trabalho é quase artesanal – se não fosse a ajuda das plataformas digitais que auxiliam no processo. Se na década de 2000, a internet presenciou mudanças na distribuição de informações, os anos 2010 parecem ter dado uma guinada na lógica de organização e curadoria de tais dados.

Pollyana Ferrari (2007) utiliza o termo bricolagem para se referir à vitrine de informações agrupadas, dando destaque para o uso do hipertexto. Segundo a autora, os dados ligados em redes virtuais podem formar uma memória coletiva, em oposição à memória fragmentada. As formas hipertextuais permitiriam ao usuário, portanto, fornecer acesso “a arquivos pessoais e coletivos, todos remixados em uma eterna bricolagem de narrativas, que podem ser textuais, imagéticas, audíveis ou sensoriais” (FERRARI, 2007, p. 84).

Ferrari adota uma perspectiva, a nosso ver, demasiadamente otimista com relação à internet, mas o fragmento acima nos ajuda a pensar que a bricolagem de informações permite ao usuário voltar-se para um conjunto de informações já disponíveis no ambiente virtual e constituir, gradativamente, uma espécie de inventário de conteúdo que lhe ajude ou lhe interesse de alguma forma individualmente – uma verdadeira colagem de bits.

Para ilustrar o assunto, há o exemplo de iniciativas de dois gigantes do setor tecnológico.

Duas empresas, que não são do ramo jornalístico, criaram recursos que permitem ao leitor elaborar o seu próprio noticiário e recebê-lo literalmente na palma das mãos. A *Apple News* e o *Google News* são ferramentas concorrentes, mas que possuem funcionamento e objetivos bastante semelhantes. Ambas permitem que o usuário dos sistemas *iOS* e *Android* colete notícias de veículos parceiros e agrupem-nas em dispositivos móveis, como celulares e *tablets*. Entre as configurações de customização, o leitor pode filtrar o conteúdo por região e idioma, gerenciar e modificar, a qualquer momento, seus assuntos de interesses, além de indicar preferências de fontes.

Atualmente, a *Apple News* está disponível apenas nos Estados Unidos, Austrália e Grã-Bretanha, mas a companhia fundada por Steve Jobs oferece *widgets* (pequenas extensões) para que usuários de outros países possam usufruir do recurso. Já o *Google News* também está disponível na versão para *desktop*. Só no Brasil, o aplicativo já foi baixado por mais de 100 milhões de pessoas e promete oferecer cobertura noticiosa de mais de 75 periódicos, vinda de 60 países, além de permitir as configurações de personalizações semelhantes a do seu principal concorrente.

Existem outras iniciativas dessa natureza e listar todas fugiria ao escopo deste trabalho. O importante é frisar que a maioria funciona sob a mesma lógica. Como a própria *Apple* explica, tais plataformas prometem que “you’re only informed about the stories you care about most”, ou de acordo com o *Google*, você terá “sua própria sala de redação personalizável em tempo real”. Temos à frente um discurso que apela para a economia de tempo, para a individualização (tratada como sinônimo de qualificação) da informação recebida e para o maior reconhecimento cognitivo e emocional do leitor com os fatos do dia.

Uma questão que se levanta entre alguns estudiosos do tema é: tal grau de customização pode provocar certo isolamento dos atores sociais no de-

bate coletivo, prejudicando as decisões públicas e o diálogo entre aqueles que assumem posicionamentos diferentes?

## Existem consequências para um debate coletivo?

Como já foi mencionado, a segmentação de notícias sempre existiu. O que pode ser considerado novo no horizonte são o grau da customização e o fato do consumo de informação (inclusive noticiosa) estar sendo realizado em dispositivos que, além de não serem jornalísticos, estão imunes à regulamentação que rege os meios de comunicação tradicionais, como o rádio e a televisão. Quando entramos no mérito das redes sociais, estamos falando de companhias (não noticiosas) responsáveis por nutrir e hierarquizar informações (noticiosas) para uma parcela considerável da população. A informação está nas mãos de conglomerados que, a princípio, não são de *conteúdo*, mas sim de *tecnologia* – com alto poder de investimento. As cinco empresas mais valiosas no mundo são justamente a *Apple*, o *Google*, a *Microsoft*, a *Amazon* e o *Facebook*<sup>6</sup>.

Entre as abordagens que se preocupam com o cenário acima, a questão central passa pelo entendimento de como tais empresas atuam: os dados coletados dos usuários são também insuportáveis para a “financeirização”<sup>7</sup> do ramo da comunicação, objetivando lucros no mercado de derivativos (MARTIN, 2013; ROSSI, 2011; BRAYAN e RAFFERTY, 2014). As informações dos usuários, incluindo reações, tendências e gostos, são agrupadas em *commodities* com probabilidades de lucros, se tornando produtos virtuais comercializados nas bolsas de valores.

6 Disponível em [https://brasil.elpais.com/brasil/2017/11/03/economia/1509714366\\_037336.html](https://brasil.elpais.com/brasil/2017/11/03/economia/1509714366_037336.html). Acesso em 07 jan. 2018.

7 Termo utilizado pelos próprios autores, incluindo a tradução encontrada em Rossi (2011). O vocábulo denota o fenômeno da difusão da lógica do mercado de valores para setores e práticas da vida social que estão além da esfera econômica.

Neste ponto, realiza-se um paralelo com a discussão de Kera (2017). Ao versar sobre a cultura *maker*, do *software livre* e do “faça você mesmo”, a autora teme que a criatividade descentralizada, típica dessa cultura, possa na verdade nutrir uma agenda isolacionista.

No fim de 2017, Chamath Palihapitiya, ex-vice-presidente do *Facebook*, afirmou que se arrepende de ter trabalhado na empresa e que a rede estaria dividindo a sociedade, prejudicando um debate cooperativo. “The short-term, dopamine-driven feedback loops that we have created are destroying how society works: no civil discourse, no cooperation, misinformation, mistruth” (PALIHAPITIYA, 2017).<sup>8</sup>

Em contrapartida às visões mais críticas, Assunção et al (2015) sublinham que, mesmo com as possibilidades de segmentação, as redes sociais são mais suscetíveis à diversidade política, em função da amplitude das conexões e do fato dos usuários terem a oportunidade de interagir com atores pertencentes a outros círculos, dentro e fora de suas redes mais próximas. Ao abordar a temática das campanhas eleitorais on-line, por exemplo, os autores acreditam que as funcionalidades de tais plataformas podem agregar pessoas com interesses distintos, aproximando-as de assuntos outrora não familiares.

As redes sociais, inseridas num contexto de lazer cotidiano, podem promover a exposição a conteúdo por parte de cidadãos que não estão particularmente interessados, contribuindo para reduzir lacunas de conhecimento. (ASSUNÇÃO, et al, 2015, p. 21)

Evidentemente seria inocente reduzir a ideia de esfera pública contemporânea ao ambiente virtual. O campo midiático, incluindo as redes sociais, não se constitui como o lugar único da pluralidade de vozes em discussão, mas sim como um agente que também tem interesses específicos, atuando

8 Disponível em: [www.youtube.com/watch?v=PMotykwOSIk](http://www.youtube.com/watch?v=PMotykwOSIk). Acesso em 17 jan. 2018.

junto a outros poderes em tensão. Apesar disso, não se pode negar que, em uma sociedade midiaticizada (BRAGA, 2006), os meios de comunicação, mesmo os não noticiosos por excelência, podem ser considerados um dos espaços privilegiados para conversação coletiva. Aldé e Marques (2015) sublinham essa influência da Internet em decisões políticas.

Embora as ferramentas de comunicação digital componham apenas uma parte de todo o aparato midiático [...], o comportamento do eleitor é afetado pelo perfil de sua rede de contatos e pelos conteúdos aos quais ele é exposto [...] (ALDÉ; MARQUES, 2015, p. 7).

Bucher (2017) traz à tona uma perspectiva otimista, mostrando que a ação dos algoritmos pode auxiliar no trabalho da imprensa. Ao entrevistar profissionais nas redações acerca da relação entre jornalismo e tecnologias da informação, a autora encontrou afirmações que defendem o fato dos algoritmos não agirem isolados no processo de agenciamento da notícia. Alguns acontecimentos seriam tão amplamente importantes que a intervenção humana, ao publicá-los, ultrapassaria qualquer ação técnica. Assim, tecnologia e olhar humano trabalhariam em conjunto para manter a população informada sobre os principais fatos que ocorrem no mundo.

The case of Omni, in particular, shows that the computational is much more than a tool or a means to an end. Rather, the computational is something to think with, and to be oriented towards. Far from just a passive thing operating in the background, the news-ranking algorithm of Omni also seems to impact the ways in which editors appoint the newsworthiness of a particular story to better accommodate its overall performance within the system. (BUCHER, 2017, p. 927).

A autora ressalta que as entrevistas com os profissionais nas redações mostraram que não há um via de mão única, pela qual o algoritmo se sobreporia do

fenômeno da noticiabilidade. Na verdade, os códigos ajudariam a remodelar as práticas humanas (a dos jornalistas), introduzindo novos *insights*, constrangimentos, possibilidades e formas de agenciamento.

## Considerações finais

A hipercustomização do noticiário, pela ação dos atores sociais em conjunto com os algoritmos, faz emergir algumas reflexões. A Internet, que em seus primórdios trazia lampejos de uma maior multiplicidade de vozes, também é palco para conflitos discursivos, tendo sido, em grande parte, incorporada ao *ethos* das grandes corporações capitalistas. Talvez outro caminho fosse mais difícil, devido à necessidade de alto investimento financeiro para ligar milhares de dispositivos por todo o mundo.

Ao se inserir em uma rede social, os usuários concordam, em tese, com o *modus operandi* de tais plataformas, ao aceitar espontaneamente os termos e condições de uso dos serviços digitais. No entanto, tal conformidade nem sempre está 100% clara na mente dos internautas, seja por falta de motivação para ler os extensos acordos ou mesmo pelo não entendimento dos vocábulos técnicos contidos em tais “contratos”. Segundo levantamento realizado pela Folha de S.Paulo, a leitura dos termos de uso de alguns dos serviços mais populares leva cerca de 4,5 horas para ser finalizada.<sup>9</sup>

Mesmo ciente de todas as regras, o usuário tem mais possibilidades de ser lesado do que as grandes companhias. Em março de 2018, dados de 87 milhões de contas do Facebook foram coletados e utilizados de forma ilegal pela empresa Cambridge Analytica que atua, entre outras frentes, com mapeamento de perfis eleitorais.

A relação entre usuário e os conglomerados também se mostra assimétrica quando tocamos na ques-

<sup>9</sup> Disponível em: [www1.folha.uol.com.br/tec/2017/12/1945132-leitura-de-termos-e-condicoes-de-servico-s-na-internet-exige-45-horas.shtml](http://www1.folha.uol.com.br/tec/2017/12/1945132-leitura-de-termos-e-condicoes-de-servico-s-na-internet-exige-45-horas.shtml). Acesso em 17 abr. 2018.

tão do fluxo dos dados. Ziegeldorf, Morchon e Wehrle (2014), ao versar sobre a “Internet das coisas”, alertam para o percurso obscuro que as informações realizam no ambiente virtual. Além das legislações sobre privacidade na internet serem diferentes entre si e até inexistente em alguns locais, os dados individuais caminham por trilhas difíceis de serem rastreadas. Existe ainda a possibilidade de quebra da criptografia de informações anônimas. Os dados pessoais passam por sistemas controlados por diferentes empresas, cada uma com sua própria política de privacidade. Quando se acessa uma rede social, por exemplo, são expostas informações privadas para a própria companhia criadora da plataforma, mas também para o provedor da rede de Internet, para o proprietário dos canais de fibra ótica e outros entes envolvidos na coleta, distribuição e organização dos dados. Um simples acesso à web depende de vários sistemas que agem em integração, mas cujos parâmetros de segurança da informação são pouco transparentes – sem considerarmos ainda vazamentos acidentais, ataques de terceiros ou espionagem.

Tais questões colocadas não pretendem clamar por um abandono em massa aos serviços digitais – seja para construir um noticiário customizado, seja para qualquer outra finalidade. O que se pretende é chamar a atenção para a urgência de tornar mais transparentes e compreensíveis as formas pelas quais podemos ser afetados pela coleta de dados e pela ação dos algoritmos.

As plataformas on-line são extremamente importantes para a dinâmica social de hoje. A utilização de um aplicativo, de uma rede social ou de um *website* permite que diferentes atores se conectem, interajam, vivenciem múltiplas experiências, consumam e ajudem a produzir notícias. Seria apressado dizer, como propõem os mais pessimistas, que o ofício do jornalismo está desaparecendo – até porque os próprios veículos de comunicação têm utilizado as redes sociais e outros serviços digitais para disseminar sua produção autoral.

A reflexão que se encerra aqui tem como objetivo nada mais do que apresentar um fenômeno em

andamento e alertar para possíveis implicações (e riscos). Muito mais do que estimular uma dicotomia entre organizações de tecnologia e empresas noticiosas, pretendemos provocar os pesquisadores em Comunicação a pensar na interlocução entre os vários entes (humanos e não humanos) envolvidos nessa relação e nos novos arranjos que se organizam na confecção de noticiários segmentados. É preciso continuar refletindo sobre a relação jornalismo e leitor, sobre o poder de agenciamento da imprensa tradicional e sobre as noções de *interesse público* e *interesse do público*, mediante a popularização de noticiários tão individualizados.

## Referências

- ALDÉ, Alessandra; MARQUES, Francisco Paulo Jamil Almeida (Orgs.). In: *Internet e poder local*. Salvador, Rio de Janeiro: Edufba, 2015.
- ASSUNÇÃO, Alysson et al. Estratégias de campanha política on-line: Marcelo Freixo nas eleições para a prefeitura do Rio de Janeiro em 2012. In: ALDÉ, Alessandra; MARQUES, Francisco Paulo Jamil Almeida (Orgs.). *Internet e poder local*. Salvador, Rio de Janeiro: Edufba, p.13-48, 2015.
- BRAGA, José Luiz. Mediatização como processo interacional de referência. *Animus*. Santa Maria, n. 2, p. 9-35, 2006, 5 v.
- BRASIL. Secretaria Especial de Comunicação Social da Presidência da República. *Pesquisa brasileira de mídia 2016: hábitos de consumo de mídia pela população brasileira*. Brasília: SECOM, 2017.
- BRAYAN, Dick; RAFFERTY, Michael. Financial derivatives as social policy beyond crisis. *Sociology*. n. 5, p. 887-903, 2014, 48 v.
- BUCHER, Taina. Machines don't have instincts: articulating the computational in journalism. *New Media & Society*, n. 6, p. 918-933, 2017, 19 v.
- CHARRON, Jean; DAMIAN-GAILLARD, Béatrice;

- TRAVANCAS, Isabel. Introdução. *Sobre jornalismo: os invisíveis do jornalismo*. n. 1, p. 14-17, 2014, 3 v.
- COSTA, Caio Tulio. Um modelo de negócio para o jornalismo digital. *Revista de Jornalismo ESPM/CJR*. São Paulo, 2014.
- DANAHER, John. The threat of algocracy: reality, resistance and accommodation. *Philos. Technol.* n. 29, p. 245-268, 2016.
- DORDA, María F.. *Serendipia programada: la construcción del gusto según las recomendaciones de Spotify*. Trabajo de fin de Máster. Getafe: Universidad Carlos III de Madrid, 2017.
- FERRARI, Pollyana. *Jornalismo digital*. São Paulo: Contexto, 2007.
- KERA, Denisa. A liminaridade da cultura maker e o hardware de fonte (na ciência): ao invés de fazer algo ser grande de novo, continue experimentando! *Liinc em Revista*, Rio de Janeiro, n.1, p. 7-28, maio 2017, 13 v.
- MARTIN, Randy. After economy? Social logics of derivatives. *Social Text*. n. 1, p. 83-106, 2013, 113 v.
- NEGROPONTE, Nicholas. *Being digital*. London: Hodder & Stoughton, 1995.
- O'NEIL, Cathy. *Weapons of math destruction*. New York: Crown, 2016.
- PARISER, Eli. *The filter bubble: what the Internet is hiding from you*. Penguin UK, 2011.
- PARK, Robert E. A notícia como forma de conhecimento. In: STEINBERG, Charles S. (org.). *Meios de Comunicação de Massa*. São Paulo: Cultrix, p. 169-185, 1996.
- PRIMO, Alex. Transformações no jornalismo em rede: sobre pessoas comuns, jornalistas e organizações; blogs, Twitter, Facebook e Flipboard. *Intexto*. Porto Alegre, RS, n. 25, p. 130-146, 2011, dez. 2011.
- RECUERO, Raquel. *Redes Sociais na Internet*. Porto Alegre: Sulina, 2011.
- \_\_\_\_; ZAGO, Gabriela da Silva; SOARES, Felipe Bonow. Mídia social e filtros-bolha nas conversações políticas no Twitter. XXVI Encontro Anual da Associação Nacional de Programas de Pós-Graduação em Comunicação (COMPÓS). 06-09 jun. *Anais*. São Paulo: Faculdade Cásper Líbero, 2017.
- ROSSI, Pedro. O protagonismo dos derivativos no capitalismo contemporâneo. *IV Encontro Internacional da Associação Keynesiana Brasileira*. Rio de Janeiro, 2011.
- VIEIRA, Agostinho. *Como as redes sociais estão afetando o compromisso social do jornalismo*. Disponível em: <<https://portal.comunique-se.com.br/como-as-redes-sociais-estao-afetando-o-compromisso-social-do-jornalismo/>>. Acesso em: 01 jan. 2018.
- ZIEGELDORF, Jan H.; MORCHON, Oscar G.; WEHRLER, Klaus. Privacy in the Internet of Things: threats and challenges. *Security and Communication Networks*. n.12, p. 2728-2742, 2014, 7 v.

**Raquel Dornelas** – Doutoranda em Comunicação Social pela Universidade do Estado do Rio de Janeiro (UERJ). E-mail: raqueldornelas@gmail.com

Recebido: 17 mar. 2018

Aprovado: 01 maio 2018



# A relevância dos algoritmos\*

## The relevance of algorithms

**Tarleton Gillespie**

Cornell University  
Ithaca, NY, Estados Unidos da America

\* Artigo publicado originalmente por Tarleton Gillespie sob o título "The relevance of algorithms", no livro *Media Technologies: Essays on Communication, Materiality, and Society* (MIT Press, 2014). Traduzido por Amanda Jurno mediante autorização do autor e da editora. Revisão: Carlos d'Andréa

## Resumo

Algoritmos (particularmente aqueles utilizados pelos mecanismos de busca, plataformas de mídia social, sistemas de recomendação e bases de dados) exercem uma função crescentemente importante em selecionar qual informação deve ser considerada mais relevante para nós, uma característica crucial da nossa participação na vida pública. À medida que assumimos as ferramentas computacionais como nossa forma primária de expressão, sujeitamos discurso e conhecimento humanos às lógicas de procedimento que sustentam a computação. Precisamos questionar os algoritmos como elementos chave de nosso sistema informacional e das formas culturais que emergem de suas sombras. Devemos ter especial atenção sobre onde e de que forma a introdução dos algoritmos nas práticas do conhecimento humano podem levar a ramificações políticas. Este artigo oferece um mapa conceitual para fazer isso. Ele propõe uma análise sociológica que não concebe os algoritmos como algo abstrato ou façanha técnica, e sim sugere como revelar as escolhas humanas e institucionais que estão por trás de suas elaborações. O objetivo é verificar como os algoritmos são convocados, recrutados e negociados como parte de esforços coletivos para conhecer e se tornar conhecido.

## Abstract

Algorithms (particularly those embedded in search engines, social media platforms, recommendation systems, and information databases) play an increasingly important role in selecting what information is considered most relevant to us, a crucial feature of our participation in public life. As we have embraced computational tools as our primary media of expression, we are subjecting human discourse and knowledge to the procedural logics that undergird computation. What we need is an interrogation of algorithms as a key feature of our information ecosystem, and of the cultural forms emerging in their shadows, with a close attention to where and in what ways the introduction of algorithms into human knowledge practices may have political ramifications. This essay is a conceptual map to do just that. It proposes a sociological analysis that does not conceive of algorithms as abstract, technical achievements, but suggests how to unpack the warm human and institutional choices that lie behind them, to see how algorithms are called into being by, enlisted as part of, and negotiated around collective efforts to know and be known.

## Palavras-chave

Algoritmos. Plataformas. Mecanismos de busca. Conhecimento. Participação.

## Keywords

Algorithms. Platforms. Search engines. Knowledge. Participation.

## Introdução

Os algoritmos desempenham um papel cada vez mais importante na seleção das informações consideradas de maior relevância para nós, um aspecto fundamental da nossa participação na vida pública. As ferramentas de busca nos ajudam a navegar em grandes bases de dados ou por toda a web. Os algoritmos de recomendação mapeiam nossas preferências em relação a outros usuários, trazendo ao nosso encontro sugestões de fragmentos novos ou esquecidos da cultura. Eles gerenciam as nossas interações em sites de redes sociais, destacando as novidades de um amigo enquanto excluem as novidades de outro.

Os algoritmos projetados para calcular o que "está em alta", o que é "tendência" ou o que é "mais discutido" nos oferecem uma camada superficial das conversas aparentemente sem fim que estão disponíveis. Juntos, eles não só nos ajudam a encontrar informações, mas nos fornecem meios para saber o que há para ser conhecido e como fazê-lo; a participar dos discursos sociais e políticos e de nos familiarizarmos com os públicos dos quais participamos. Além disso, são hoje uma lógica central que controla os fluxos de informação dos quais dependemos, com o "poder de possibilitar e atribuir significados, gerenciando como a informação é percebida pelos usuários, a 'distribuição do sensível'" (LANGLOIS, 2013).

Os algoritmos não são necessariamente *softwares*: em seu sentido mais amplo, são procedimentos codificados que, com base em cálculos específicos, transformam dados em resultados desejados. Os procedimentos dão nome tanto ao problema quanto aos passos pelos quais ele precisa passar para ser resolvido. Podemos considerar como algoritmos, por exemplo, instruções de navegação ou fórmulas matemáticas usadas para prever o movimento de um corpo celestial. "Algoritmos fazem coisas e sua sintaxe incorpora uma estrutura de comando para permitir que isso aconteça" (GO-

FFEY, 2008, p. 17). Podemos pensar, portanto, que os computadores são fundamentalmente máquinas algorítmicas - projetadas para armazenar e ler dados, aplicar procedimentos matemáticos de forma controlada e oferecer novas informações como resultado. Porém tratam-se de procedimentos que poderiam ser feitos manualmente - e, de fato, eram feitos (LIGHT, 1999).

Mas, à medida que adotamos ferramentas computacionais como nossos principais meios de expressão e passamos a fazer não só da matemática mas de toda a informação "digital", passamos a sujeitar o discurso e o conhecimento humano a essas lógicas procedimentais que sustentam toda a computação. E há implicações específicas quando usamos algoritmos para selecionar o que é mais relevante a partir de um *corpus* de dados composto por rastros das nossas atividades, preferências e expressões.

Esses algoritmos, os quais chamaremos de *algoritmos de relevância pública*, estão - através dos mesmos procedimentos matemáticos - produzindo e certificando conhecimento. A avaliação algorítmica da informação, assim, representa uma *lógica de conhecimento* particular baseada em suposições específicas sobre o que é o conhecimento e como alguém deveria identificar seus componentes mais relevantes. O fato de estarmos recorrendo a algoritmos para identificar o que precisamos saber é tão marcante quanto termos recorrido aos especialistas credenciados, ao método científico, ao senso comum ou à palavra de Deus.

O que precisamos é um questionamento dos algoritmos enquanto elemento central do nosso ecossistema informacional (ANDERSON, 2011), e das formas culturais que emergem às suas sombras (STRIPHAS, 2010), prestando atenção especialmente em como e onde a introdução desses algoritmos nas práticas de conhecimento humano podem ter ramificações políticas. Este ensaio é um mapa conceitual exatamente pra isso. Aqui destaco seis dimensões dos algoritmos de relevância pública que têm valor político:

1. *Padrões de inclusão*: as escolhas por trás do que gera um índice, em primeiro lugar; o que é excluído; e como os dados são preparados para o algoritmo.
2. *Ciclos de antecipação*: as implicações das tentativas dos provedores dos algoritmos de conhecer a fundo e prever completamente os seus usuários; e como importam as conclusões às quais eles chegam.
3. *Avaliação de relevância*: os critérios pelos quais os algoritmos determinam o que é relevante; como esses critérios nos são ocultados; e como eles implementam escolhas políticas acerca de um conhecimento considerado apropriado e legítimo.
4. *A promessa da objetividade algorítmica*: a maneira como o caráter técnico do algoritmo é situada como garantia de imparcialidade; e como essa alegação é mantida diante de controvérsias.
5. *Entrelaçamento com a prática*: como os usuários reconfiguram suas práticas para se adequar aos algoritmos dos quais dependem; e como podem transformar algoritmos em espaços de disputa política, às vezes até mesmo para questionar as políticas do próprio algoritmo.
6. *A produção de públicos calculados*: como a apresentação algorítmica dos públicos, para eles mesmos, molda uma noção de si desse público; e quem está em melhor posição para se beneficiar desse conhecimento.

Considerando a rapidez com que mudam estas tecnologias e os usos que são feitos delas, esta lista deve ser vista como provisória e não como completa. Mas a meu ver, acredito que essas sejam as linhas de pesquisa mais importantes para compreender os algoritmos enquanto ferramentas emergentes para o conhecimento e discurso públicos.

Pode ser sedutoramente fácil interpretar tudo isso de forma errada. Na tentativa de dizer algo relevante sobre como os algoritmos transformam o discurso público, devemos resistir firmemente à tentação de tratar a tecnologia como condutora nas explicações. Ainda que recentes estudos sociológicos sobre a internet tenham se dedicado a desfazer um determinismo tecnológico simplista que contaminou trabalhos anteriores, este determinismo continua a ser uma instância analítica atraente. Uma análise sociológica não deve conceber os algoritmos como realizações técnicas abstratas, mas desvendar as escolhas humanas e institucionais que estão por trás desses mecanismos frios. Suspeito que uma abordagem mais frutífera seria nos voltarmos tanto para a sociologia do conhecimento, quanto para a sociologia da tecnologia - para observarmos como essas ferramentas são convocadas, alistadas como parte de, e negociadas em torno de esforços coletivos para conhecer e se tornar conhecido. Isso pode nos ajudar a revelar que algoritmos aparentemente sólidos são, de fato, realizações frágeis. Também deve nos lembrar que os algoritmos são hoje tecnologias de comunicação; assim como as tecnologias de transmissão e de publicação, eles são hoje "instrumentos científicos da sociedade em geral" (GITELMAN, 2006, p. 5) e estão envolvidos e influenciando as maneiras pelas quais ratificamos o conhecimento para a vida cívica, porém de maneiras mais "protocólicas" (GALLOWAY, 2004) - ou organizadas computacionalmente para dizer de outro jeito - do que qualquer outro meio anteriormente.

### **Padrões de inclusão**

Os algoritmos são máquinas inertes e sem sentido até serem combinados com bancos de dados para com eles funcionar. Uma pesquisa sociológica sobre um algoritmo deve sempre levar em consideração os bancos de dados aos quais ele está ligado; não fazê-lo seria o mesmo que estudar o que foi dito em um protesto público, sem

perceber que alguns dos protestantes foram barreados na entrada do parque.

Para os usuários, algoritmos e bancos de dados são conceitualmente conjugados: usuários geralmente os tratam como um único mecanismo em funcionamento. E, aos olhos do mercado, os criadores da base de dados e os provedores do algoritmo são geralmente os mesmos, ou trabalham em consonância econômica e, muitas vezes, ideológica. “Juntos, estruturas de dados e algoritmos são, para um computador, as duas metades da ontologia do mundo” (MANOVICH, 1999, p. 84). No entanto, podemos tratar os dois como analiticamente distintos: antes que os resultados possam ser fornecidos algorítmicamente, a informação deve ser coletada, preparada para o algoritmo e, às vezes, excluída ou rebaixada.

## Coleção

Vivemos em um momento histórico no qual, mais do que nunca, quase todas as atividades públicas incluem o armazenamento de extensos registros, catalogação e arquivamento de documentos - e fazemos isso ainda mais nas redes de comunicação projetadas de tal forma que cada entrada, cada página visualizada e cada clique deixe um rastro digital. Transformar tais rastros em bases de dados envolve um conjunto complexo de técnicas de informação (STALDER, MAYER, 2009): o *Google*, por exemplo, rastreia os sites de indexação da web e seus metadados. Ele digitaliza as informações do mundo real, de acervos de bibliotecas a imagens de satélite ou registros fotográficos de ruas da cidade; convida os usuários a fornecerem seus detalhes pessoais e sociais como parte de seu perfil no *Google+*; mantém registros detalhados de cada pesquisa realizada e cada resultado clicado; adiciona informações com base na localização de cada usuário; armazena os rastros das experiências de navegação na web reunidas via suas redes massivas de publicidade.

Compreender o que está incluído nesses bancos de dados requer atenção às políticas de coleta dos serviços de informação, mas também deveria se estender para além das práticas efetivamente envolvidas. Não só para detectar casos de má conduta, embora existam alguns, mas para entender como um provedor de informações pensa sobre a coleta de dados que empreende. A resistência política ao projeto *StreetView* do *Google*, na Alemanha e na Índia, lembra-nos que a resposta à pergunta “Como é esta esquina?” tem diferentes implicações para pessoas que querem ir até aquela esquina, para pessoas que moram lá e para pessoas que acreditam que essa informação não deveria ser disponibilizada publicamente. Essa resposta também nos revela o que o *Google* pensa sobre o que é “público”, uma interpretação que está sendo amplamente implantado através do serviço prestado pela empresa.

## Preparada para o algoritmo

“Dado bruto é um oxímoro” (GITELMAN, JACKSON, 2013). Os dados já vêm dissecados e persistentemente bagunçados. Contudo, existe uma ordem premeditada necessária para que os algoritmos possam funcionar com esses dados. Mais do que qualquer coisa, os algoritmos são projetados e apreciados para serem funcionalmente automáticos; para, quando acionados, agirem sem qualquer intervenção ou supervisão regular de humanos (WINNER, 1977). Isso significa que as informações dos banco de dados devem ser transformadas e institucionalizadas de tal forma que os algoritmos possam agir sobre elas automaticamente. Os dados devem ser “imaginados e enunciados em contraste com a continuidade dos fenômenos” (GITELMAN, JACKSON, 2013). Reconhecer quais são as formas pelas quais os dados são “limpos” é um importante contraponto para a aparente automaticidade dos algoritmos. Assim como é possível conhecer sobre esculturas, estudando seus moldes invertidos, podemos

entender mais sobre os algoritmos ao analisarmos de perto como as informações são orientadas para enfrentá-los, como elas são *preparadas para o algoritmo*.

Nas primeiras arquiteturas de banco de dados, a informação era organizada em hierarquias rigorosas e, como foi revelado, inflexíveis. Com o desenvolvimento das arquiteturas relacionais e orientadas aos objetos, as informações puderam ser organizadas de maneira mais flexíveis, onde é possível associar os *bits* de dados de várias formas uns com os outros, as categorias podem mudar ao longo do tempo e os dados podem ser explorados sem ser preciso navegar ou mesmo entender a estrutura hierárquica na qual eles estão arquivados. As implicações sociológicas dos designs dos bancos de dados têm sido amplamente ignoradas; mas é preciso ter em mente que os estilos das bases de dados criaram políticas, além de apenas criar ferramentas informacionais essenciais para o funcionamento dos algoritmos. Como Rieder (2012) observa, com a ampla aceitação dos bancos de dados relacionais surge uma “ontologia relacional” que entende os dados como atomizados, “objetos regulares, uniformes e apenas vagamente conectados que podem ser requisitados de maneiras potencialmente ilimitadas quando recuperadas”, deslocando assim o poder expressivo do design estrutural do banco de dados para a busca.

Porém, mesmo com esses bancos de dados mais flexíveis, a categorização continua a ser de vital importância para seu design e seu gerenciamento. O processo de categorização é uma poderosa intervenção semântica e política: quais são as categorias, o que pertence a elas e quem decide como implementá-las na prática, são todas asserções poderosas sobre como as coisas são e como elas devem ser (BOWKER; STAR, 2000). Uma vez instituída, uma categoria desenha uma demarcação que será tratada com reverência por um futuro algoritmo. Como exemplo, trazemos o incidente *#amazonfail*. Em 2009, mais de 157 mil livros *gay-frien-*

*dly*<sup>1</sup> desapareceram instantaneamente das listas de vendas do site da *Amazon* porque foram acidentalmente categorizados como “adultos”. Naturalmente, sistemas informacionais complexos como esses estão propensos a erros. Mas esse erro em particular também nos revelou que o algoritmo da *Amazon*, que calcula a “lista dos livros mais vendidos”, é instruído a ignorar livros categorizados como adultos. Ou seja, mesmo na ausência de erros, qualquer que seja o critério usado pela *Amazon* para determinar o que é considerado “adulto”, ou não, está sendo aplicado e reificado - e se torna aparente apenas na inexplicável ausência de alguns livros da lista e na presença de outros.

## Exclusão e rebaixamento

Embora os produtores de banco de dados compartilhem o apetite pela coleta de informações, eles se distinguem mais pelo que escolhem excluir.

O arquivo, ao lembrar tudo e apenas certo conjunto de fatos / descobertas / observações, consistente e ativamente se envolve no esquecimento de outros conjuntos. Assim, a força impositiva do arquivo opera através da exclusão invisível. A invisibilidade é uma característica importante nesse caso: o arquivo se apresenta como o conjunto de todas as declarações possíveis, ao invés da lei do que pode ser dito (BOWKER, 2006, p. 12-14).

Mesmo nas condições atuais de abundância digital (KEANE, 1999), em que é mais barato e mais fácil errar mantendo as informações, sempre há uma sobra.

Os sites podem, eles próprios, se recusarem a ser indexados pelos coletores de dados (como ferramentas de busca, por exemplo). Elmer (2009) revela que o *robot.txt*, um trecho de código que

1 Nota da tradutora: o termo “gay-friendly” não costuma ser traduzido para português e é usado em referência a lugares ou conteúdos que são receptivos e/ou voltados para os membros da comunidade LGBT.

previne uma página ou site de ser indexado por ferramentas de busca, apesar de ter sido projetado inicialmente como uma ferramenta para preservar a privacidade dos criadores individuais, tem sido usado por instituições governamentais como uma forma alternativa de “reter” documentos públicos da apreciação do público. Mas além da auto-exclusão, algumas informações coletadas inicialmente são subsequentemente removidas antes mesmo que um algoritmo chegue até elas. Apesar de os serviços de informação em larga-escala se gabarem por serem abrangentes, esses sites são, e sempre devem ser, censores de informações também. Dados indexados excluem *spams* e vírus; vigiam violação de direitos autorais e pornografia; e retiram o que é obsceno, condenável ou politicamente controverso das bases de dados (Gillespie no prelo).

Conteúdos ofensivos podem ser simplesmente removidos dos dados indexados, ou uma conta pode ser suspensa, antes mesmo de chegar ao conhecimento de outro usuário. Contudo, junto com o algoritmo, é possível lidar com conteúdos problemáticos de formas mais sutis. O *YouTube* “rebaixa algorítmicamente” vídeos provocantes para que não apareçam nas listas dos mais assistidos ou na página inicial gerada para novos usuários. O *Twitter* não censura conteúdos profanos de tuítes públicos, mas ele os remove da avaliação algorítmica que determina quais termos figuram nos “*Trending*”.

Os padrões específicos pelos quais as informações são excluídas ou incluídas nos bancos de dados, e depois gerenciadas de maneiras particulares, são reminiscências dos debates do século XX (TUSHNET, 2008) sobre as formas em que as decisões feitas pela mídia comercial sobre o que deve ser sistematicamente deixado de fora e quais categorias de fala simplesmente não se qualificam podem moldar a diversidade e o caráter do discurso público. Essas escolhas, sejam feitas por um editor de jornal ou pelos mecanismos de indexação de uma ferramenta de pesquisa, ajudam a estabelecer e a confirmar padrões

de debate viável, legitimidade e decoro. Mas neste caso, os algoritmos podem ser apresentados como automáticos, já que são os padrões de inclusão nos bancos de dados que pré-determinam o que vai ou não aparecer entre seus resultados.

## Ciclos de antecipação

Algoritmos de busca determinam o que oferecer com base nas informações do usuário. Mas a maioria das plataformas hoje faz do seu negócio saber mais, muito mais, sobre o usuário do que apenas a pesquisa que ele acabou de fazer. Os sites esperam ser capazes de antecipar o usuário quando o algoritmo é acionado, o que requer tanto o conhecimento coletado naquele momento, quanto o conhecimento já acumulado sobre o usuário e sobre outros usuários considerados parecidos com ele em termos estatísticos e demográficos (BEER, 2009) - unindo o que Stalder e Mayer (2009) chamam de “segundo índice”. Se podemos dizer que as emissoras de TV e Rádio forneciam não só conteúdo para as audiências, mas também audiências para os anunciantes (SMYTHE, 2001), podemos dizer também que os provedores digitais não apenas fornecem informações para os usuários, mas usuários para seus algoritmos que são feitos e refeitos a cada momento de uso, porque cada clique e cada busca incrementam a ferramenta.

Muitos dos estudos sobre coleta de dados e práticas de rastreamento dos provedores de informação contemporâneos têm focado em questões importantes sobre privacidade provocadas por eles. Zimmer (2008) argumenta que as ferramentas de busca não só pretendem incessantemente indexar a *web* mas também desenvolver uma “coleta perfeita” de tudo relativo aos seus usuários. Para fazer isso, os provedores de informação precisam não só rastreá-los, mas construir infraestruturas técnicas e modelos de negócio que conectem sites individuais em um pacote de serviços (como as várias ferramentas e serviços do *Google*) ou em um ecossiste-

ma ainda mais amplo (como com o “*social graph*” do Facebook e seus botões de “curtir” espalhados ao longo pela web) para então criarem incentivos para que os usuários permaneçam neles. Isto permite que o provedor seja “passivo-agressivo” (BERRY, 2012) na forma como agrega as informações coletadas em vários sites e as transforma em um perfil coerente e cada vez mais detalhado do usuário. Os provedores também se beneficiam do caráter cada vez mais participativo da web, em que usuários são encorajados a oferecer todo tipo de informações sobre si mesmos e a se sentirem empoderados ao fazê-lo. À medida que nossas micro-práticas migram cada vez mais para essas plataformas, torna-se sedutor para os provedores de informação (apesar de não ser compulsório) rastrear essas atividades e transformá-las em mercadoria de várias formas (GILLESPIE; POSTIGO, 2012). Inclusive, os usuários podem não estar cientes de que suas atividades estão sendo rastreadas pelos maiores anunciantes online, em toda a web, e têm pouco ou nenhum meio de contestar este acordo mesmo que o saibam (TUROW, 2012).

Contudo, privacidade não é a única questão politicamente relevante nesse caso. Nestes ciclos de antecipação, os bits de informação são mais fáceis de serem lidos pelo algoritmo, e, portanto, tendem a substituir os usuários. O Facebook sabe muito sobre seus usuários; mas, mesmo assim, só sabe o que é possível saber. As informações mais conhecidas (geolocalização, plataforma computacional, informações de perfil, amigos, atualizações de status, links seguidos no site, tempo no site, atividade em outros sites que hospedam o botão de “curtir” ou cookies) são uma interpretação do usuário, um “dossiê digital” (SOLOVE, 2004) ou uma “identidade algorítmica” (CHENEY-LIPPOLD, 2011) que é imperfeita, mas suficiente. O que é menos legível, ou que não pode ser conhecido sobre os usuários, é deixado de lado ou é aproximado de forma grosseira por esses algoritmos. Como Balka (2011) descreveu, os sistemas de informação produzem “corpos

de sombra” ao enfatizar alguns aspectos dos seus sujeitos e ignorar outros. Esses corpos de sombra persistem e se proliferam através dos sistemas de informação, e as diferenças que separam o usuário antecipado do próprio usuário que é representado podem ser politicamente problemáticas ou politicamente produtivas.

Mas os algoritmos não só fazem previsões exaustivas sobre os usuários; às vezes, eles fazem aproximações suficientes. Talvez tão importante quanto a constante vigilância dos usuários são as conclusões que os provedores querem elaborar com base em relativamente pouca informação sobre eles. *Hunch.com*, um serviço de recomendação de conteúdo, gabou-se por conseguir saber as preferências de um usuário com 80-85 por cento de precisão, com base nas respostas a apenas cinco perguntas. Apesar de reduzir radicalmente a complexidade de uma pessoa a cinco pontos em um gráfico, o que é importante é que esta é uma precisão suficiente para os seus propósitos<sup>2</sup>. Dado que tais sites se sentem confortáveis em oferecer caricaturas dos usuários, as questões que parecem nos classificar de forma suficiente para eles, particularmente em relação às nossas preferências enquanto consumidores, tendem a ganhar relevância como formas de medir o público. E, em certa medida, nós somos convidados a nos formalizar nessas categorias. Diante desses provedores, somos encorajados a escolher nos menus que eles oferecem uma opção para que sejamos corretamente antecipados pelo sistema e para que ele nos ofereça as informações certas, as recomendações certas, as pessoas certas.

Além de conhecer detalhes pessoais e demográficos sobre cada usuário, os provedores de informação realizam uma grande pesquisa para tentar entender, e depois operacionalizar, como

<sup>2</sup> Nota do autor: ZUCKERMAN, Ethan. 26 maio 2011. Eli Pariser talks about the filter bubble. In: The Boston Phoenix. Disponível em: <http://thePhoenix.com/Boston/arts/121405-eli-pariser-talks-about-the-filter-bubble/>. Acesso em: 22 abr. 2013.

os humanos normalmente procuram, se engajam e compreendem as informações. Mais notavelmente no estudo das interações humano-computador (IHC), o entendimento da psicologia e da percepção humana é aplicado ao design dos algoritmos e às formas pelas quais os seus resultados deveriam ser representados aos usuários. Os designers esperam ser capazes de prever as tendências e capacidades psicofisiológicas dos usuários, além de seus hábitos e preferências. Mas, também nessas antecipações, valores implícitos e, muitas vezes, políticos podem estar inscritos na tecnologia: os hábitos de percepção ou interpretação de alguns usuários são tomados como modelo universal, hábitos contemporâneos são imaginados como atemporais, objetivos computacionais particulares são tratados como auto-evidentes.

Também estamos presenciando um novo tipo de poder informacional, reunido nessas enormes bases de dados sobre as atividades e as preferências dos usuários, que está ele mesmo modificando a paisagem política. Além das suas técnicas, os provedores de informação que acumulam esses dados, as indústrias terceirizadas que reúnem e compram dados de usuários como mercadoria, e aqueles que transacionam dados sobre os usuários por quaisquer outros motivos (ou seja, empresas de cartão de crédito) têm por essa razão uma voz mais forte tanto no mercado, quanto nos corredores do poder legislativo, e estão cada vez mais se envolvendo nos debates políticos sobre a proteção do consumidor e direitos digitais. Estamos vendo a implantação da mineração de dados nas arenas da organização política (HOWARD, 2005), no jornalismo (ANDERSON, 2011) e na publicação de conteúdo (STRIPHAS, 2009), onde os segredos extraídos das quantidades massivas de dados dos usuários são tomados como diretrizes convincentes para a produção de conteúdo no futuro - seja a próxima campanha micro-segmentada ou o próximo fenômeno pop.

## A avaliação da relevância

Quando os usuários clicam em “Pesquisar”, carregam seu “Feed de Notícias” do Facebook, ou pedem recomendações do Netflix, os algoritmos precisam identificar instantânea e automaticamente quais, dos trilhões de bits de informação, melhor atendem os critérios em questão, e que vão melhor satisfazer um usuário específico e seus objetivos presumidos. Embora estes cálculos nunca tenham sido simples, eles se tornaram mais complexos à medida que o uso público desses serviços foi amadurecendo. Os algoritmos de pesquisa, por exemplo, eram antes baseados apenas em dizer qual a frequência com que determinados termos de pesquisa apareciam nas páginas indexadas da web. Agora, esses algoritmos incorporam informações contextuais sobre os sites e sobre onde eles estão hospedados; consideram a frequência e como o site é relacionado por outros; e convocam técnicas de processamento de linguagem natural para melhor “entender” tanto a consulta, quanto os recursos que o algoritmo pode oferecer como resposta. De acordo com o Google, seu algoritmo de pesquisa examina mais de duzentos indicadores para cada consulta feita no site<sup>3</sup>.

Esses indicadores são os meios pelos quais o algoritmo estima a “relevância” de um conteúdo. É nessa questão que os sociólogos dos algoritmos devem fincar os pés: “relevante” é um julgamento fluido e carregado de significados, tão aberto à interpretação quanto alguns dos termos avaliativos

3 Nota do autor: GOOGLE. 2013. Facts about Google and competition. In: Google. Disponível em: <http://www.google.com/competition/howgooglesearchworks.html>. Acesso em: 22 abr. 2013.

O Google e o Bing já se envolveram em uma guerra de competição por “indicadores”, primeiro quando o Bing anunciou que usava 1.000 indicadores e o Google seguindo com a informação que seus 200 indicadores tinham até 50 variações, levando seu total a quase 10.000. Veja: SULLIVAN, Danny. 11 nov. 2010. Dear Bing we have 10,000 ranking signals to your 1,000. Love, Google. In: Search Engine Land. Disponível em: <https://searchengineland.com/bing-10000-ranking-signals-google-55473>. Acesso em 22 abr. 2013.

sobre os quais estudiosos da mídia já se debruçaram, como “notável” ou “popular”. Como não há uma métrica independente capaz de dizer quais seriam *realmente* os resultados de pesquisa mais relevantes para qualquer busca, os engenheiros precisam decidir quais resultados parecem “corretos” e ajustar seu algoritmo para atingir esse resultado, ou fazer alterações com base em evidências dos usuários, considerando cliques rápidos e pesquisas sem continuação como uma aproximação, não de relevância exatamente, mas de satisfação com o conteúdo oferecido. Acusar um algoritmo de ser enviesado implica considerar que exista um julgamento imparcial de relevância, o qual a ferramenta estaria falhando em alcançar. Uma vez que tal medida não está disponível, as disputas sobre as avaliações algorítmicas não têm um terreno sólido para se basear.

## Critérios

Para ser possível dizer que um determinado algoritmo faz suposições avaliativas, do tipo que tem consequências para o esforço do conhecimento humano, seria preciso fazer uma análise crítica do algoritmo a fim de questionar seus critérios subjacentes. Porém, em quase todos os casos, esses critérios de avaliação estão ocultos e devem permanecer assim. O algoritmo de “*Trendings*” do *Twitter*, que informa ao usuário sobre os termos mais discutidos naquele momento em uma determinada área geográfica, deixa inclusive a definição do que seria “mais discutido” sem especificação. Os critérios que eles usam para medir o que é “mais discutido” são descritos apenas em termos gerais: a velocidade do surgimento de um determinado termo; se ele já apareceu na lista “*Trendings*” do *Twitter* antes; e se ele circula dentro de um determinado grupo ou se expande através de vários grupos de usuários. O que não é declarado pela empresa é como esses critérios são medidos; como são avaliados em relação uns

aos outros; quais outros critérios também foram incorporados à avaliação; e se algum desses critérios foi substituído. Isso faz com que os algoritmos estejam eternamente abertos à suspeita dos usuários de que seus critérios seriam enviesados para favorecer os provedores comerciais ou politicamente, ou que incorporariam premissas embutidas e não-verificáveis que agem de forma incompreensível, até mesmo para os designers (GILLESPIE, 2012).

Um provedor de informações como o *Twitter* não pode ser muito mais explícito ou preciso quanto ao funcionamento de seus algoritmos. Fazê-lo seria dar meios para que seus concorrentes duplicassem e superassem seu serviço facilmente, além de exigir uma explicação mais técnica do que a maioria dos usuários está preparada para receber. Isso prejudicaria a capacidade de mudança dos critérios conforme necessário, mas acima de tudo, forneceria um roteiro para quem quisesse “burlar o sistema” e colocar seus sites no topo dos resultados de pesquisa ou fazer suas *hashtags* aparecerem na lista de “*Trendings*”. Embora alguns sites de recomendações colaborativas, como o *Reddit*, tenham tornado públicos seus algoritmos de classificação das histórias e comentários dos usuários, eles precisam estar constantemente procurando e corrigindo casos de “desvalorização” organizada pelos usuários, e essas táticas não podem ser tornadas públicas. Com poucas exceções, a tendência é mesmo que os provedores sejam pouco claros<sup>4</sup>.

## Objetivos comerciais

Uma segunda abordagem poderia implicar na consideração cuidadosa dos contextos econômicos e culturais dos quais o algoritmo veio.

4 Nota do autor: De acordo com a já mencionada possibilidade de um algoritmo ser perfeitamente transparente, existe uma gama de escolhas abertas para o desenvolvedor sobre o quão evidente ser. Pode ser simples como decidir ser mais franco na caracterização da ferramenta; fornecer uma explicação para a razão pela qual certos anúncios são oferecidos junto a uma página; ou fornecendo documentações mais criteriosas do site.

Qualquer sistema de conhecimento emerge em meio aos objetivos econômicos e políticos da disponibilização de informações, e será moldado pelos objetivos e estratégias das poderosas instituições que buscam capitalizá-lo (HESMON-DHALGH, 2006). As pressões enfrentadas pelas ferramentas de busca, plataformas de conteúdo e provedores de informações podem moldar sutilmente o design do próprio algoritmo e a apresentação de seus resultados (VAINDHYANATHAN, 2011). À medida que passa a se apresentar como uma lógica de conhecimento legítima, novos esforços comerciais são feitos sob medida para o algoritmo (por exemplo, a otimização dos sites para aparecerem nas ferramentas de busca - em inglês, SEO, *search engine optimization*), reificando escolhas feitas e forçando outras adicionais.

Por exemplo, as primeiras críticas se preocupavam se as ferramentas de busca iriam oferecer propagandas como *links* ou conteúdo em destaque, apresentadas como resultado de cálculos algorítmicos. A rápida e clara rejeição pública a essa manobra demonstrou quão forte é nossa crença nesses algoritmos: os usuários não queriam que o conteúdo que os provedores pagassem para serem exibidos fosse misturado com o conteúdo selecionado algorítmicamente. Mas a preocupação agora é multidimensional: a paisagem do “Feed de Notícias” do Facebook, por exemplo, não pode mais ser descrita como dois territórios distintos, um social e um comercial. Em vez disso, entrelaça os resultados dos cálculos algorítmicos (quais atualizações de status e outras atividades dos seus amigos devem ser listadas no Feed, quais *links* serão recomendados para esse usuário, quais amigos estão ativos no site naquele momento), elementos estruturais (ferramentas para contribuir com uma atualização de status ou para comentar em um elemento informativo, *links* para grupos e páginas) e elementos colocados ali com base em relações de patrocínio (*banners* publicitários, aplicativos de sites de terceiros). Para mapear este terreno complexo, é necessário compreender

a fundo as relações econômicas e os pressupostos sociais que ele representa.

## Premissas epistemológicas

Por fim, devemos considerar se os critérios avaliativos do algoritmo são estruturados por princípios políticos ou organizacionais específicos que têm eles mesmos ramificações políticas. Não é apenas se um algoritmo pode ser parcial para este ou aquele fornecedor, ou pode favorecer seus próprios interesses comerciais sobre outros. É uma questão sobre a importância das premissas filosóficas sobre o conhecimento relevante nas quais o algoritmo se baseia. Alguns estudos que analisaram vieses na seleção das ferramentas de busca (em ordem de publicação, INTRONA; NISSENBAUM, 2000, HALVAIS, 2008, ROGERS, 2009, GRANKA, 2010) observaram algumas tendências estruturais em retornar como resultado o que já era popular, sites de língua inglesa e fornecedores de informações comerciais. Ao debater sobre o que significaria exigir neutralidade nos resultados de pesquisa, os estudiosos jurídicos (GRIMMELMANN, 2010, PASQUALE; BRACHA, 2008) se referiam a muito mais do que oferecer dicas sobre parceiros comerciais nos resultados. Os critérios que os algoritmos de informação pública levam em consideração são miríades; cada um está equipado com um patamar que define o que será empurrado para os resultados, como um resultado será posicionado acima de outro e assim por diante.

Dessa forma, as avaliações realizadas pelos algoritmos dependem sempre de pressupostos inscritos nos seus códigos que versam sobre o que e como importa ser identificado por eles. Quando uma simples ferramenta de pesquisa conta o número de vezes que um termo pesquisado aparece nas páginas da web indexadas, ela está reificando uma lógica particular que assume que as páginas que incluem o termo pesquisado provavelmente são relevantes para alguém interessado neste termo. Quando o Google desenvolveu o

*PageRank*, que determina o valor de uma determinada página com base no número de *links* em referência a essa página, ele construiu uma lógica diferente: uma página linkada em vários lugares na *web*, em sites considerados de alta qualidade, é vista como sendo “ratificada” por outros usuários e, portanto, têm maior probabilidade de ser relevante para este usuário também. Para Finkelstein, ao preferir *links* advindos de sites vistos como sendo de alta qualidade, o *Google* mudou de uma abordagem mais populista para uma “democracia de acionistas”.

Um *link* não é um voto, mas tem influência proporcional ao poder relativo do eleitor (em termos de popularidade). Como os blocos de interesses comuns, ou facções sociais, podem afetar os resultados de uma pesquisa em um grau que depende do seu peso relativo na rede, os resultados do cálculo algorítmico das ferramentas de busca refletem lutas políticas da sociedade (FINKELSTEIN, 2008, p.107).

Quando um site de discussão de notícias decide qual a proporção de queixas negativas para o número de visualizações é suficiente para justificar e esconder automaticamente um tópico de comentários, essa proporção representa a análise do *site* sobre a volatilidade do discurso público, ou pelo menos a volatilidade que o *site* prefere para a comunidade de usuários que eles acham que têm (BRAUN, 2011). Muita experiência e julgamentos podem ser incorporados a esses artefatos cognitivos (HUTCHINS 1995, LATOUR, 1987), mas são julgamentos que são logo submergidos e automatizados.

A maioria dos usuários não se debruça sobre os critérios algorítmicos e tende a tratá-los como ferramentas não problemáticas a serviço de uma atividade maior: encontrar uma resposta, resolver um problema, entreter. No entanto, embora a tecnologia possa ser igualmente “encaixapretada” (LATOUR 1987, PINCH; BIJKER, 1984) por designers e usuários, isso não deveria nos levar a crer que ela permaneça estável. Na

verdade, os algoritmos podem ser facilmente, instantaneamente, radicalmente e invisivelmente alterados. Embora grandes atualizações possam acontecer apenas ocasionalmente, os algoritmos estão regularmente sendo “ajustados”. As mudanças podem ocorrer sem que a interface com o algoritmo mude minimamente: o “Feed de Notícias” e a barra de pesquisa podem parecer os mesmos de ontem, ainda que as avaliações que acontecem por baixo deles tenham sido completamente refeitas. A metáfora da caixa-preta não nos ajuda nesse caso, já que o funcionamento do algoritmo é obscuro e maleável, “provavelmente tão dinâmico que uma fotografia instantânea nos daria poucas chances de acessar seus vieses” (PASQUALE, 2009). Na verdade, o que nós nos referimos enquanto um algoritmo muitas vezes não é apenas um, mas são muitos algoritmos. Ferramentas de busca como o *Google* participam regularmente de testes “A/B”<sup>5</sup>, apresentando *rankings* diferentes para diferentes subconjuntos de usuários para obter dados sobre a velocidade e satisfação dos clientes em tempo real. As preferências encontradas nos resultados do teste são, então, incorporadas aos ajustes da ferramenta em uma atualização posterior.

Cada algoritmo tem como premissa tanto um pressuposto sobre a análise adequada de relevância, quanto em uma representação desse pressuposto em uma técnica de avaliação (computacional). Podem existir premissas implícitas incorporadas na ideia de relevância de um site, podem existir atalhos incorporados em sua representação técnica dessa idéia, e podem existir fricções entre elas duas.

## A promessa da objetividade algorítmica

Mais do que meras ferramentas, os algoritmos também são estabilizadores da confiança, garantias práticas e simbólicas de que suas avaliações são justas e precisas, livres de subjetividade, erro

5 Nota do autor: CHRISTIAN, Brian. 25 abr. 2012. The A/B Test: Inside the technology that's changing the rules of business. In: *Wired.com*. Disponível em: [https://www.wired.com/2012/04/ff\\_abtesting/](https://www.wired.com/2012/04/ff_abtesting/). Acesso em: 22 abr. 2013.

ou tentativas de influência. Mas, embora os algoritmos possam parecer automáticos e intocáveis pelas intervenções de seus provedores, esta é uma ideia cuidadosamente elaborada. “As ferramentas de busca se gabam por serem automatizadas, exceto quando não o são” (GRIMMELMANN, 2008, p. 950). Na verdade, nenhum serviço de informações pode ser completamente isento de interferência humana ao entregar informações: embora um algoritmo possa avaliar qualquer site como o mais relevante para sua busca, esse resultado não aparecerá se for pornografia infantil, não aparecerá na China se for um discurso político dissidente, e não aparecerá na França se promover o nazismo. No entanto, é muito importante para os provedores desses algoritmos que eles pareçam ser isentos de interferência humana. A legitimidade desses mecanismos deve ser performada junto à disponibilização da própria informação.

As articulações oferecidas pelo provedor do algoritmo junto à ferramenta servem para proporcionar o que Pfaffenberger (1992) chama de “controle logonômico”; ou seja, para definir a ferramenta entre as práticas dos usuários e concedê-la uma legitimidade que é levada às informações fornecidas e, de quebra, ao provedor. A cuidadosa articulação de um algoritmo como imparcial (mesmo quando essa caracterização é mais ofuscante do que a explicativa) o certifica como um ator sociotécnico confiável concede relevância e credibilidade aos seus resultados e mantém a aparente neutralidade do provedor em face as milhões de avaliações que faz. Esta articulação do algoritmo é tão crucial para a sua manutenção social, quanto seu design material e suas obrigações econômicas.

Em grande parte, cabe ao provedor descrever seu algoritmo como tendo uma forma particular; tendo, portanto, um certo conjunto de valores; e, assim, conferindo-lhe um certo tipo de legitimidade. Isso inclui caracterizar cuidadosamente a ferramenta e o seu valor para várias audiências, às vezes de várias formas diferentes: um algoritmo pode ser

defendido como uma ferramenta de avaliação imparcial para os críticos dos seus resultados e, ao mesmo tempo, ser prometido como uma ferramenta para a promoção seletiva e para anunciantes em potencial (GILLESPIE, 2010). Como observa Mackenzie (2005), esse processo requer mais que uma descrição única e enfática: depende da “repetição e citação” (*idem*, 2005, p. 81) e, ao mesmo tempo, requer que se encubra o “conjunto autoritário de práticas” (*idem*, 2005, p. 82) que dá força ao algoritmo. Quando um provedor de informações se vê criticado pelos resultados que fornece, a legitimidade do seu algoritmo deve ser reparada tanto discursiva, quanto tecnicamente. E os usuários são cúmplices nisso:

uma sociedade obcecada pelos principais resultados do Google Notícias fez com que esses resultados se tornassem importantes, e seria insensato supor o contrário (que as pessoas estão obcecadas com os resultados porque eles são importantes) sem uma descrição do porquê o algoritmo é superior a, por exemplo, ‘o julgamento sobre notícias’ dos editores da mídia tradicional. (PASQUALE, 2009).

Esta articulação se dá primeiro na apresentação da ferramenta, na sua implantação dentro de um serviço de informação mais amplo. Chamá-los de “resultados”, “melhores”, “principais histórias” ou “tendências” diz não só sobre o que o algoritmo está medindo, mas também sobre o que deve ser entendido como medida. Uma parte igualmente importante deste trabalho discursivo está na descrição de como o algoritmo funciona. O que pode parecer uma explicação clara sobre que acontece por trás dos bastidores pode não ser realmente uma visada do verdadeiro processo que ali se dá, mas de “bastidores performados” (HILGARTNER, 2000) e cuidadosamente elaborados a fim de legitimar ainda mais esse processo e seus resultados. A descrição do sistema *PageRank* do Google, o elemento mais antigo do seu complexo algoritmo de busca, foi publicada primeiro em um artigo científico (já sendo essa

uma entrega elaborada sobre seus funcionamentos matemáticos), mas foi mitificada em seguida - como a característica definidora da ferramenta, como o elemento central que fez o *Google* se tornar destaque frente a seus concorrentes e como uma lógica computacional fundamentalmente democrática - mesmo quando o algoritmo estava sendo redesenhado para considerar centenas de outros critérios no processo de busca.

Acima de qualquer coisa, os provedores dos algoritmos de informações devem assegurar que seus algoritmos são imparciais. A performance da *objetividade algorítmica* tornou-se fundamental para a manutenção dessas ferramentas como mediadoras legítimas do conhecimento relevante. Nenhum provedor tem sido mais inflexível sobre a neutralidade do seu algoritmo quanto o *Google*, que responde regularmente às solicitações de alteração nos resultados de busca reivindicando que o algoritmo não deve ser adulterado. A empresa celebroumente se retirou completamente do mercado chinês, em 2010, ao invés de censurar seus resultados, embora tenha consentido com as regras da China antes, e pode ter se retirado para não ter de admitir que estava perdendo para as concorrentes chinesas. Mesmo apesar dessa postura, o *Google* alterou seus resultados quando surgiram reclamações sobre uma imagem de Michelle Obama, editada de forma racista, aparecendo no topo dos resultados de busca de Imagens. A empresa também fornece um mecanismo de "busca segura" para não exibir palavrões e imagens sexuais para menores, e se recusa a auto-completar pesquisas que indiquem serviços de troca de arquivos em formato torrent. No entanto, o *Google* afirma regularmente que não altera seus índices, nem manipula seus resultados. Morozov (2011) acredita que esta é uma maneira de desviar a responsabilidade:

a suspensão espiritual do *Google* para a 'neutralidade algorítmica' trai o crescente desconforto da empresa em ser o gatekeeper de informações

mais importante do mundo. Seus fundadores preferem tratar a tecnologia como uma força autônoma e totalmente objetiva ao invés de passar noites sem dormir, preocupando-se com vieses inerentes em como seus sistemas operam - sistemas estes que cresceram de forma tão complexa que nenhum engenheiro no *Google* os entende completamente. (MOROZOV, 2011)

Essa afirmação da objetividade algorítmica desempenha, em muitos aspectos, um papel equivalente à norma de objetividade no jornalismo ocidental. Assim como os motores de busca, os jornalistas desenvolveram táticas para determinar o que é mais relevante, como reportá-lo e como garantir sua relevância - um conjunto de práticas que são relativamente invisíveis para sua audiência; uma meta que eles admitem ser mais difícil de seguir do que parece; e um princípio que ajuda a deixar julgamentos de valor e políticas pessoais de lado, mas não os elimina. Essas práticas institucionalizadas são movidas por uma promessa conceitual que, no discurso do jornalismo, é regularmente expressa (ou exagerada) como uma espécie de totem. Os jornalistas usam a norma de objetividade como um "ritual estratégico" (TUCHMAN, 1972), para dar legitimidade pública às táticas de produção do conhecimento que são inerentemente precárias. "Estabelecer jurisdição sobre a capacidade de analisar objetivamente a realidade é uma reivindicação de um tipo especial de autoridade" (SCHUDSON; ANDERSON, 2009, p. 96).

Objetividades jornalísticas e algorítmicas não são a mesma, de forma alguma. A objetividade jornalística depende de uma promessa institucional de adequada diligência, que é incorporada e transmitida através de um conjunto de normas que os jornalistas aprendem em treinamento e no trabalho. As escolhas feitas pelos jornalistas representam uma competência apoiada em um compromisso profundamente infundido, filosófico e profissional para deixar de lado seus próprios preconceitos e convicções políticas. Já a promessa do algoritmo

se baseia muito menos em normas institucionais e competências adquiridas, e mais em uma promessa tecnicamente influenciada de neutralidade mecânica. Quaisquer que sejam as escolhas feitas, elas são apresentadas como sendo tanto livres de intervenção das mãos humanas, quanto submersas no frio funcionamento da máquina.

Mas em ambas, a legitimidade depende da acumulação de diretrizes para a *proceduralização* da seleção de informações. Os discursos e práticas de objetividade passaram a servir como uma regra constitutiva do jornalismo (RYFE, 2006). A objetividade faz parte do entendimento dos jornalistas sobre si próprios e sobre o que significa ser um jornalista. Participa da avaliação do seu trabalho, por editores, colegas e leitores, e é um sinal determinante do que os próprios jornalistas reconhecem como conteúdo jornalístico. A promessa da objetividade algorítmica também foi nitidamente incorporada nas práticas de trabalho dos provedores de algoritmos, definindo o que se constitui enquanto função e propósito de um serviço de informações. Quando o *Google* inclui no seu manifesto de “Dez coisas que sabemos ser verdade” que “Nossos usuários confiam em nossa objetividade e nenhum ganho de curto prazo poderia justificar a violação dessa confiança”, não se trata nem de artimanha, nem de um “*Maria-vai-com-as-outras*”<sup>6</sup> corporativo. É uma compreensão profundamente enraizada do caráter público do serviço de informação do *Google* que influencia e legitima muitos dos seus projetos técnicos e comerciais, e ajuda a obscurecer a confusa realidade sobre o serviço que a empresa presta.

Ainda assim, essas reivindicações devem competir com outras articulações no diálogo público que podem ser favoráveis, ou não, aos objetivos econômicos e ideológicos das partes interessadas. Bijker

<sup>6</sup> Nota da tradutora: a expressão usada pelo autor é “*corporate Kool-Aid*”. Em inglês, a expressão “*Beber o Kool-Aid*” é usada em referência a alguém que segue uma ideia condenada ou perigosa por pressão dos colegas. Assim, escolhemos traduzi-la como “*Maria-vai-com-as-outras*” que tem um significado parecido em português.

(1997) chama de “quadros tecnológicos” concorrentes essas caracterizações discursivas de uma tecnologia feitas por grupos de atores que também têm interesse na operação, no significado e no valor social dessa tecnologia. O que os usuários de um algoritmo de informação o consideram, sejam eles astutos ou ignorantes, é importante. A forma como a imprensa retrata tais ferramentas vai fortalecer ou prejudicar os cuidadosos esforços discursivos dos seus fornecedores. Isso significa que, embora o algoritmo pareça possuir em si uma aura de neutralidade tecnológica, ou incorporar ideais populistas e meritocráticos, a sua aparência não depende apenas do seu *design*, mas também das realidades mundanas dos ciclos de notícias, assessorias de imprensa, *blogs* de tecnologia, discussões de fãs, revoltas de usuários e maquinações dos concorrentes do seu provedor.

Existe um paradoxo fundamental na articulação dos algoritmos. A objetividade algorítmica é uma declaração importante para um provedor, particularmente de algoritmos que fornecem informações vitais e voláteis para o consumo público. Articular o algoritmo como uma intervenção claramente técnica ajuda um provedor de informações a responder acusações de viés, erro e manipulação dos resultados. Ao mesmo tempo, como pode ser visto com o *PageRank* do *Google*, existe um valor sociopolítico no destaque do populismo dos critérios que o algoritmo usa. Afirmar que um algoritmo de um determinado site é uma representação democrática da opinião coletiva em toda a *web* lhe confere autoridade. E há um valor comercial na alegação de que o algoritmo retorna resultados “melhores” do que os dos seus concorrentes, o que coloca a satisfação do consumidor acima da noção de exatidão (VAN COUVERING, 2007). Ao examinar a articulação de um algoritmo, devemos prestar atenção especial à forma como esta tensão, entre neutralidade tecnicamente assegurada e o sabor social da avaliação que está sendo feita, é gerenciada - e, às vezes, como ela se rompe.

## Entrelaçamento com a prática

Apesar de poderem ser estudados como ferramentas computacionais abstratas, os algoritmos são construídos para serem incorporados às práticas do mundo, que produzem as informações que eles processam, e ao mundo vivido pelos seus usuários (COULDRY, 2012). Isso se dá principalmente quando o algoritmo é o instrumento de um negócio que tem como produto a informação que entrega (ou as propagandas a que se associa). Se os usuários falharem ou se recusarem a se encaixar nas suas práticas e a conceder significado a elas, o algoritmo irá falhar. Isso significa que não devemos considerar seus “efeitos” nas pessoas, mas um “entrelaçamento” multidimensional entre algoritmos postos em prática e as táticas dos usuários que fazem face a eles. Claro, essa relação é um alvo em movimento, porque os algoritmos mudam, assim como as populações de usuários e as atividades com as quais se deparam. Ainda assim, isso não deveria significar que não exista relação entre eles. À medida que esses algoritmos se abrigam na vida das pessoas e nas suas rotinas informacionais, os usuários moldam e rearticulam os algoritmos com os quais se deparam. Os algoritmos também afetam a maneira que as pessoas procuram informações, como elas percebem e pensam sobre os horizontes de conhecimento, e como elas se compreendem no e pelo discurso público.

É importante conceber este entrelaçamento não como uma influência unidirecional, mas como um ciclo recursivo entre os cálculos do algoritmo e os “cálculos” das pessoas. O algoritmo que ajuda os usuários a navegarem pelas fotos do *Flickr* é construído no arquivo de fotos postadas, o que significa que ele é projetado para compreender e refletir as escolhas feitas pelos fotógrafos. O que as pessoas fazem e não fotografam já é um tipo de cálculo avaliativo, embora seja histórico, multivalente, contingente e sociologicamente fundamentado. Mas estes não eram os únicos impulsionadores do de-

*sign* do *Flickr*; a sensibilidade para as práticas fotográficas teve que competir com custos, eficiência técnica, obrigações legais e imperativos comerciais. Além disso, a população dos usuários do *Flickr* e os tipos de fotos que publicava mudaram à medida que o site cresceu em popularidade; quando foi forçado a competir com o *Facebook*; quando introduziu preços diferenciados; quando foi comprado pelo *Yahoo*, e assim por diante.

Vários usuários do *Flickr* postam fotos com o propósito claro de que elas sejam vistas: alguns são fotógrafos profissionais procurando emprego, alguns são pessoas procurando comunidades de amadores semelhantes a elas, outros estão simplesmente orgulhosos do seu trabalho. Assim, tal como o algoritmo precisa ser sensível aos fotógrafos, eles têm interesse em serem sensíveis para o algoritmo, sabendo que aparecer como resultado das buscas certas pode colocar suas fotografias diante das pessoas certas. Assim como a ênfase de *Hollywood* em gêneros específicos convida os roteiristas a escrever de maneira genérica<sup>7</sup>, o algoritmo do *Flickr* pode induzir reorientações sutis nas práticas dos fotógrafos em direção à sua lógica construída, ou seja, induzir os fotógrafos a buscar fotografar de modo a aderir a certas categorias emergentes ou a orientar sua escolha de tema e composição em direção a coisas que o algoritmo parece privilegiar. “Deixamos rastros não sobre a forma como fomos, mas sobre uma negociação tácita entre nós e os nossos auditores imaginados” (BOWKER, 2006, p. 6-7).

## Algoritmicamente identificável

Essa negociação tácita consiste, em primeiro lugar, na reorientação cotidiana e estratégica das práticas que muitos usuários realizam, perante uma ferramenta que eles sabem que pode ampliar seus esforços. Há, para os produtores de informações, um estímulo poderoso e compreensível para

<sup>7</sup> Nota do autor: Christian Sandvig, comunicação pessoal.

tornar seu conteúdo, e eles próprios, identificáveis para um algoritmo. Toda uma indústria, de otimização para mecanismos de busca (SEO - *Search Engine Optimization*, em inglês), promete impulsionar os sites para o topo dos resultados de pesquisa. Mas podemos pensar na otimização (deliberada, profissional) como apenas a linha de frente de um processo muito mais variado, orgânico e complexo pelo qual produtores de conteúdo de todos os tipos se orientam perante os algoritmos. Quando usamos *hashtags* em nossos *tuítes* - uma inovação criada pelos usuários e adotada posteriormente pelo *Twitter* - não estamos apenas nos juntando a uma conversa ou esperando ser lidos por outros, estamos redesenhando nossa manifestação para ser melhor reconhecida e distribuída pelo algoritmo de busca do *Twitter*. Alguns usuários podem trabalhar para serem percebidos pelo algoritmo: sabe-se que adolescentes marcam suas atualizações de *status* com nomes de marcas conhecidas, não relacionadas ao assunto, esperando que o *Facebook* privilegie essas atualizações nos "Feeds de Notícias" dos seus amigos<sup>8</sup>. Outros podem trabalhar para fugir de um algoritmo: sabemos que usuários de *Napster* e *P2P*, que compartilham músicas protegidas por direitos autorais, escrevem os nomes dos artistas com pequenos erros ortográficos para que os usuários consigam encontrar as músicas da "Britny Speers", mas o *software* da indústria musical não<sup>9</sup>.

Isso é burlar o sistema? Ou é uma maneira fundamental pela qual, até certo ponto, orientamo-nos perante os meios de distribuição através dos quais esperamos falar? Com base nos critérios do algoritmo em questão (ou de acordo com nossa melhor estimativa do seu funcionamento), nós já nos tornamos *algoritmicamente identificáveis* de várias formas. Isso não é tão diferente do que assessores de imprensa que orientam seus esforços para encaixar

seus clientes nas rotinas da indústria de notícias: programar um comunicado de imprensa para sair na transmissão da noite, ou fornecer gravações de vídeo para uma emissora ávida por imagens emocionantes, são técnicas para lidar com uma mídia que pode ampliar seus esforços. Hoje, para todos nós, as redes sociais e a *web* oferecem uma espécie análoga de "visibilidade mediada" (THOMPSON, 2005, p. 49), e nós nos beneficiamos de forma semelhante quando lidamos com esses algoritmos.

### Backstage Access - acesso aos bastidores

Mas quem está em melhor posição de compreender e operar os algoritmos públicos que tanto importam para a circulação pública de conhecimento? A visão sobre como funcionam os algoritmos de informação é uma forma de poder: vital para a participação no discurso público, essencial para ganhar visibilidade online, constituidor de credibilidade e das oportunidades que resultam dela. Como mencionamos anteriormente, os critérios e os códigos dos algoritmos geralmente são obscuros - mas não da mesma forma para todos. Para a maioria dos usuários, o conhecimento sobre os algoritmos é vago, simplista e, às vezes, equivocado; eles podem tentar empurrar o algoritmo tanto de maneiras que simplesmente consideram melhores (*hashtags*, *metadados*) ou que fundamentalmente equivocadas em relação aos critérios do algoritmo (como retuitar várias vezes a mesma mensagem na esperança de figurar entre os "*Trendings*" do *Twitter*). Os profissionais de *SEO* e os *spammers* também têm pouco acesso, mas desenvolveram habilidades técnicas para deduzir os critérios do algoritmo a partir de testes e engenharia reversa. Comunidades de entusiastas em tecnologia e críticos se engajam em tentativas semelhantes de descobrir o funcionamento desses sistemas, seja por diversão, conhecimento, vantagem pessoal ou determinadas interrupções. Os legisladores, que estão apenas começando a questionar as implicações dos algoritmos para um

8 Nota do autor: danah boyd, comunicação pessoal.

9 Nota do autor: ABC News. 5 mar. 2001. Napster faced with big list, trick names. In: ABC News. Disponível em: <http://abcnews.go.com/Entertainment/story?id=108389>. Acesso em 22 abr. 2013.

comércio justo ou para o discurso político, receberam até agora apenas as explicações mais gerais possíveis: os provedores de informação afirmam que seus algoritmos são segredos comerciais que não devem ser divulgados em espaços públicos.

Além disso, de fato algumas das partes interessadas recebem acesso garantido ao algoritmo, ainda que sob condições controladas. Os anunciantes recebem um tipo de acesso aos bastidores de funcionamento para saber como fazer o melhor lance<sup>10</sup>. Os provedores de informação que oferecem Interfaces de Programação de Aplicativos (APIs - *Application Programming Interfaces*, em inglês), para seus parceiros comerciais e desenvolvedores, oferecem um vislumbre sob o capô, mas os vincula com contratos e acordos de não-divulgação no mesmo instante. O acesso, a compreensão e os direitos sobre os algoritmos que desempenham um papel crucial no discurso e no conhecimento públicos tendem a mudar entre as diferentes partes interessadas e em circunstâncias específicas - mudando também o poder disponível para essas partes, e para aqueles que elas representam, de construir para os algoritmos, navegá-los e regulá-los.

## Domesticação

Por mais que essas ferramentas nos incitem a nos tornarmos legíveis para elas, nós também as inserimos nas nossas práticas, alterando seu sentido e até mesmo seu design, algumas vezes. Silverstone (1994) sugere que à medida que as tecnologias são oferecidas ao público, elas passam por um processo de "domesticação": essas tecnologias entram em nossas casas literalmente, mas também de forma figurada - os usuários as transformam em propriedade, incorporando-as em suas rotinas e imbuindo-as com significados adicionais que os provedores não puderam antecipar. Algoritmos de informação

10 Nota da tradutora: o sistema de anúncios pagos nas plataformas segue o mesmo modelo de um leilão - há um preço inicial, mas esse valor vai sendo ajustado de acordo com os outros "competidores" pelo mesmo público.

pública certamente são importantes para o modo como os usuários encontram informações, comunicam-se uns com os outros e percebem o mundo ao seu redor. Mas, mais do que isso, os usuários expressam preferências por suas ferramentas de busca favoritas, opinam sobre as recomendações de um site como sendo incômodas, intuitivas ou precisas. Alguns usuários confiam em uma determinada ferramenta, enquanto outros desconfiam dela, usando-a com cuidado ou nem a utilizando de todo. Os usuários do *iPhone* da *Apple* trocam dicas entre si sobre como fazer a *Siri*, a assistente virtual da empresa, a falar seu repertório de réplicas<sup>11</sup> e depois compartilham a indignação com as respostas dadas a questões políticas polêmicas<sup>12</sup>. Usuários satisfeitos com o *Facebook* tornam-se críticos amanhã quando o algoritmo do "Feed de Notícias" é alterado, aparentemente por motivações econômicas - enquanto continuam a postar atualizações de status durante e depois da revolta. Os usuários, perante as assimetrias de poder da coleta de dados e vigilância online, desenvolveram uma série de táticas de "ofuscação" para escapar ou contaminar as tentativas dos algoritmos de conhecê-los (BRUNTON; NISSENBAUM, 2011). Embora seja crucial considerarmos as formas pelas quais as ferramentas algorítmicas moldam nosso contato com as informações, não devemos insinuar que os usuários estejam sob a influência dessas ferramentas. A realidade é mais complicada e mais particular.

Os usuários também podem recorrer a esses algoritmos, refletirem sobre si com base nos seus dados; muitos sites permitem nos apresentarmos para os outros e para nós mesmos, incluindo o nosso perfil público, o desempenho das nossas amizades, a expressão das nossas preferências ou um re-

11 Nota do autor: <http://siri-sayings.tumblr.com/>. Acesso em: 22 abr. 2013.

12 Nota do autor: WORTHAM, Jenna. 30 nov. 2011. Apple says Siri's abortion answers are a glitch. In: *The New York Times*. Disponível em: <http://bits.blogs.nytimes.com/2011/11/30/apple-says-siris-abortion-answers-are-a-glitch/>. Acesso em: 22 abr. 2013.

gistro das nossas atividades recentes. O recurso da linha do tempo do *Facebook* organiza as atividades dos usuários como lembranças cronológicas sobre si mesmos; o prazer de ver o que é selecionado algoritmicamente oferece uma espécie de encantamento que vai além da composição das fotos e das postagens em primeiro lugar. Mas os algoritmos também podem funcionar como uma “tecnologia do *self*” (FOUCAULT, 1988), particularmente convincente, quando parecem ratificar, independentemente, a visibilidade pública de uma determinada pessoa. É comum, hoje, procurar a si mesmo no *Google*: ver-me aparecer como o melhor resultado da busca pelo meu nome, oferece uma espécie de garantia da minha tênue existência pública. Existe uma sensação de validação quando um tópico sobre seu animal de estimação aparece no “*Trendings*” e no *Twitter*; quando a *Amazon* lhe recomenda um livro que você ama ou quando a função “*Genius*” do *iTunes* da *Apple* compõe uma lista de reprodução interessante a partir da sua biblioteca de músicas. Seja quando nós efetivamente adaptamos nossas compras da *Amazon* para parecermos versados (assim como as famílias pesquisadas pela *Nielsen* exageravam ao informar quanto assistiam a *PBS* e *C-Span*)<sup>13</sup> ou se simplesmente gostamos quando o algoritmo confirma nosso senso de identidade, os algoritmos são um convite poderoso para nos entendermos através de lentes independentes que eles prometem fornecer.

Os algoritmos não são apenas o que seus *designers* fazem deles, ou o que eles fazem a partir da informação que processam. São também o que nós fazemos deles dia após dia - mas com uma ressalva: como a lógica, a manutenção e o *redesign* desses algoritmos permanecem nas mãos dos provedores de informação, eles claramente estão numa posição privilegiada para reescrever nosso entendimento

13 Nota da tradutora: o autor fala sobre quando os usuários ajustam seus algoritmos para que eles pareçam “mais inteligentes” a partir das recomendações oferecidas. O exemplo refere-se às famílias pesquisadas para avaliação da *Nielsen* que afirmavam assistir a canais considerados “mais cultos” ou “mais sérios”.

sobre eles, ou para gerar uma longa incerteza sobre os critérios que dificultam tratarmos os algoritmos como sendo verdadeiramente nossos.

## Lógicas de conhecimento

Teorizar é fácil, mas é substancialmente mais difícil documentar como os usuários podem mudar suas visões de mundo para acomodar as lógicas subjacentes e os pressupostos implícitos dos algoritmos que usam regularmente. Há uma razão para que as lógicas de funcionamento desses algoritmos não apenas moldem as práticas dos usuários, mas também os levem a internalizar suas normas e prioridades: Bucher (2012) argumenta que o algoritmo *EdgeRank*, usado pelo *Facebook* para determinar quais atualizações de *status* são prioritariamente exibidas no “*Feed de Notícias*” de um usuário, encoraja uma “subjetividade participativa” nos usuários, que reconhecem que gestos de afinidade (como comentar a foto de um amigo) são um critério central para esse algoritmo. Longford (2005) argumenta que o código da plataforma comercial nos “habituata”, através de incessantes solicitações e configurações padrão cuidadosamente projetadas, a dar mais informações pessoais sobre nós mesmos. Mager (2012) e Van Couvering (2010) propõem que os princípios do capitalismo estão incorporados no funcionamento das ferramentas de busca.

Mas, nós não precisamos recorrer a tais teorias de dominação ideológica para sugerir que os algoritmos, projetados para oferecer conhecimentos relevantes, também oferecem formas de conhecer - e que, à medida que se tornam mais comuns e confiáveis, suas lógicas se tornam auto-afirmativas. A ferramenta de pesquisa do *Google*, entre suas 200 indicações, presume que um conhecimento relevante é assegurado em grande parte pela ratificação pública e está ajustada para dar grande peso às opiniões dos que já são autenticados publicamente. Esta mistura da sabedoria das massas e das autoridades coletivamente certificadas é a solução encontrada

pelo *Google* para a antiga tensão entre o conhecimento e o senso-comum, no permanente problema de “como conhecer”. Não é sem precedentes, e não é uma maneira fundamentalmente errada de conhecer, mas é específica, com suas próprias ênfases e miopias. Hoje, a solução do *Google* é operacionalizada em uma ferramenta que bilhões de pessoas usam todos os dias, a maioria das quais a experimenta simplesmente como algo que “funciona” de forma simples e sem problemas. Até certo ponto, o *Google* e seu algoritmo ajudam a afirmar e a normalizar essa lógica de conhecimento como “correta”, tão correta quanto seus resultados parecem ser.

### **A produção de públicos calculados**

Ito, Boyd e outros autores recentemente introduziram o termo “públicos em rede” (BOYD, 2010; ITO, 2008; VARNELIS, 2008) para destacar as comunidades de usuários que podem se unir através das mídias sociais e a forma como as tecnologias estruturam como esses públicos podem se formar, interagir e, às vezes, se desfazer. “Enquanto os públicos em rede têm muito em comum com outros tipos de públicos, a forma como a tecnologia os estrutura introduz affordances distintas que moldam como as pessoas se engajam com esses ambientes” (BOYD, 2010, p.39). Visto que os algoritmos são um componente tecnológico central nesses ambientes mediados, eles também ajudam a estruturar os públicos que podem emergir usando a tecnologia digital.

Algumas questões têm sido levantadas sobre as maneiras que o funcionamento dos algoritmos de informação, e os modos pelos quais escolhemos navegar por eles, poderiam minar nossos esforços de sermos cidadãos envolvidos. A capacidade de personalizar resultados de pesquisa e notícias online foi a primeira, e talvez melhor articulada, dessas preocupações. Com as ferramentas de busca contemporâneas, os resultados que dois usuários recebem para a mesma busca podem ser bem diferentes; em um serviço

de notícias ou em uma rede social, as informações oferecidas podem ser adaptadas especificamente às preferências do usuário (pelo usuário ou pelo provedor), de modo que, na prática, as histórias apresentadas, como as mais importantes, podem ser tão distintas de usuário para usuário que não exista sequer um objeto comum de diálogo entre elas. Sunstein (2001) e, mais recentemente, Pariser (2011) argumentam que, quando os serviços de informação algorítmica podem ser personalizados a esse ponto, a diversidade do conhecimento público e do diálogo político pode ser minada. Somos levados - por algoritmos e por nossa própria preferência pelos que pensam de forma semelhante - para dentro de “filtros bolha” (*ibid.*), onde encontramos apenas as notícias que esperamos encontrar e as perspectivas políticas que já nos são caras.

Mas os algoritmos não estruturam apenas nossas interações com os outros enquanto membros de públicos em rede. Eles também trafegam por *públicos calculados* que eles mesmos produzem. Quando a *Amazon* recomenda um livro que “clientes como você” compraram, está invocando e afirmando conhecer um público com o qual somos convidados a sentir afinidade - embora a população em que essas recomendações se baseiam não seja transparente e certamente não coincida com toda sua base de consumidores. Quando o *Facebook* oferece como configuração de privacidade que as informações de um usuário sejam vistas por “amigos e amigos de amigos”, ele transforma um conjunto distinto de usuários em uma audiência - trata-se de um grupo que não existia até aquele momento, e que só o *Facebook* sabe sua composição precisa. Esses grupos gerados por algoritmos podem se sobrepor, podem ser uma aproximação imprecisa ou podem não ter nada a ver com os públicos que o usuário procurou.

Alguns algoritmos vão além, fazendo afirmações sobre os públicos que pretendem conhecer

e o lugar dos usuários no meio deles. Em outro trabalho, argumento que o algoritmo do “Trendings” do *Twitter* promete aos usuários um vislumbre do que um público particular (nacional ou regional) está falando em um determinado momento, mas se trata de um público construído, moldado pelos critérios específicos do *Twitter* e que não são especificados em grande parte (GILLESPIE, 2012). O *Klout*, um serviço online que rastreia as atividades e a reputação dos usuários no *Facebook*, no *Twitter* e em outros lugares, promete calcular a influência dos usuários através dessas várias plataformas de redes sociais. Suas medidas são intuitivas em sua definição, mas completamente opacas em seus mecanismos. A disputa entre os “públicos em rede” forjados pelos usuários e os “públicos calculados” oferecidos pelos algoritmos complica ainda mais a dinâmica da sociabilidade em rede.

Junto a outras medidas de opinião pública, como sondagens ou levantamentos, o problema central é a extrapolação, na qual se presume que um subconjunto represente uma população inteira. Com os algoritmos, a população pode ser a base de usuários inteira, às vezes, composta por centenas de milhões de pessoas (mas apenas a base de usuários a qual o provedor do algoritmo tem acesso). Em vez disso, o problema central aqui é que a intenção por trás dessas representações calculadas do público não é atuarial, de modo algum. Algoritmos, que alegam identificar o que está “em voga”, dedicam-se a estimar uma aproximação calculada de um público a partir dos rastros das atividades de seus participantes, e em seguida lhes apresentam um relatório sobre o que eles mais falaram. Mas, por trás disso, podemos nos perguntar: qual é a vantagem de fazer tais caracterizações para esses provedores e como isso molda o que eles estão medindo? Quem está sendo escolhido para ser medido a fim de produzir essa representação e quem é deixado de fora desse cálculo? E, talvez o mais importante, como estas tecnologias, agora não apenas tecnologias

de avaliação, mas de representação, ajudam a constituir e codificar os públicos que eles afirmam medir, públicos que de outra forma não existiriam, exceto quando o algoritmo os convoca?

Essas questões são muito importantes, e vão importar mais ainda, na medida em que as representações do público produzidas pelos algoritmos de informação sejam tomadas como legítimas, pelos usuários ou pelas autoridades, e incorporadas ao projeto modernista mais amplo de reflexividade (GIDDENS, 1990). “A sociedade está envolvida no monitoramento de si mesma, examinando-se, retratando-se de várias formas e usando os entendimentos resultantes para organizar suas atividades” (BOYER; HANNERZ, 2006, p. 9). O que o *Twitter* afirma importar para os “americanos” ou o que a *Amazon* diz que os adolescentes leem são formas de conhecimento autoritário que podem e serão invocadas por instituições cujo objetivo é regular essas populações.

A crença de que tais algoritmos, combinados com quantidades massivas de dados de usuários, são melhores para nos dizer coisas sobre a natureza do público ou sobre a constituição da sociedade, também provou ser sedutora para os estudiosos. A ciência social voltou-se avidamente para as técnicas computacionais ou para o estudo da socialidade humana através do “big data” (LAZER et al, 2009); para uma visão crítica, ver (BOYD; CRAWFORD, 2012), na esperança de desfrutar do tipo de *insights* que as ciências biológicas alcançaram, buscando algoritmicamente agulhas nesses palheiros digitais com todos esses dados. A abordagem é sedutora: ter milhões de dados gera maior legitimidade e é empolgante ver como os algoritmos parecem detectar padrões que os pesquisadores não puderam enxergar de outras formas. “Para um certo tipo de cientista social, os padrões de tráfego de milhões de e-mails parecem ter caído do céu” (NATURE, 2007). Mas essa abordagem metodológica deve considerar as complexidades descritas até aqui,

particularmente quando os dados de um pesquisador foram gerados por algoritmos comerciais. As técnicas de pesquisa computacional não são barômetros do social. Elas produzem hieróglifos: moldados pela ferramenta pela qual são esculpidos, exigindo uma interpretação sacerdotal, eles contam histórias poderosas, mas muitas vezes mitológicas - geralmente a serviço dos deuses.

Finalmente, quando nós somos os dados, o que devemos fazer com as associações que os algoritmos afirmam identificar sobre nós enquanto sociedade - o que não sabemos ou que talvez não queremos saber? No inquietante exemplo de Ananny (2011), ele percebeu que o mercado de aplicativos do *Android* recomendava um aplicativo sobre a localização de "predadores sexuais"<sup>14</sup> para os usuários que baixaram o *Grindr*, uma ferramenta de rede social baseada em localização para homens gays. Ele especula sobre como os algoritmos do mercado do *Android* podem ter feito essa associação - até mesmo as operadoras do *Android* não conseguiram explicar com facilidade. O algoritmo cometeu um erro? O algoritmo fez uma associação direta, simplesmente combinando aplicativos que continham "sexo" na descrição? Ou o mecanismo de recomendação do *Android* realmente identificou uma associação sutil que, embora não desejemos, é recorrentemente feita na nossa cultura, entre homossexualidade e comportamento predatório? Zimmer (2007) observa um caso semelhante: uma busca pela frase "ela inventou" retornava a pergunta "você quis dizer 'ele inventou'?" Ou seja, isso aconteceu até que *Google* alterou os resultados. Embora preocupante em sua política de gênero, a resposta do *Google* foi completamente "correta", explicada pelo triste fato de que, em todo o *corpus* da *web*, a palavra "inventado" é precedida muito mais vezes por "ele" do que por "ela". O algoritmo reconheceu isso - e presumiu erroneamente que a pesquisa "ela inventou" fosse apenas

um erro tipográfico. Aqui, o *Google* se prova muito menos sexista do que nós somos. Em resposta ao exemplo de Ananny, Gray sugere que, assim como devemos examinar os algoritmos que fazem esse tipo de associações, também devemos investigar os "algoritmos culturais" que essas associações representam (isto é, associando sistematicamente a homossexualidade com a predação sexual) em um conjunto massivo e distribuído de "pontos de dados" - nós mesmos.

## Considerações finais

Compreender os algoritmos e seu impacto no discurso público, então, requer pensar não apenas sobre como eles funcionam, onde são implantados ou o que os movimenta financeiramente. Este não é simplesmente um chamado para desvendar seu funcionamento interno e destacar seus critérios implícitos. É uma investigação sociológica que não interessa aos provedores desses algoritmos, que nem sempre estão na melhor posição para sequer perguntar. Requer analisar por que os algoritmos estão sendo vistos como uma lógica de conhecimento confiável; como eles desmoronam e são reparados ao entrar em contato com o vai e vem do discurso público; e onde os pressupostos políticos podem estar gravados não só em seu design, mas também serem constitutivos da sua utilização e sua legitimidade generalizada.

Vejo a emergência do algoritmo enquanto uma ferramenta de informação confiável, como a mais recente resposta à fundamental tensão do discurso público. Os meios pelos quais produzimos, circulamos e consumimos informações em uma sociedade complexa devem ser necessariamente operados através da divisão do trabalho: alguns produzem e selecionam informações, e o resto de nós, pelo menos naquele momento, somente pode conceder aquilo que valha a pena e só pode aceitá-las, se serve de consolo. Toda mídia pública anterior a essa também enfrentou esse desafio, desde

<sup>14</sup> Nota da tradutora: o termo usado pelo autor em inglês é "sex-offender".

os pregoeiros, até os jornais e as transmissões via televisão. Em cada uma delas, quando entregamos o fornecimento de conhecimento para outros, ficamos vulneráveis a suas escolhas, métodos e subjetividades. Às vezes, é um fornecimento positivo de competência, perspicácia editorial, gosto refinado. Mas também somos cautelosos à intervenção, às falhas humanas e aos interesses particulares, e nos encontramos com mecanismos apenas secundários de confiança social para confirmar o que é verdadeiro e relevante (SHAPIN, 1995). Seus procedimentos estão em grande parte indisponíveis para nós e são inevitavelmente seletivos, enfatizando algumas informações e descartando outras, e suas escolhas podem trazer consequências. Existe a nítida possibilidade de erro, viés, manipulação, comodismo, influência comercial ou política, ou falhas sistêmicas. O processo de seleção sempre pode ser uma oportunidade de organizar as informações por razões outras que não a relevância: por propriedade, por interesses próprios comercial ou institucional, ou por ganho político. Juntas, essas questões representam uma vulnerabilidade fundamental que nunca poderemos resolver completamente. Nós podemos apenas criar garantias da melhor maneira possível.

A partir dessa perspectiva, devemos ver os algoritmos não apenas como códigos com consequências, mas sim como o mais recente mecanismo construído socialmente e institucionalmente gerenciado para convencer o julgamento público: uma nova lógica de conhecimento. Podemos considerar a lógica algorítmica como oposta, e até talvez suplantar, à editorial enquanto *lógica concorrente*. A lógica editorial depende das escolhas subjetivas de especialistas, que são eles próprios feitos e autorizados como tal através de processos institucionais de treinamento e certificação, ou validados pelo público através dos mecanismos do mercado. A lógica algorítmica, em contrapartida, depende das escolhas procedimentalizadas de uma máquina, projetadas por operadores humanos para automatizar alguma representação do julgamento humano ou desenterrar padrões atra-

vés de traços sociais coletados. Ambas lutam com, e afirmam resolver, o problema fundamental do conhecimento humano: como identificar informações relevantes cruciais para o público através de meios inevitavelmente humanos, mas de formas livres de erros humanos, vieses ou manipulações. Ambas abordagens algorítmicas e editoriais do conhecimento são profundamente importantes e profundamente problemáticas; grande parte dos estudos em comunicação, mídia, tecnologia e público lida com uma ou ambas técnicas e suas armadilhas.

Uma investigação sociológica sobre os algoritmos deve aspirar revelar o complexo funcionamento desta máquina de conhecimento, tanto o processo pelo qual ela escolhe as informações para os usuários, quanto o processo social pelo qual ela se torna um sistema legítimo. Mas pode haver algo, no final, impenetrável em relação aos algoritmos. Eles são projetados para trabalhar sem intervenção humana, são deliberadamente ofuscados, e trabalham com informações em uma escala que é difícil de compreender (pelo menos sem outras ferramentas algorítmicas). E talvez mais do que isso, nós queremos nos livrar do dever de sermos céticos sobre informações que nunca poderemos garantir com certeza. Através desses mecanismos, chegamos a um acordo sobre esse problema (se não o resolvemos) e, assim, eles são soluções com as quais não apenas podemos contar, mas em que *devemos* acreditar. Mas esse tipo de fé (VAIDHYANATHAN, 2011) torna difícil reconhecer sobriamente suas falhas e fragilidades.

Assim, de várias formas, os algoritmos continuam fora do nosso alcance e eles são projetados para continuar mesmo. Isso não quer dizer que não devemos aspirar a iluminar seu funcionamento e a seu impacto. Nós deveríamos. Mas talvez nós também precisemos nos preparar para nos depararmos, mais e mais, com associações inesperadas e indescritíveis que eles vão desenhar para nós, às vezes; a incerteza fundamental sobre com quem estamos falando ou quem estamos ouvindo; e as

implicações palpáveis, porém opacas, que se movem silenciosamente por baixo do conhecimento quando ele é gerenciado por algoritmos.

## Agradecimentos

Eu gostaria de agradecer aos meus colegas do *Culture Digitally* pela ajuda e conselhos nesse artigo e pelo generoso apoio do *Collegium de Lyon* e do *The European Institutes for Advanced Study (EU-RIAS) Fellowship Programme*.

## Refêrencias

ANANNY, Mike. The curious connection between apps for gay men and sex offenders. *The Atlantic*, 2011, 14 v.

ANDERSON, Chris W. Deliberative, agonistic, and algorithmic audiences: Journalism's vision of its public in an age of audience transparency. *International Journal of Communication*, p. 529-547, 2011, 5 v.

BALKA, Ellen. *Mapping the body across diverse information systems: Shadow bodies and how they make us human*. Trabalho apresentado no Annual Meeting for the Society for Social Studies of Science, Cleveland, OH. 2011

BEER, David. Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, n. 6, p. 985-1002, 2009, 11 v.

BERRY, David Michael. Introduction. In: BERRY, David Michael (Ed.). *Life in Code and Software: Mediated Life in a Complex Computational Ecology*. London: Open Humanities Press, 2012.

BIJKER, Wiebe E.. *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change*. Cambridge, MA: MIT Press, 1997.

BOWKER, Geoffrey C. *Memory practices in the sciences*. Cambridge, MA: Mit Press, 2006.

BOWKER, Geoffrey C.; STAR, Susan Leigh. *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press, 2000.

BOYD, Danah; CRAWFORD, Kate. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, n. 5, p. 662-679, 2012, 15 v.

BOYD, danah. Social network sites as networked publics: Affordances, dynamics, and implications. PAPANCHARISSI, Zizi (Ed.). *A networked self: Identity, community, and culture on social network sites*. Routledge, 2010.

BOYER, Dominic; HANNERZ, Ulf. Introduction: Worlds of journalism. *Ethnography* 7, p. 5-17, 2006, 1 v.

BRACHA, Oren; PASQUALE, Frank. Federal Search Commission-Access, Fairness, and Accountability in the Law of Search. *Cornell Law Review*, n. 6 p. 1149-1210, 2008, 93 v.

BRAUN, Josh. *Electronic components and human interventions: Distributing television news online*. Tese de Doutorado. Cornell University, 2011.

BRUNTON, Finn; NISSENBAUM, Helen. Vernacular resistance to data collection and analysis: A political theory of obfuscation. *First Monday*, n. 5, 2011, 16 v.

BUCHER, Taina. Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, n. 7, p. 1164-1180, 2012, 14 v.

CHENEY-LIPPOLD, John. A new algorithmic identity: Soft biopolitics and the modulation of control. *Theory, Culture & Society*, n. 6, p. 164-181, 2011, 28 v.

COULDRY, Nick. *Media, Society, World: Social Theory and Digital Media Practice*. Cambridge, UK: Polity Press, 2012.

ELMER, Greg. Exclusionary rules? The politics of protocols. In: CHADWICK, Andrew; HOWARD, Philip

- N. (Ed.). *Routledge handbook of internet politics*, p. 376-283, 2008.
- FINDELSTEIN, Seth. Google, links, and popularity versus authority. In: TUROW, Joseph; TSUI, Lokman (Ed.). *The Hyperlinked Society: Questioning Connections in the Digital Age*. Ann Arbor: University of Michigan Press, p. 104-120, 2008.
- FOUCAULT, Michel. Technologies of the self. In: MARTIN, Luther; GUTMAN, Huck; HUTTON, Patrick (Ed.). *Technologies of the Self: A Seminar with Michel Foucault*. London: Tavistock, p. 16-49, 1988.
- GALLOWAY, Alexander R. *Protocol: How control exists after decentralization*. Cambridge: MA, MIT press, 2004.
- GIDDENS, Anthony. *The Consequences of Modernity*. Stanford, CA: Stanford University Press, 1990.
- GILLESPIE, Tarleton; POSTIGO, Hector Postigo. 4 maio 2012. "Five more points." In: *Culture Digitally*. Disponível em: <<http://culturedigitally.org/2012/05/five-more-points>>. Acesso em: 9 mar. 2018.
- GILLESPIE, Tarleton. 2012. "Can an algorithm be wrong?." In: *Limn 2*. Disponível em: <<http://limn.it/can-an-algorithm-be-wrong/>>. Acesso em: 22 abr. 2013.
- GILLESPIE, Tarleton. *Speaking from Platforms*. New Haven, CT: Yale University Press. Forthcoming, 2012.
- GILLESPIE, Tarleton. The politics of "platforms." *New Media & Society*, n. 3, p. 1-18, 2010, 12 v.
- GITELMAN, Lisa; JACKSON, Virginia. Introduction. In: GITELMAN, Lisa (Ed.). *Raw data is an oxymoron*. Cambridge: MA, MIT Press, 2013. p. 1-4.
- GITELMAN, Lisa. *Always already new. Media, history, and the data of culture*. Cambridge: MA, MIT Press, 2006.
- GOFFEY, Andrew. Algorithm. In: FULLER, Matthew (Ed.). *Software Studies: A Lexicom*. Cambridge, MA: MIT Press, p. 15-20, 2008.
- GRANKA, Laura A. The politics of search: A decade retrospective. *The Information Society*, n. 5, p. 364-374, 2010, 26 v.
- GRIMMELMAN, James. *The Google dilemma*. New York Law School Law Review, p. 939-950., 2008, 53 v.
- GRIMMELMANN, James. Some skepticism about search neutrality. In: SZOKA, Berin; MARCUS, Adam (Ed.). *The Next Digital Decade: essays on the future of the internet*. Washington, DC: TechFreedom, 2010.
- HALAVAIS, Alexander. *Search Engine Society*. Cambridge, UK: Polity Press, 2008.
- HESMONDHALGH, David. Bourdieu, the media and cultural production. *Media, culture & society*, n. 2, p. 211-231, 2006, 28 v.
- HILGARTNER, Stephen. *Science on Stage: Expert Advice as Public Drama*. Stanford, CA: Stanford University Press, 2000.
- HOWARD, Philip N. *New Media Campaigns and the Managed Citizen*. Cambridge, UK: Cambridge University Press, 2005.
- HUTCHINS, Edwin. *Cognition in the Wild*. Cambridge, MA: MIT Press, 1995.
- INTRONA, Lucas D.; NISSENBAUM, Helen. Shaping the Web: Why the politics of search engines matters. *The information society*, n. 3, p. 169-185, 2000, 16 v.
- ITO, Mizuko. Introduction. In: VARNELIS, Kazys (Ed.). *Networked Publics*. Cambridge, MA: MIT Press, p. 1-14, 2008.
- KEANE, John. Public life in the era of communicative abundance. *Canadian Journal of Communication*, n. 2, p. 165-178, 1999, 24 v.
- LANGLOIS, Ganaele. Participatory culture and the new governance of communication: The paradox of participatory media. *Television & New Media*, n. 2, p. 91-105, 2013, 14 v.

- LATOURE, Bruno. *Aramis, or the Love of Technology*, trans. Catherine Porter. Cambridge, MA: Harvard University Press, 1996.
- LATOURE, Bruno. *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, MA: Harvard University Press, 1987.
- LAZER, David, et. al. Computational social science. *Science*, p. 721-723, 2009, 323 v.
- LIGHT, Jennifer S. When computers were women. *Technology and culture*, n. 3, p. 455-483, 1999, 40 v.
- LONGFORD, Graham. Pedagogies of digital citizenship and the politics of code. *Techné: Research in Philosophy and Technology*, v. 9, n. 1, p. 68-96, 2005, 9 v.
- MACKENZIE, Adrian. The performativity of code: Software and cultures of circulation. *Theory, Culture & Society*, n. 1, p. 71-92, 2005, 22 v.
- MAGER, Astrid. Algorithmic ideology: How capitalist society shapes search engines. *Information, Communication and Society*, n. 5, p. 769-787, 2012, 15 v.
- MANOVICH, Lev. Database as symbolic form. *Convergence*, n. 2, p. 80-99, 1999, 5 v.
- MOROZOV, Evgeny. Don't be evil. *The New Republic*, p. 13-07, 2011.
- NATURE. "A matter of trust." *Nature*, p. 637-638, 2007, 449 v.
- PARISER, Eli. *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin Press, 2011.
- PASQUALE, Frank. 18. nov. 2009. *Assessing algorithmic authority*. In: Madisonian.net. Disponível em: <<http://madisonian.net/2009/11/18/assessing-algorithmic-authority/>>. Acesso em: 22 abr. 2013.
- PFAFFENBERGER, Brian. Technological dramas. *Science, Technology & Human Values*, n. 3, p. 282-312, 1992, 17 v.
- PINCH, Trevor; BJIKER, Wiebe. The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science*, n. 3, p.399-411, 1984, 14 v.
- RIEDER, Bernhard. "ORDER BY column\_name: The relational database as pervasive cultural form". Trabalho apresentado na *Conferência The Lived Logics of Database Machinery*, Londres, 2012.
- ROGERS, Richard. *The Googlization question, and the inculpable engine*. In: BECKER, Konrad; 2009.
- STALDER, Felix (Ed.). *Deep Search: The Politics of Search Engines*. Edison, NJ: Transaction Publishers, p. 173-184, 2009.
- RYFE, David. The nature of news rules. *Political Communication*, n.2, p. 203-214, 2006, 23 v.
- SCHUDSON, Michael; ANDERSON, Chris Anderson. Objectivity, professionalism, and truth seeking in journalism. WAHL-JORGENSEN, Karin; HANITZSCH, Thomas (Ed.). *The handbook of journalism studies*. Routledge, 2009.
- SHAPIN, Steven. Trust, honesty, and the authority of science. In: BULGER, Ruth Ellen; BOBBY, Elizabeth Meyer; FINEBERG, Harvey (Ed.) *Society's Choices: Social and Ethical Decision Making in Biomedicine*. Washington, DC: National Academies Press, p. 388-408, 1995.
- SILVERSTONE, Roger. *Television and Everyday Life*. London: Routledge, 1994.
- SMYTHE, Dallas W. On the audience commodity and its work. In: DURHAM, Meenakshi;
- KELLNER, Douglas (Ed.) *Media and Cultural Studies: KeyWorks*. Malden, MA: Blackwell Publishing, 1981, p. 230-56.
- SOLOVE, Daniel J. *The digital person: Technology and privacy in the information age*. New York: New

York University Press, 2004.

STALDER, Felix; MAYER, Christine. The Second Index: Search Engines, Personalization and Surveillance. In: BECKER, Konrad; STALDER, Felix (Ed.) *Deep Search: The Politics of Search beyond Google*. London: Transaction Publishers, 2009, p. 98-115.

STERNE, Jonathan. *The audible past: Cultural origins of sound reproduction*. Durham NC: Duke University Press, 2003.

STRIPHAS, Theodore G. How to have culture in an algorithmic age. In: *Differences and Repetitions*. 13 jul.2010. Disponível em: <<http://striphas.blogspot.com.br/2010/07/how-to-have-culture-in-algorithmic-age.html>>. Acesso em: 9 mar. 2018.

STRIPHAS, Theodore G. *The Late Age of Print: Everyday Book Culture from Consumerism to Control*. New York: Columbia University Press, 2009.

SUNSTEIN, Cass R. *Republic.com 2.0*. Princeton, NJ: Princeton University Press, 2001.

THOMPSON, John. The new visibility. *Theory, Culture & Society*, n. 6, p. 31-51, 2005, 22 v.

TUCHMAN, Gaye. Objectivity as strategic ritual: An examination of newsmen's notions of objectivity. *American Journal of Sociology*, n. 4, p. 660-679, 1972, 77 v.

TUROW, Joseph. *The daily you: How the new advertising industry is defining your identity and your worth*. New Haven, CT: Yale University Press, 2012.

TUSHNET, Rebecca. Power without responsibility: Intermediaries and the First Amendment. *George Washington Law Review*, n.4, p. 986-1016, 2007, 76 v.

VAIDHYANATHAN, Siva. *The Googlization of everything: (and why we should worry)*. Berkeley: University of California Press, 2011.

VAN COUVERING, Elizabeth. Is relevance relevant? Market, science, and war: Discourses of search engine quality. *Journal of Computer - Mediated Communication*, n. 3, p. 866-887, 2007, 12 v.

VAN COUVERING, Elizabeth. *Search engine bias: the structuration of traffic on the World-Wide Web*. 2010. Tese de Doutorado. The London School of Economics and Political Science (LSE).

VARNELIS, Kazys (Ed.). *Networked Publics*. Cambridge, MA: MIT Press, 2008

WINNER, Langdon. *Autonomous technology: Technics-out-of-control as a theme in political thought*. Cambridge, MA, MIT Press, 1977.

ZIMMER, Michael. 9 maio 2007. Google: "Did you mean: 'He invented'?" In: Michael Zimmer. Disponível em: <<http://www.michaelzimmer.org/2007/05/09/google-did-you-mean-he-invented/>>. Acesso em: 9 mar. 2018.

ZIMMER, Michael. The externalities of search 2.0: The emerging privacy threats when the drive for the perfect search engine meets Web 2.0. *First Monday*, n. 3, 2008, 13 v.

**Tarleton Gillespie** – Pesquisador-chefe da Microsoft Research New England, Professor Associado adjunto na Cornell University (EUA), membro do Social Media Collective ([socialmediacollective.org](http://socialmediacollective.org)) e um dos fundadores do blog Culture Digitally ([culturedigitally.org](http://culturedigitally.org)). **E-mail:** [tlg28@cornell.edu](mailto:tlg28@cornell.edu)



# **Examinando uma técnica algorítmica: o classificador de bayes como uma leitura interessada da realidade**

## **Scrutinizing an algorithmic technique: the Bayes classifier as interested reading of reality**

**Bernhard Rieder**

Universidade de Amsterdã e Digital Methods Initiative  
Amsterdã, Holanda

## Resumo

Este artigo descreve a noção de “técnica algorítmica” como um meio-termo entre algoritmos concretos, algoritmos implementados e a crescente área de estudos e teorizações sobre software. As técnicas algorítmicas especificam princípios e métodos para realizar atividades em um meio de software. Portanto, estes princípios e métodos constituem unidades de conhecimento especializado no âmbito da criação e desenvolvimento de softwares. Neste estudo, proponho que as técnicas algorítmicas são um objeto adequado de estudo nas áreas de ciências sociais e humanas, visto que elas capturam os princípios e técnicas centrais que se escondem detrás de softwares e raramente utilizam uma linguagem ou são descritas de forma acessível. Para propor meu estudo de caso, eu concentro o foco de análise no campo da ordenação informacional. Em primeiro lugar, discuto a ampla trajetória histórica do raciocínio formal ou “mecânico” aplicado a questões comerciais e governamentais. Em segundo lugar, foco minha investigação em uma técnica algorítmica específica: o classificador de Bayes. Essa técnica é explicada através da leitura da obra original de M. E. Maron escrita no começo dos anos 1960. A obra é apresentada como uma maneira de sujeitar uma realidade empírica e “dataficada” a uma leitura interessada que atribua significado para cada variável na relação com um objetivo operacional. Após a discussão sobre o classificador de Bayes em relação a questões de poder, o artigo retoma sua motivação inicial e defende uma crescente atenção às técnicas algorítmicas nos estudos de software.

## Abstract

This paper outlines the notion of ‘algorithmic technique’ as a middle ground between concrete, implemented algorithms and the broader study and theorization of software. Algorithmic techniques specify principles and methods for doing things in the medium of software and they thus constitute units of knowledge and expertise in the domain of software making. I suggest that algorithmic techniques are a suitable object of study for the humanities and social science since they capture the central technical principles behind actual software, but can generally be described in accessible language. To make my case, I focus on the field of information ordering and, first, discuss the wider historical trajectory of formal or ‘mechanical’ reasoning applied to matters of commerce and government before, second, moving to the investigation of a particular algorithmic technique, the Bayes classifier. This technique is explicated through a reading of the original work of M. E. Maron in the early 1960 and presented as a means to subject empirical, ‘datafied’ reality to an interested reading that confers meaning to each variable in relation to an operational goal. After a discussion of the Bayes classifier in relation to the question of power, the paper concludes by coming back to its initial motive and argues for increased attention to algorithmic techniques in the study of software.

## Palavras-chave

Técnicas algorítmicas. Classificador de Bayes. Estatística. Relações de poder.

## Keywords

Algorithmic Techniques. Bayes classifier. Statistics. Power relations.

## Introdução

Desde a popularização da Internet, os procedimentos algorítmicos para recuperação, ordenação e filtragem de informações vêm sendo o foco de atenção de especialistas em Ciências Sociais e Humanas. Os buscadores são um dos exemplos mais notáveis. Sua imensa popularidade tornou a discussão dos procedimentos de classificação relevantes para um público mais amplo, e estabeleceu o campo de estudos sobre pesquisa na *web* (ZIMMER, 2010), o que indica um interesse contínuo da comunidade acadêmica. A *web* também fornece muitas outras instâncias em que os algoritmos selecionam, hierarquizam, sugerem, e realizam outras atividades: os vendedores *online* fazem “recomendações” automáticas de produtos; *sites* de encontros calculam os coeficientes de “compatibilidade” entre os membros e os organizam de acordo com isso; os agregadores de notícias geram páginas iniciais de acordo com níveis de “importância”; os serviços de redes sociais filtram as atualizações de *status* de amigos baseadas em métricas de “proximidade”; e, os serviços de microblogging dão destaque a *trending topics* baseados em picos repentinos de atividade. Os termos entre aspas destacam que estamos lidando com tarefas culturais e, portanto, altamente ambíguas, expressas e delegadas a processos maquínicos.

Um livro popular de Eli Pariser (2011, p.72) enumera muitos outros casos de “classificação algorítmica” e argumenta que a tendência de personalizar a informação oferece o risco de nos fecharmos “filtros bolha” que nos restringem a “informações que estão em conformidade com as nossas ideias de mundo” (*idem*, 2011, p. 52). Embora a opinião de Pariser sobre o fenômeno seja limitada em termos de evidências, o livro tem o mérito de ter apresentado para um público amplo um fenômeno que sem dúvida é real e significativo: processos algorítmicos - e os amplos

arranjos nos quais estão inseridos - têm de fato começado a “transformar o mundo que vivemos ao controlar o que vemos e o que não vemos” (PARISER, 2011, p.48).

Entretanto, continua existindo um contraste entre o diagnóstico cada vez mais aceito de que o “software estrutura e possibilita grande parte do mundo contemporâneo” (FULLER, 2008, p.1) e a atenção dada a casos específicos e métodos precisos. Se o software realmente se tornou uma técnica de poder ou, precisamente, um meio para projetar e implementar técnicas complexas para “conduzir condutas” (FOUCAULT, 2004a, p. 192), é desconcertante que a análise crítica de objetos técnicos concretos, procedimentos, e práticas sejam extremamente raras. Assim como Pariser, é comum que os estudiosos das ciências humanas e sociais prefiram teorizar os efeitos sociais e políticos do software ao invés de examinar as lógicas e arranjos técnicos que os enunciados algorítmicos produzem. Os estudos de software e outros campos relacionados ao tema começaram a desenvolver este tipo de trabalho com foco na linguagem algorítmica, no código, ao contrário da *literatura* algorítmica, que são programas projetados e escritos com configurações específicas e objetivos concretos. Embora a compreensão do primeiro tema seja crucial, o panorama atual dos softwares existentes não segue teleologicamente a perspectiva da mera existência de máquinas de computação e linguagens de programação. O software, assim como uma linguagem, permite a expressão e a mecanização de uma ampla gama de ideias e objetivos, mesmo que seus princípios básicos, e suas convenções e conhecimentos históricos acumulados estruturam certos resultados.

Alguns dos esforços mais promissores para “estudar o algoritmo” (LAZER et al., 2014, p.1205) vieram de jornalistas que adaptaram suas tradicionais investigações para “relatórios de responsabilidade algorítmica” (DIAKOPOULOS, 2015, p.398), usando técnicas de engenharia reversa

para reconstruir procedimentos de decisão. Os métodos de pesquisa emergentes para a auditoria de algoritmos apontam para uma direção similar (cf. CUSTERS et al., 2013; SANDVIG et al., 2014; SWEENEY, 2013). Mas existem limites para o que se pode inferir do exterior, já que por conta de sua arquitetura, sistemas de informação são frequentemente complexos e adaptativos. No entanto, esses sistemas que não surgem de um vácuo intelectual. Partindo da observação de que o desenvolvimento de procedimentos algorítmicos tem como base ricos reservatórios de conhecimento disponibilizados por disciplinas como ciência da computação e engenharia de software, este artigo formula uma abordagem para a análise de software que se posiciona entre a teorização ampla e a investigação empírica de aplicações concretas de algoritmos de ordenação informacional. Esse meio-termo gira em torno do conjunto finito de abordagens bem conhecidas para a filtragem e classificação de informações que sustentam a maioria dos sistemas em funcionamento. Proponho chamar esses métodos padronizados, que estão no centro da prática de desenvolvimento de software e da educação em Ciências da Computação, de técnicas algorítmicas.

Na verdade, aprender a programar, no sentido de dominar uma linguagem de programação, é apenas uma pequena parte da educação em informática no nível universitário. Uma porção ainda maior se concentra nas diferentes técnicas, geralmente de matemática aprofundada, que podem ser expressas com códigos, e como aplicá-las aos notórios “problemas do mundo real” que os estudantes podem encontrar em um contexto profissional concreto. Técnicas algorítmicas são, portanto, unidades de conhecimento especializadas no campo da produção e desenvolvimento de software. O que Wing (2006) e outros pesquisadores chamaram de pensamento computacional é precisamente isso: a capacidade de abstrair de uma situação concreta para um nível onde técni-

cas algorítmicas familiares possam ser implementadas. Isso coloca os especialistas<sup>1</sup> em software em consonância com outros especialistas em métodos, como métodos estatísticos, pesquisadores operacionais, ou consultores, que dominam um conjunto de métodos e procedimentos que são aplicados facilmente em diversas situações.

Ainda que as técnicas algorítmicas sejam, em última instância, destinadas à implementação de softwares, muitas vezes elas surgem com ideias que têm pouco a ver com a Computação. Por exemplo, não há nada intrinsecamente computacional sobre noção que mensagens *spam* podem ser identificadas pela inspeção do seu conteúdo textual. A proposta de Graham (2002) para utilizar um classificador estatístico que trata cada palavra em uma mensagem como um indicador de “possibilidade de spam” transporta essa forte intuição inicial para um território computacional, ainda que os programas possam implementar a ideia de várias maneiras. O termo “classificador estatístico”, no entanto, aponta não apenas para um espaço amplo na teoria da probabilidade, mas também para uma lista de técnicas algorítmicas bem documentadas para a realização de classificações estatísticas.

Cada técnica gira em torno de uma ideia central, um núcleo conceitual que normalmente é estabelecido através da combinação de linguagem natural e notação matemática. A técnica fornece uma lógica geral e especificações de cálculo formal, mas para que a implementação efetiva desses elementos em um sistema funcione efetivamente, requer que muitas decisões sejam tomadas em relação às unidades que serão consi-

1 Existem diferenças consideráveis entre os tipos de “fabricantes de software”, mas todos eles empregam técnicas algorítmicas de diversas maneiras. As técnicas de ordenação informacional que este artigo aborda resultam de atividades intensivas de pesquisa acadêmica em Informática e Ciência da Computação, mas que agora estão disponíveis para a maioria dos desenvolvedores de software através de bibliotecas de software e por conta da extensa documentação existente.

deradas: a especificação dos parâmetros que serão utilizados; os ajustes que deverão ser feitos; os resultados que serão produzidos, e assim por diante. Eu argumentaria que a manifestação mais concreta da acumulação de conhecimento em Ciência da Computação e disciplinas relacionadas é a disponibilidade de um arquivo de técnicas cada vez maior, pronto e acessível para aqueles que sejam suficientemente habilidosos para aplicá-las. Codificar, então, é uma forma de expressar essas técnicas em termos que um computador possa entender, e os algoritmos concretos são o resultado de encontros situados entre ambientes de computação, técnicas algorítmicas e requisitos locais. Acredito que a investigação crítica sobre softwares e os vários papéis sociais que desempenham pode se beneficiar substancialmente de um entendimento aprofundado das técnicas algorítmicas. Para desenvolver esse argumento com mais detalhes, proponho investigar um exemplo específico do campo da ordenação informacional.

A menção da filtragem de *spam* já aponta para a técnica que vou examinar mais de perto neste artigo, uma vez que os *classificadores bayesianos* são normalmente usados para este tipo de tarefa. Também conhecidos como filtros Bayesianos, os classificadores de Bayes apontam para a observação de que tarefas como pesquisa, classificação e filtragem estão intimamente relacionadas em uma perspectiva técnica (cf., BELKIN; CROFT, 1992). Portanto, utilizo o termo “ordenação informacional” para abordar essas práticas como um todo. Os classificadores de Bayes fornecem um método específico para fazer uso da inferência estatística na classificação de um novo elemento, por exemplo, um e-mail recebido, baseado em um modelo de decisão derivado de elementos previamente categorizados, por exemplo, mensagens já marcadas como *spam*. Existem três razões principais para escolher esta técnica específica. Primeiro, os

classificadores de Bayes são um *pars pro toto*<sup>2</sup> para a ordenação informacional contemporânea. Na medida em que eles são probabilísticos (suas classificações não são binárias, mas sim medidas em níveis de certeza), adaptáveis (eles “aprendem” a partir da experiência) e apropriados para personalização, permitem uma discussão sobre aspectos centrais para o campo mais amplo de aprendizagem de máquina. Em segundo lugar, os princípios que os classificadores de Bayes se baseiam são suficientemente simples de serem definidos, não requerendo familiaridade com probabilidade matemática. Em terceiro lugar, os filtros bayesianos estão envolvidos em umnexo histórico - as técnicas iniciais para recuperação de informações e mineração de texto - que está intimamente relacionado com suas aplicações politicamente relevantes no presente, servindo assim de forma instrutiva para a argumentação teórica mais ampla que pretendo fazer. Enquanto os classificadores de Bayes são em primeiro lugar apenas uma das muitas técnicas aplicáveis às tarefas que nos levam a questionar o poder social dos algoritmos, como a filtragem de conteúdo ou recomendação, e avaliação automatizada de crédito ou qualificação para um emprego, eles também representam instâncias para uma *leitura interessada* de uma realidade empírica e “dataficação”<sup>3</sup> que atravessa esses exemplos.

Enfrentar a questão do poder social dos algoritmos através do exame crítico de uma técnica em particular implica um desvio significativo de abor-

2 Nota dos tradutores: a expressão significa o mesmo que uma metonímia, ou seja, tomar um termo para designar o significado de um dado referente com o qual guarde uma relação semântica ou conceitual.

3 Eu uso este termo referência à noção de “dataficação” de Mayer-Schönberger e Cukier: “os dados referem-se a uma descrição de algo que permite gravar, analisar e reorganizar. [...] Para dataficação um fenômeno é necessário colocá-lo em um formato quantificado para que possa ser tabulado e analisado” (Mayer-Schönberger & Cukier, 2013, p.78). Eu considero que a dataficação não é um processo neutro, mas como sempre implica uma seleção proposital, ou seja, concepção e mediação.

dagens mais comuns que discutem “algoritmos” de forma ampla, generalizando a vasta diversidade de técnicas existentes, ou de abordagens que examinam um domínio de aplicação específico (como a filtragem de conteúdo ou avaliação de crédito) ou uma plataforma (por exemplo, o Facebook) em que os algoritmos desempenham um papel crucial. No entanto, a abordagem desenvolvida neste artigo está alinhada com a forma específica na quais técnicas algorítmicas podem ser ao mesmo tempo diversas e gerais: enquanto todas as técnicas implicam uma maneira particular de realizar ações, é comum que elas possam ser aplicadas em uma ampla gama de domínios, caso esses domínios possuam certos requisitos básicos. Como eu apresentarei com mais detalhes abaixo, esses requisitos podem ser bastante gerais: para a aplicação de um classificador de Bayes, por exemplo, é apenas necessário apertar as entidades para classificá-las na forma de objetos que possuam propriedades quantificáveis. Assim como decidir sobre o mecanismo de feedback ou de “supervisão” para treinar o modelo estatístico. Portanto, além de e-mails, uma entidade pode ser praticamente qualquer coisa, o que inclui pessoas, situações e ideias.

Existem basicamente três posições ideais e típicas a serem tomadas quando se trata de avaliar o significado político do que há muito foi chamado de “informatização”, um termo com bela simplicidade utilizado para capturar a reconfiguração digital da qual tantas práticas se tornaram dependentes. A primeira posição considera que os computadores, os softwares e os especialistas associados a eles seguem uma lógica singular, por exemplo, a da razão instrumental. Se considerarmos essa razão instrumental inadequada, a luta política gira em torno da contenção ou reversão da informatização. Essa atitude se manifesta no trabalho de pensadores como Heidegger e Illich, ou, de forma mais recente, no estudo de Golumbia (2009) sobre o “computacionalismo” (*computationalism*). A segunda posição vê a tecnologia apenas como um epi-

fenômeno de forças sociais ou econômicas, o que significa que estudar essas próprias forças nos diria tudo o que precisamos saber. Nem o construtivismo social, em um nível micro, nem a crítica marxista, no nível macro, verificariam uma necessidade particular de estudar técnicas algorítmicas em detalhadamente, pois são consideradas consequências e não causas. Embora a tecnologia possa tornar-se relevante através da reificação ou do aumento da produtividade, os lugares reais de luta estão em outra parte. Uma terceira posição, no entanto, sustenta que a tecnologia da informação pode ter um significado político considerável sem necessariamente encapsular uma lógica singular. Isso significa que diferentes formas de informatização são concebíveis. Nesta posição, os computadores e os softwares são certamente entendidos por representar sua própria “substância”, mesmo que sua especificidade não se baseie em uma marca ideológica singular, mas nas formas particulares que representam e intervêm no mundo.

Essa posição implica em uma reforma e uma participação crítica que pode ser encontrada em campos como o movimento sobre “valores no design” (NISSENBAUM, 2005) ou no trabalho de Philip Agre (1997). Se alguém defende uma das duas primeiras posições, há poucas razões para examinar técnicas particulares em detalhe, pois tudo pode ser considerado mera expressão de grandes lógicas ou ideologias. No entanto, a terceira posição considera que as técnicas de exame são importantes, pois conduzem a uma compreensão não só do que elas realmente fazem e da forma como o fazem, mas também do que elas poderiam fazer ou do que elas poderiam fazer diferentemente. Se o software é visto como uma forma genuína de expressão e intervenção cultural, o horizonte conceitual mobilizado em artefatos técnicos é um objeto relevante de estudo para as Ciências Humanas e Sociais, bem como um lugar para a afirmação e disputa política.

Este artigo adota enfaticamente a terceira posição, mas se esforça para permanecer sensível

aos argumentos apresentados pelas outras duas. Apesar da minha breve e redutora apresentação, não devemos deixar de lado nem os efeitos homogeneizadores dos princípios, ideias e convenções informacionais, nem o seu profundo emaranhamento com configurações sociais e econômicas maiores. A escolha por trabalhar com a terceira posição não pode ser entendida como algo dado, há uma necessidade de criar, identificar, articular e defender essa escolha. Com esse espírito, uso o exemplo do classificador de Bayes tanto como meu caso estudo como para defender meu argumento.

No entanto, discutir uma técnica que produz um modelo de decisão baseado no encontro entre dados e *feedback* exige uma introdução mais ampla. Quando os editores Joseph Becker e Robert M. Hayes (1963, p. V) inauguraram a *Information Sciences Series*, uma influente série de livros publicada por Wiley & Sons a partir de 1963 com a afirmação de que “a informação é o ingrediente essencial na tomada de decisões”, eles ligaram explicitamente a jovem disciplina ao pragmatismo orientado por objetivos de pesquisa de gestão e operações, ao invés de se associarem ao ideal científico de análise e descrição desinteressada. Isso aponta para uma tradição de aplicação do raciocínio formal ou “mecânico” a questões comerciais e governamentais que compartilham poucas ideias com a famosa álgebra do pensamento de Leibniz e com o puro universalismo da lógica matemática que muitas vezes é enfatizado em histórias sobre a Computação. Mas a aplicação do cálculo para questões práticas e, em particular, para o gerenciamento do aumento do tamanho e da complexidade organizacional (BENIGER, 1986) tem uma longa história que culmina com o surgimento de estatísticas como a maneira predominante de olhar e agir em um mundo visto como dinâmico e opaco.

## **Estatística e realismo contábil**

Embora a interface entre cálculo e raciocínio tenha sido amplamente estudada através da história do “nobre” campo da lógica, foi nas arenas obscu-

ras do comércio e do governo que a Computação se tornou uma técnica aplicada para a tomada de decisões. O advento da contabilidade de dupla entrada no século XV e a circulação de *algorismi* - manuais para aprender aritmética aplicada - permitiram o crescimento no comércio de longa distância, no qual grandes navios e rotas mais longas exigiram entidades organizacionais maiores que, por sua vez, demandaram meios necessários para gerenciar esta complexa logística e compartilhar a cota de riscos e lucros. Portanto, foi esta “informatização elevada para o status de uma ciência empírica” (SWETZ, 1987, p. 295) que forneceu ferramentas robustas para controlar e decidir, e assim estabilizou, padronizou e sistematizou como os comerciantes deveriam gerenciar seus negócios e interagir com seus pares. A incerteza não foi eliminada, mas se quantificou como risco, exigindo seguro contra perdas ao invés da tentativa de evitá-las. A *Política Aritmética (Political Arithmetik)* de William Petty (1655) é notável por aplicar essas técnicas comerciais a questões governamentais, ou seja, por gerar recomendações concretas sobre investimento público e organização econômica com base no que hoje chamamos de análise custo-benefício. Estatística, como o nome indica, está profundamente interligada com tentativas de desenvolver uma “Ciência do Estado”, ou seja, um meio racional de governar, gerenciar e decidir.

Até hoje, o termo estatística manteve um duplo significado. Por um lado, refere-se à coleta de fatos ou dados, que, no século XIX, se transformou em uma “avalanche” (HACKING, 1990, p.2) ou “dilúvio” (COHEN, 2005, p. 113) de números tabulados relacionados a diversos temas e sujeitos. Este tipo de “descrição por números” é o que queremos dizer quando falamos sobre estatísticas de acidentes ou de emprego. Por outro lado, “em 1889, os usuários de estatística [...] ficaram descontentes com a mera apresentação de fatos numéricos e procuraram por métodos de análise mais refinados” (GIGERENZER et al., 1989, p. 58), o que levou ao surgimento de uma série de técnicas matemáticas usadas para analisar

e encontrar padrões nos dados coletados. Por exemplo, as dependências entre variáveis. Ambos os significados se referem a práticas epistêmicas, isto é, práticas baseadas na produção e na definição de conhecimento. No entanto, definição de conhecimento é o que efetivamente mobiliza o raciocínio mecânico em seu sentido mais completo, como uma transformação puramente formal de um conjunto de símbolos em outro que produz, no processo, um excedente epistêmico. A detecção de um nível significativo de correlação entre duas variáveis não é simplesmente uma “apresentação de um fato numérico”; e sim uma operação cognitiva que gera *interpretação* da relação entre números e, por extensão, entre o mundo que eles pretendem descrever.

Ao mesmo tempo, a estatística está ligada a uma determinada classe de problemas ou de “problematização”, ou seja, as formas específicas de apresentar e enquadrar coisas como problemas (FOUCAULT, 1984, p. 17). Como um subcampo das matemáticas, a estatística fornece técnicas de raciocínio para situações em que variadas entidades se comportam de tal maneira que não podem mais ser explicadas por uma única unidade. Por exemplo, a mecânica estatística se materializou quando ficou claro que uma descrição de um comportamento empírico dos gases com base na medição de moléculas individuais seria totalmente impossível. Da mesma forma, como observa Foucault (2004b, p. 107), o estudo das epidemias e das dinâmicas econômicas no século XIX abalou o domínio da família como modelo de compreensão e regulação da sociedade. Em vez disso, a população surgiu como uma entidade conceitual adequada, vista como uma entidade capaz de gerar fenômenos e dinâmicas que não poderiam ser reduzidas às suas partes constituintes. Tanto as moléculas como as pessoas não podem mais ser descritas em termos deterministas quando encontradas como “múltiplos vivos” (MACKENZIE; MCNALLY, 2013). Em ambos os casos, a estatística resolveria a suposta contradição entre incerteza e controle (HACKING, 1990) ao fornecer conceitos

e técnicas para raciocinar com e sobre múltiplos. A estrutura e as dinâmicas de conjuntos de entidades e casos similares, porém cambiantes, tornam-se, ao mesmo tempo, *explanandum* e *explanans*. A estatística reconhece o problema dos múltiplos como uma característica definidora do mundo e desenvolve noções como a regularidade e variação, distribuição e tendência, ou dependência e correlação para examinar, descrever e atuar sobre isto.

É secundário aceitar plenamente ou não a validade e utilidade dos conceitos e ferramentas estatísticas; o que conta é que o raciocínio estatístico, bem antes do advento dos computadores, tornou-se uma forma aceita e “confiável” (PORTER, 1995) de se relacionar com o mundo, tanto em termos de compreensão (interpretação) quanto de intervenção (tomada de decisão), seja “nas operações de governo, ao gerir negócios e finanças, nas atividades de ciência e engenharia e até mesmo em alguns aspectos da vida cotidiana” (COHEN, 2005, p. 17). Mas, como observa Desrosières (2001, p. 346), a estatística permite diferentes “atitudes” em relação à “realidade” que devem descrever, diferentes “orquestrações da realidade”. O “realismo metrológico” (DESROSIÈRES, 2001, p. 340), a correspondência ou a teoria da equivalência da verdade, é apenas uma delas. O que encontramos no comércio e no governo, onde a estatística passou a desempenhar um papel ainda maior do que na ciência, é “realismo contábil”, no qual o “‘espaço de equivalência’ não é composto de quantidades físicas (espaço e tempo), mas de um equivalente geral: dinheiro (DESROSIÈRES, 2001, p. 342). Aqui, a marca de referência (*benchmark*) para a validade de uma técnica não é mais a correspondência desinteressada, mas sua utilidade para alcançar metas específicas.

De forma não surpreendente, devido aos seguidores do pragmatismo emergente do realismo contábil, a Ciência da Informação precocemente passou a utilizar diferentes métodos estatísticos para solucionar seus múltiplos problemas. O famoso documento de Bush (1945) não foi o único a

lamentar pela “crescente montanha de pesquisas” (*idem*, 1945, p.112) na qual cientistas estavam se atolando. Foi na década 1950 e 1960 que várias técnicas e algoritmos utilizados hoje em dia foram usados para organizar informações pela primeira vez. Nessa época, apenas poucas publicações não chamaram a atenção para a crescente massa de informações escritas e sua crescente velocidade de produção.

No processo de desenvolvimento de abordagens mecânicas para resolução desse problema, o conhecimento, há muito entendido como um todo coerente e consistente que precisa ser mapeado por indivíduos esclarecidos, passa a ser definido como uma massa complexa, dinâmica e extensa de documentos que precisam ser processados e “extraídos” em relação a um requisito específico e situado. Isso significou um investimento de longo alcance no substituto computacional do conhecimento, a informação, e o desenvolvimento da “recuperação de informação” (MOOERS, 1950) como um conjunto de métodos para tomar decisões sobre documentos.

## Uma abordagem estatística para o significado da computação

É curiosa a recuperação de informação empregar o termo “informação”, que tem um significado cotidiano profundamente ambíguo e que se assemelha ao termo “conhecimento armazenado”. Ao contrário da Teoria da Comunicação de Shannon, que se preocupava com os aspectos técnicos da transmissão de sinal, a embrionária Ciência da Informação visava o que Weaver chamava de “nível B” de comunicação, ou seja, o “problema semântico” (SHANNON; WEAVER, 1949, p. 4), isto é, o *significado*.

Muitas vezes referido em termos menos intimidantes, como “assunto de um documento” (*aboutness*) ou “relevância”, manusear o significado através da ordenação informacional tornou-se o foco

central, mesmo que isso inicialmente significasse pouco mais que selecionar e classificar documentos em relação a uma consulta ou “pergunta”. Mas o campo evoluiu em um ritmo acelerado. Podemos dividir os métodos contemporâneos para a filtragem de informações em dois grupos: *filtragem baseada em conteúdo*, que as utiliza propriedades dos documentos para decidir o que fazer com eles, e a *filtragem colaborativa*, baseada na “recomendação social”. Uma parte considerável do primeiro grupo foi essencialmente estabelecida em 1975<sup>4</sup> e é possível argumentar que o segundo grupo é ao menos parcialmente antecipado pelo trabalho de análise de citações de trabalhos realizados em 1960.

A filtragem baseada em conteúdo foi o campo onde os classificadores de Bayes se desenvolveram pela primeira vez através dos usos de texto - palavras, sentenças, parágrafos etc. - como um suporte para o significado, como uma matéria-prima a partir da qual inferir “assunto de um documento” pela Computação. Este esforço se baseou na realização atribuída a Alonzo Church e a Alan Turing que afirmavam “que números eram um aspecto não essencial da computação - eles eram apenas uma maneira de interpretar os estados internos de uma máquina” (BARR; FEIGENBAUM, 1981, p. 4). Portanto, isso significava que o *tratamento digital de informações não numéricas* (MOOERS, 1950) havia se tornado um esforço plausível. Como a inteligência artificial, sua prima mais ilustre, a recuperação da informação ampliou o alcance do raciocínio mecânico para um domínio anteriormente considerado como um privilégio humano: o tratamento do conhecimento. Ao invés de reconstruir o histórico geral de recuperação de informações, esta seção se propõe a investigar de maneira particular como concretamente isso foi concebido.

<sup>4</sup> Este é o ano em que o altamente influente “modelo de espaço vetorial” foi formalmente introduzido (SALTON; WONG; YANG, 1975). O curioso é que Salton foi o orientador de Amit Singhal, chefe atual da divisão de pesquisa do Google.

O classificador de Bayes de fato se resume a aplicação de estatística na missão da Ciência da Informação em processar conhecimento e significado. A histórica começa no fim dos anos 1950, quando M. E. Maron<sup>5</sup> “estava pensando seriamente sobre o problema de recuperação de informação” (MARON, 2008, p. 971). Filósofo analítico e cibernético, Maron estava trabalhando na Ramo-Wooldridge Corporation e na (infame) Rand Corporation antes de se tornar professor contratado da Escola de Informatização da Universidade Berkeley em 1966. Nesse momento, ele estava especificamente insatisfeito com a prática de indexação de documentos, na qual se atribuía uma palavra-chave a um item para descrever o seu assunto em “dois possíveis valores temáticos” (*idem*, 2008, p. 971): um tema poderia ser atribuído ou não, mas nada era intermediário. A primeira melhoria feita por Maron em colaboração com seu colega J. L. Kuhns foi a proposta de indexadores para especificar um valor, em algum lugar entre 0 e 1, que indicasse a relevância de uma palavra-chave para um documento. Esse método, chamado de *probabilistic indexing*, tornou possível, através do teorema de Bayes e da inferência estatística inversa, fornecer uma lista de resultados ranqueada automaticamente para uma pesquisa sobre o assunto, ao invés de apresentar apenas um conjunto de documentos não ordenados (MARON; KUHNS, 1960).

Deixando de lado essa história complicada e suas várias interpretações, o teorema de Bayes fornece um método simples para calcular a probabilidade de uma hipótese baseada em um conhecimento previamente existente. Por exemplo, se conhecemos a porcentagem de mulheres em uma população e as porcentagens *a priori* de mulheres e homens com cabelos longos, podemos calcular a probabilidade da hipótese de que uma pessoa com cabelos longos que vemos de costas seja uma mulher. No

5 Este artigo se centra no trabalho de M. E. Maron, embora para dar uma visão mais completa do tema seria necessário incluir as contribuições feitas por Solomonoff (1957)

nosso caso, a hipótese diz respeito à probabilidade de um documento (marcado com um número de termos ponderados) ser relevante para uma consulta (representada por uma série de termos de pesquisa que também podem ser ponderados). Curiosamente, Maron propôs desde o início substituir a probabilidade *a priori* de um documento (o equivalente da “porcentagem de mulheres em uma população”), o que normalmente seria dividido pelo número de documentos na coleção, por estatísticas de uso do documento. Essa escolha introduz no processo o tão discutido princípio da vantagem cumulativa ou o “efeito de Matthew” (MERTON, 1968): uma vez que os documentos com maior uso acabarão na parte mais alta da lista de resultados, seu uso se tornará ainda maior. No final, a fórmula de relevância desenvolvida por Maron se parece com a fórmula abaixo, usando um produto escalar<sup>6</sup> para calcular a “proximidade” entre os termos da consulta (‘WordsQuery’ - palavras da busca) e os termos do documento (‘WordsDoc’ - palavra do documento).

$$P(\text{DocumentIsRelevant}) = (\text{WordsQuery} \cdot \text{WordsDoc}) * P(\text{DocumentUse}). (1)$$

Na busca pelo termo [hidráulica], um documento marcado com 0.7 para esse termo teria um “número de relevância” superior a um documento marcado com um valor de 0.5. Os documentos acessados com mais frequência também recebem um valor maior. Os resultados seriam então classificados de acordo com seu número de relevância. Combinações de termos e termos de pesquisa ponderados também se tornaram possíveis. Esse método ainda era baseado numa forma de indexação manual, mas representava “um ataque teórico que substitui a indexação binária tradicional por uma abordagem estatística” [...] para fazer previsões so-

6 Um produto escalar é a soma dos produtos de duas sequências de números. Por exemplo, o produto escalar entre a consulta {hidráulica: 0.7, carro: 0.5} e os termos do documento {hidráulica: 0.6, carro: 0.8} seria  $0.7 * 0.6 + 0.5 * 0.8$ , ou seja, 0.82. Quanto mais termos se sobrepõem entre uma consulta e um documento maior é o peso dos termos, e, portanto, mais “relevante” é o documento em questão.

bre a relevância dos documentos na coleção” (THOMPSON, 2008 p. 964).

No entanto, a técnica de indexação e classificação que acabei de descrever não é um classificador de Bayes. Ela introduziu o teorema de Bayes na recuperação de informações. Mas um segundo experimento foi mais longe e usou o texto completo do documento, acabando assim com o indexador humano - considerado lento, pouco confiável e tendencioso conforme os primeiros cientistas das Ciências da Informação. Com base na afirmação de que “a estatística sobre tipo, frequência, localização, ordem etc. das palavras selecionadas são adequadas para fazer previsões razoavelmente boas sobre o assunto de documentos que contenham essas palavras” (MARON, 1961, p.405), Maron concebeu uma técnica para a classificação automática de documentos: o classificador de Bayes.

A classificação, nos termos da abordagem probabilística proposta por Maron, significava que os textos deveriam ser classificados em categorias de temas especificados pelo usuário: “com base em alguma noção mais ou menos clara de categoria, devemos decidir se um documento arbitrário deve ou não ser atribuído a ela” (MARON, 1961, p. 404). Um simples exemplo contemporâneo é, de fato, a filtragem de *spams*: os e-mails são documentos e as categorias são *spam* e *não spam*. A técnica de Maron concebeu documentos de texto como objetos que contém certas propriedades (ou características). Isso simplesmente significa que cada texto foi representado por uma lista finita de palavras que contém, bem como sua frequência. Essas propriedades foram consideradas, para dizer ao menos algo sobre o assunto que o documento discute, ou o suficiente para a tarefa pragmática de recuperação de informação. Hoje, tais listas de propriedades quantificadas são geralmente referidas como *vectores* ou *vectores de características*.

O primeiro passo foi, então, selecionar uma série de documentos característicos para “trei-

namento” de cada uma das categorias temáticas (portanto, a intervenção humana foi então reorganizada e não eliminada) e, a partir daí gerar uma lista de palavras combinadas identificadas nos documentos para cada categoria. Nem todas as palavras foram consideradas. Palavras muito frequentes ou muito raras foram descartadas. A seleção resultante foi submetida a uma técnica semelhante ao que agora denominamos como frequência do termo - *métrica inversa de frequência de documento* ( $tf * idf$ ) introduzida por Karen Spärck Jones (1972) no início da década de 1970: as palavras que são distribuídas uniformemente em todas as categorias são consideradas “pistas” inadequadas e, portanto, rejeitadas. Para as palavras suficientemente “específicas” foi gerado um índice, onde a cada palavra se atribuiu um valor de relevância por categoria, determinando “certas relações de probabilidade entre as palavras portadoras de conteúdo individuais e as categorias temáticas” (MARON, 1961, p.405). Em poucas palavras: se um termo aparece frequentemente nos documentos de treinamento atribuídos a uma determinada categoria, mas é raro para outras, ela se torna uma forte *pista* ou *indicador* dessa categoria.

Uma vez que a fase de treinamento se completou, a intervenção humana não era mais necessária: com base na ideia de que um documento deveria ser atribuído a uma categoria se contiver bons indicadores para isso, novos documentos poderiam ser classificados automaticamente. O procedimento de cálculo foi muito semelhante à classificação descrita acima, mas, ao invés vez de calcular a correspondência exata entre os termos da consulta e os termos presentes no documento, foi calculada a correspondência entre a lista de palavras para uma categoria e para cada novo documento. Na medida em que muitas palavras foram incluídas, cada documento poderia receber um valor de relevância para diversas categorias. Por exemplo, o documento  $n$  era 0.8 relevante para categoria  $i$ , 0.5 para categoria  $j$ , e assim

por diante, o que resulta em uma classificação probabilística ao invés de binária. O sistema era adaptativo, já que permitia a aprendizagem além do treinamento inicial: se um usuário decidisse reclassificar um documento, as listas de palavras para categorias seriam recalculadas, somando um novo “conhecimento” ao modelo estatístico.

Este é o esboço básico de um classificador de Bayes e serve como uma ilustração surpreendentemente representativa do grande campo de técnicas contemporâneas de aprendizagem de máquina. Na medida em que essa abordagem assume a independência estatística entre as palavras, ela é o que mais especificamente pode ser designado como um classificador de Bayes “ingênuo”. Porém, esse classificador foi utilizado como ponto de partida por quase todas as técnicas modernas, mesmo que os procedimentos matemáticos fossem aplicados de forma distinta como os classificadores de máxima entropia ou as máquinas de vetor de suporte. Há uma série de aspectos importantes a serem considerados.

Em primeiro lugar, o procedimento que acabei de descrever é precisamente o que o termo “técnica algorítmica” tenta teorizar. O que acabei de descrever de maneira simples ainda não é um algoritmo em uma compreensão mais restritiva do termo, mas descreve um método de classificação que implica uma maneira de olhar e atuar no mundo. Ele enquadra e formaliza documentos de texto como listas de frequências de palavras, formula uma sequência de estágios do treinamento, passando pela classificação até o ajuste, especificando uma série de funções protomatemáticas para pesar e calcular. Qualquer desenvolvedor de software poderia ser capaz de criar programa de trabalho baseado na minha descrição, mas cada implementação seria diferente, pois muitos detalhes continuam sem ser especificados e requerem decisões. As palavras deveriam ser reduzidas aos seus radicais? Quando cortar as palavras frequentes e não frequentes? Como calcular a espe-

cificidade de uma palavra? Essas e outras questões precisam ser respondidas quando uma técnica algorítmica é trazida para realizar uma tarefa específica em um ambiente operacional específico. Um algoritmo concreto é o resultado desse processo. Isso significa que o estudo de técnicas algorítmicas não é suficiente para dar sentido aos comportamentos e compromissos específicos de sistemas reais e concretos. Porém, uma compreensão sólida de técnicas comuns pode fundamentar formas de teorização em relação aos raciocínios reais que sustentam a criação de softwares e, assim, informar maneiras mais concretas de análise empírica. Ao oferecer categorias analíticas que abrangem esses casos, essa abordagem pode estabelecer uma base para a comparação entre diferentes implementações - por exemplo, unidades e características, treinamento e configuração de feedbacks, e modalidades de decisão.

Em segundo lugar, o exemplo dado de classificação de texto é apenas uma aplicação de uma técnica mais geral. O trabalho de Maron baseia-se em um enquadramento específico de textos como listas de palavras quantificadas, ignorando completamente aspectos como ordem, sintaxe ou estilo. Assim, o significado é concebido de forma altamente redutora. Embora existam sistemas que utilizem entendimentos mais amplos e aprofundados da linguagem, ambos os enquadramentos podem coexistir, já que qualquer sistema operacional requer e depende de alguma forma, de seleção, formalização e redução.

Desse modo, a dataficação traduz hipóteses fundamentais em relação ao domínio da aplicação em estruturas de dados e as reifica. Isso geralmente implica não apenas em redução, mas também em *redução a uma forma comum*, um movimento que sustenta a generalização e explica como as técnicas podem ser aplicadas em uma grande variedade de campos. Os classificadores de Bayes são aplicáveis a todas as entidades que podem ser representadas como recursos, ou seja, listas de propriedades aos

quais valores podem ser atribuídos. Músicas podem ser consideradas vetores tendo como base o fato de que usuários as escutaram<sup>7</sup>; locais e situações podem ser classificados de acordo com as características e o comportamento das populações associadas; pessoas buscando crédito bancário podem ser aprovadas ou não com base nos seus perfis no Facebook. Mas em todos esses casos, para começar esse trabalho, alguém deve tomar as decisões em relação aos tipos de filtros e como formalizá-los. Outras técnicas de classificação envolveriam uma diferente dramaturgia com outros atores e passos a seguir, mas o campo de aplicação é comparativamente amplo. Se pensarmos em interfaces de usuário, técnicas de rastreamento e - crescentes - sensores como dispositivos que canalizam aspectos de práticas humanas como estruturas de dados formais, não é de se admirar que muitas das técnicas pioneiras de recuperação de informacional da década de 1950 estejam vivendo uma segunda "primavera" com a popularização da internet. Quanto mais fazemos coisas envolvendo aparelhos digitais, maior será o número de entidades e fenômenos que podem ser formalizados como objetos e propriedades de acordo como foi descrito. A crescente proliferação de problemas envolvendo grandes conjuntos de entidades em ambientes previamente computadorizados torna quase inevitável a aplicação de soluções algorítmicas.

Em terceiro lugar, o classificador de Bayes organiza a tomada de decisão de uma forma profundamente diferente do enquadramento comum de algoritmos como *fórmulas*. Até nas publicações acadêmicas predomina essa concepção, a qual aparenta imaginar um grupo de desenvolvedores numerando variáveis que levem em consideração e especificam de que forma serão agrupadas e pesadas.

7 No trabalho de Maron, um documento foi representado com uma lista de palavras e suas frequências, por exemplo, doc1: {palavra1: 5, palavra2, 10, ...}. Uma classificação de música de acordo com os usuários pode representar cada música com uma lista de usuários que escutam com frequência: música1: {usuário1: 5, usuário2: 10, ...}.

Os desenvolvedores do Facebook, por exemplo, estariam preocupados com os critérios para a filtragem no Feed de Notícias, meticulosamente organizando métricas como a afinidade entre os usuários, engajamento com publicações e funções de tempo para produzir uma receita de decisão clara que é guardada a sete chaves como um segredo precioso. Mas essa concepção é cada vez mais incompleta e ultrapassada; técnicas como os classificadores de Bayes propõem meios para derivar modelos de decisão a partir do encontro entre dados, um propósito e um mecanismo de *feedback*. No caso da filtragem de *spam* bayesiana, ninguém precisa compilar manualmente uma lista de palavras que sejam *spam*. O que acontece é que, em uma primeira etapa, as categorias são definidas com base no que a classificação deve atingir seu objetivo; no nosso caso, a classificação de e-mails em *spam* e *não spam*. Quando um usuário começa a marcar mensagens, uma lista de palavras é gerada automaticamente, produzindo uma base para a tomada de decisões. Não é uma fórmula clara, mas um modelo estatístico adaptativo contendo potencialmente centenas de milhares de variáveis. Quando um avaliador de empréstimo usa dados do Facebook para decidir se um candidato é suscetível a pagar um empréstimo, não há a necessidade de um sociólogo Bourdieusiano marcar todos os possíveis elementos do perfil em termos de classe social pertencente. É suficiente ter uma série de perfis de usuários que já pagaram ou que foram inadimplentes com seus empréstimos e gerar um modelo em que cada item do perfil se torne um indicador de "credibilidade". O Facebook pode desenvolver o seu mecanismo de filtragem do Feed de Notícias de maneira muito similar, usando algum mecanismo que estabeleça o "tempo gasto no site" ou a "probabilidade de clicar em um anúncio" como critério para determinar quais parâmetros de visibilidade são ótimos para cada usuário, individualmente. Em vez de selecionar e avaliar as variáveis manualmente, o classifi-

gador deriva - ou "aprende" - os parâmetros ideais da relação entre dados (postagens que possuem propriedades diferentes), comentários (o envolvimento de um usuário com essas postagens) e um propósito (aumentar ainda mais o envolvimento). O sistema é capaz de executar o seguinte comando: "mostrar ao usuário as postagens similares às que anteriormente tiveram um elevado grau de envolvimento". Portanto, o classificador de Bayes não é nem uma receita estática para tomada de decisões, nem uma teoria relacionada com uma atribuição ontológica. É um método para fazer com que dados sejam significativos em relação a um desejo de selecionar algo em concreto. É um dispositivo para a produção automatizada de leitura interessadas de uma realidade empírica. O objetivo de Maron não era dizer algo profundo sobre a relação entre texto e significado, mas projetar um sistema que produzisse no domínio de sua aplicação resultados "bons". Eu uso o termo "interessada" para enfatizar que o processo epistêmico não é apenas marcado por algum tipo de viés não desejado, mas totalmente projetado em torno de um propósito explícito que supera quaisquer escrúpulos epistemológicos e ontológicos que se possa ter. Isso se alinha com a noção sugerida por Deserosières (2011) de realismo contábil, na qual "o objetivo não é mais a verdade, mas a performatividade - isto é, a melhor equação possível de entrada / saída" (LYOTARD, 1984, p.46).

Em quarto lugar, a técnica não determina a performatividade do algoritmo resultante. O classificador de Bayes fornece a capacidade de fazer leituras interessadas, mas não especifica nem a finalidade nem a forma como uma decisão dá um retorno ou se conecta novamente ao mundo exterior. O Facebook pode decidir treinar os mecanismos do Feed de Notícias com base em qualquer critério que a empresa queira utilizar. Mas isso faz dos classificadores de Bayes ferramentas "neutras"? Mackenzie (2015 p. 444) argumenta que "à medida que a aprendizagem de máquinas é generalizada, as formas de valor que

circulam sob a forma de mercadorias se alteram", enfatizando que os diferentes "estilos" técnicos de processamento "trazem diferentes tipos de valor" (MACKENZIE, 2015, p. 436). Visto como leituras interessadas da realidade dataficação, é possível avançar ainda mais e argumentar que as técnicas de mineração de dados incorporam formas de cognição e estilos específicos de percepção que, por um lado, não são antropomórficos, já que consistem em procedimentos que só podem ser implementados por rápidos aparatos computacionais e, por outro lado, estão completamente enredados com arranjos operacionais. No nível da significação, as técnicas de mineração de dados atribuem significado a cada variável em relação a um propósito. No nível de performatividade, a mudança para infraestruturas digitais cada vez mais integradas significa que cada decisão classificatória pode ser devolvida para o mundo instantaneamente, mostrando um anúncio específico, escondendo um post específico, recusando um empréstimo a um candidato específico, estabelecendo o preço de um produto a um nível específico, e assim por diante. Nenhum dado novo permanece inocente. Se considerarmos o poder de operar como uma "rede de relações" (FOUCAULT, 1975, p. 31), podemos apreciar a forma como a mineração de dados oferece formas específicas de estabelecer, organizar e modular as relações entre as entidades dataficação para servir objetivos estratégicos. Mesmo que os classificadores de Bayes não determinem como seus resultados são usados no nível de interface, eles representam um novo conjunto de tecnologias que produzem conhecimento profundamente "interessado", usá-lo para tomar decisões com efeitos concretos é então introduzir uma "microfísica"<sup>8</sup> que tem o potencial de afetar profundamente as relações de poder.

Em quinto lugar, como já fora mencionado, os filtros de Bayes são ideais para personalização e não

8 Foucault (1975, p. 31) introduz o termo para abordar as diferentes, sutis e difusas tecnologias de poder que operam de várias maneiras em corpos.

deveria ser uma surpresa que a primeira tentativa (RICH, 1983) de personalizar a recuperação de informações para indivíduos e não apenas categorias de usuários (novato, especialista etc.) baseou-se em um método probabilístico próximo ao que acabamos de descrever. Embora ainda não haja nenhum imperativo técnico para produzir um modelo estatístico separado para cada usuário, a recuperação da informação renunciou, desde o início, às noções universalistas de ordem que encontramos na maioria dos sistemas de classificação de conhecimento usados em bibliotecas e enciclopédias. Até os primeiros sistemas enfatizaram a consulta em banco de dados e tentaram transferir o máximo de expressividade possível aos usuários, os quais eram imaginados como cientistas que sabem muito bem o que eles estão à procura. Os sistemas não foram conceitualizados como pontos de acesso convenientes a árvores de conhecimento estáveis, mas como motores de ordem, capazes de projetar estruturas latentes presentes de várias maneiras nos dados. O debate sobre o “filtro bolha” nos mostra como as sociedades liberais são marcadas pela tensão entre o universalismo e o perspectivismo. Isso também demonstra como a Internet aumentou ainda mais este problema, unindo, de um lado, um número sem precedentes de pessoas diante das mesmas interfaces e, por outro, tornando possível proporcionar a cada pessoa algo diferente. Os classificadores de Bayes formam parte de um grupo maior de técnicas e estão no centro desse problema.

De acordo com a categoria mais ampla de técnicas de ordenação informacional, os classificadores de Bayes produzem significado - e não apenas o significado dos textos - de maneiras muitas vezes complexas e diretas através da tomada de decisão informada por objetivos e propósitos específicos. De certa forma, desde o início, a famosa distinção de Bacon entre o que *é* e o que *deve ser* desaparece na forma de uma descrição que é produzida em um horizonte prescritivo. Nós não decidimos mais (ou somente) com base no

que sabemos; nós sabemos com base na decisão que devemos tomar.

Para encerrar este artigo, em primeiro lugar, eu quero voltar a ideia inicial e argumentar por que vale a pena estudar técnicas algorítmicas.

### Por que estudar técnicas algorítmicas?

Podemos afirmar que desde a Revolução Industrial nosso relacionamento com ferramentas e máquinas tem sido altamente ambíguo. Tecnologias raramente passam despercebidas, são aclamadas ou ridicularizadas, pelo menos até que se desvanecem no tecido da vida cotidiana. Desde a Segunda Guerra Mundial, os estudiosos das Ciências Humanas se propuseram a realizar estudos e avaliações, muitas vezes de forma bastante pessimista. Mas na maioria dos casos, não se reconheceu formas concretas de engajamento com objetos e procedimentos técnicos. Ao mesmo tempo, a suposta autonomia do desenvolvimento tecnológico tem sido regularmente lamentada. Pode ser que essa autonomia seja consequência da falta de engajamento com a tecnologia como tecnologia? Ou seja, a tecnologia considerada como um vasto conjunto de objetos, técnicas e práticas que só podem ser enquadradas ao conceito de grego de *techne* com extrema redução.

Para Simondon (2014), o objeto técnico ganha sua dimensão objetiva, ou seja, ganha seu lugar nas relações econômicas, sociais e psicossociais com base em sua dimensão objetiva. Isto é, em sua operação técnica. Tecnologia (*la technique*) é vista como um dos caminhos fundamentais o qual os seres humanos se relacionam com o mundo, ao lado da religião, ciência ou arte: “O que reside nas máquinas é a realidade humana, o gesto humano fixado e cristalizado em estruturas que funcionam” (SIMONDON, 1958, p. 12). Como “seres que funcionam” (SIMONDON, 2014, p. 138) é que máquinas adquirem significado e não apenas pela incorporação cultural. A avaliação do papel social que um objeto técnico

deve desempenhar, portanto, tem que partir do seu "significado funcional" (SIMONDON, 2014, p. 28). Isto é, do ele que faz e de como faz. Considerando essa perspectiva, gostaria de apresentar quatro argumentos a favor de um engajamento mais extenso com técnicas algorítmicas.

O primeiro argumento se baseia na atenção insuficiente aos aspectos e conhecimentos técnicos, o que aumenta o risco de a crítica envolver uma figura irreal. O diagnóstico de Winner que diz que "nossas concepções padrão sobre tecnologia revelam uma desorientação que beira a dissociação com a realidade" (WINNER, 1978, p.8) é mais apropriado do que nunca. Espero que o exemplo do classificador de Bayes tenha deixado claro que as técnicas contemporâneas para ordenação informacional e, portanto, o que as informações de muitas entidades podem representar, não são simplesmente transposições ou extensões de princípios mais antigos de catalogação e classificação utilizadas no espaço digital, mesmo que as interfaces empreguem metáforas similares. Como tentei mostrar, há uma certa continuidade ou ressonância com formas anteriores e aplicações de raciocínio mecânico. Através da estatística, a "teoria da probabilidade se tornou o árbitro da racionalidade prática" (GIGERENZER et al., 1989, p. 255), mas o computador aumentou consideravelmente o seu alcance e poder. No entanto, os equívocos são abundantes quando se trata de entender como softwares funcionam. Embora as lógicas não monotônicas, multivaloradas, e probabilísticas tenham sido fundamentais para a Ciência da Computação desde o seu início, ainda podemos ler que "as regras lógicas não permitem ambiguidades substantivas; uma proposição ou sucede ou não sucede" (GOLUMBIA, 2009, p. 194). Os classificadores de Bayes são exemplos perfeitos de técnicas que responderão a se um documento pertence a uma determinada categoria, não com um "sim" ou com um "não", mas com um "0.7".

Embora o verdadeiramente contínuo, posto em dúvida pela teoria quântica, esteja fora do domínio

dos computadores (digitais), o "graduável e difuso" (GOLUMBIA, 2009, p.21) tornou-se passível de ser expresso em termos informacionais durante um período muito longo de tempo. Isso é essencial para a apreciação da dimensão política do software. Ao pensar em *controle* através de algoritmos, como uma metáfora guia, devemos confiar menos nos filósofos calculadores de Leibniz e passar a confiar mais nas análises de custo-benefício de Petty, que operam hoje através de conceitos como probabilidade, média, margem, distribuição, correlação, ordem, limiar, variação, tendência, distância, equilíbrio e assim por diante. O Google pode ser uma empresa de publicidade em primeiro lugar, mas em segundo lugar é empresa de estatística. No fim das contas, é preciso estabelecer se fenômenos politicamente relevantes como os "filtros bolha" são um diagnóstico preciso, um uma fantasia mitológica ou simplesmente um pânico moral agradável. A descrição técnica é um passo necessário para a desmistificação.

O segundo argumento consiste em olhar para artefatos técnicos concretos, sejam técnicas ou implementações concretas. Perceber a oposição enganosa entre objetividade e subjetividade e, em particular, entender como a dicotomia entre julgamento e cálculo podem ser abordadas sob uma luz diferente. O termo "qualculação" (*qualculation*) (COCHOY, 2002) é útil aqui, porque mesmo os procedimentos puramente informacionais dependem de um trabalho extensivo que não é baseado em cálculos. A decisão de Maron de formalizar textos como frequências de palavras foi projetada para permitir sua computação, e não para ser considerado o resultado da computação. Callon e Law (2005) argumentam que noção de cálculo não deve ser contrastada com o julgamento (humano), mas sim estendida a este julgamento para poder ser aplicada. Isso não significa que todas as formas de raciocínio são equivalentes. Pode-se afirmar com segurança que Dean Baquet, editor executivo do The New York Times, e Krishna Bharat, criador do agregador automático de notícias Google News, têm uma influência considerável sobre as notícias que as

peças leem todos os dias; mas seu “poder” é configurado de forma bastante diferente, operando através de diferentes técnicas e mecanismos. Seguindo Callon e Law (2005), podemos descrever ambas as configurações como qualculativas, já que em ambos os casos as decisões que levam aos resultados combinam elementos de cálculo e julgamento - até o *The New York Times* não ignora os números de vendas. No entanto, as diferenças na forma como eles calculam são significativas. Como fora previamente salientado, o raciocínio mecânico não elimina o poder, mas reconfigura e desloca os critérios humanos da definição de resultados para a definição de procedimentos, mecanismos ou técnicas que vão produzir resultados. A diferença entre o *The New York Times* e o *Google News* é importante, mas não pode ser reduzido a uma mera diferença entre julgamento e cálculo. Ao estudar os classificadores de Bayes podemos ver que são abundantes os momentos que requerem tomadas de decisão: a seleção de documentos de treinamento; a eliminação de palavras; o estabelecimento de limiares e os pontos de corte; o compromisso com uma teoria de linguagem que leva frequências de palavras como indicadores de significado; as várias formas de coordenação entre o classificador e a interface, e a escolha de utilizar um método probabilístico e este método específico. Paradoxalmente, parece que quanto mais calculamos, mais é necessário julgar. Essa mudança de perspectiva deveria fazer com que trocássemos a obsoleta questão se os algoritmos são ou não “objetivos” por uma pesquisa sobre compromissos interpretativos, propósitos e parâmetros de referência incorporados nos conjuntos específicos de cálculos que os dão suporte. O estudo de técnicas algorítmicas não pode substituir a análise empírica de lugares onde decisões e implementações são finalmente feitas, mas pode ir salientando o esforço de informar a este respeito.

O terceiro argumento é que cada vez mais está ficando claro que os algoritmos - geralmente baseados em técnicas de probabilidade - estão desempenhando um papel crucial ao decidir como a informação circula, como as pessoas se encontram e se relacionam, e

como a conduta é efetivamente realizada. E enquanto quase todas as técnicas deixam espaço para muitas decisões e podem ser desenvolvidas em várias direções, seu horizonte conceitual ainda implica uma orientação epistemológica e formas específicas de intervenção. Uma crítica precisa mostrar mais frequentemente o conhecimento da relação paradoxal entre padronização e variação que caracterizam os softwares. Apesar da ênfase no “propósito” e no critério de desempenho descrito anteriormente, a técnica do classificador de Bayes não se limita em sua finalidade. Em vez disso, apresenta uma maneira específica de ler e agir sobre o mundo em relação a um propósito específico. É possível e desejável uma teoria mais ampla sobre os classificadores de Bayes e, também por extensão, sobre o aprendizado de máquinas. O argumento de Desrosières (2001) de que o realismo contábil implementa o dinheiro como principal equivalente nos traz de volta ao emaranhamento das grandes configurações sociais e econômicas que mencionei previamente. Embora eu não acredite que haja necessariamente uma conexão fatal entre o aprendizado de máquina e o capitalismo avançado, é difícil ignorar que as técnicas de leitura interessada se encaixam perfeitamente em um sistema que tenta extrair o valor monetário através de infinitas explorações de diferenças. Remetendo novamente ao vocabulário de Simondon, uma teoria geral das técnicas algorítmicas terá que aceitar que objetividade e as dimensões objetivas não podem ser mantidas separadas indefinidamente.

O quarto e último argumento apontam que a falta de envolvimento com os procedimentos tecnológicos concretos cristaliza a fenda cultural e epistemológica existente entre humanistas e engenheiros. Existe uma grande lacuna entre técnicas que orientam implementações reais e os apelos bem-intencionados de Pariser (2011), que pede às empresas e engenheiros para criar algoritmos que favoreçam a “diversidade” e o “acaso” ao invés de nos fechar em bolhas. Se quisermos tornar a filtragem de informação uma questão política em termos mais precisos, essa lacuna precisa ser superada de forma mais fundamenta-

da, ao invés de se basear em uma noção superficial que a visão administrativa de interdisciplinaridade nos proporciona. Se a “diversidade” e o “acaso” não significam aleatoriedade, teremos que pensar sobre o significado concreto que queremos dar a esses termos, e esse significado terá que encontrar uma maneira de se expressar como através de um aparato computacional. Isso não é simplesmente uma tradução de uma disciplina para outra, como de Ciências Políticas para Ciências da Computação, mas sim um processo muito mais complicado que requer uma produção minuciosa de associações precárias entre formas de expressão que são fundamentalmente diferentes. Embora eu concorde com Mackenzie (2015) que não é essencial ter uma compreensão detalhada de como os métodos de mineração de dados funcionam, acredito que uma compreensão sobre o nível mais geral de técnicas algorítmicas é tão possível quanto necessária.

Minha apresentação técnica do classificador de Bayes certamente omitiu muitos aspectos que qualquer cientista da informação ou da computação consideraria essencial. No entanto, meu objetivo era estabelecer um nível de descrição no qual um encontro entre princípios culturais e técnicos fosse possível, de tal maneira que nenhum dos “lados” fosse reduzido a uma caricatura. Podemos certamente imaginar uma ampla crítica de um estilo particular de mineração de dados, por exemplo, o estilo probabilístico. Mas quando entendemos que “probabilístico”, aqui, significa que cada propriedade de uma entidade formal é enquadrada como um indicador de associação dentro de um conjunto de categorias predefinidas, nossa crítica talvez seja capaz de produzir sugestões consistentes. Então nos perguntamos: poderíamos mostrar aos usuários quais propriedades mais contribuíram para uma decisão ser feita? Poderíamos disponibilizar na interface a propriedade ou valor de seleção? Poderíamos tornar os princípios de treinamento explícitos? Poderíamos fazer com que as associações probabilísticas desaparecessem ao longo do tempo? Poderíamos definir exceções? Essas per-

guntas sugestivas são apenas um começo, mas talvez possam ilustrar como a compreensão mais profunda e pré-formal das técnicas algorítmicas nas ciências humanas e sociais poderia levar ao surgimento de uma zona de intercâmbio na qual a falsa oposição entre cultura e tecnologia (SIMONDON, 1958, p. 9) já não se aplicaria.

## Agradecimentos

O autor gostaria de agradecer a Fernando Vander Vlist e os dois anônimos pareceristas pelo valioso *feedback*.

## Referências

- AGRE, Philip. *Toward a critical technical practice: Lessons learned in trying to reform AI. Social Science, Technical Systems and Cooperative Work: Beyond the Great Divide*. Erlbaum, 1997.
- BARR, Avron; FEIGENBAUM, Edward; ROADS, C. *The Handbook of Artificial Intelligence*. Stanford, Calif: Heuris Tech Press, 1982, 1 v.
- BECKER, Joseph; HAYES, Robert Mayo. *Information Storage & Retrieval: Tools, elements, Theories*. New York: Wiley, 1963.
- BELKIN, Nicholas J.; CROFT, W. Bruce. Information filtering and information retrieval: Two sides of the same coin? *Communications of the ACM*, n. 12, p. 29-38, 1992, 35 v.
- BENIGER, James R. *The Control Revolution: Technological and Economic Origins of the*. Cambridge: Harvard University Press, 1986.
- BUSH, Vannevar et al. As we may think. *The atlantic monthly*, n. 1, p. 101-108, 1945, 176 v.
- CALLON, Michel; LAW, John. On qualculation, agency, and otherness. *Environment and Planning D: Society and Space*, n. 5, p. 717-733, 2005, 23 v.
- COCHOY, Franck. *Une sociologie du packaging ou*

- l'âne de Buridan face au marché: Les emballages et le choix du consommateur.* Paris: Presses universitaires de France, 2002.
- COHEN, I. Bernard. *The triumph of numbers: How counting shaped modern life.* Nova York: WW Norton & Company, 2005.
- CUSTERS, B. H. M. et al. *Discrimination and Privacy in the Information Society.* Data mining and profiling in large databases. (Eds.). Berlin: Springer, 2013.
- DESROSIÈRES, Alain. How Real Are Statistics? Four Possible Attitudes. *Social Research*, p. 339-355, 2001.
- DIAKOPOULOS, Nicholas. Algorithmic accountability: Journalistic investigation of computational power structures. *Digital Journalism*, n. 3, p. 398-415, 2015, 3 v.
- FOUCAULT, Michel. *Surveiller et punir: naissance de la prison.* Paris: Gallimard, 1975.
- \_\_\_\_\_. *L'usage des plaisirs: histoire de la sexualité II.* Paris: Gallimard, 1984.
- \_\_\_\_\_. *Naissance de la biopolitique.* Paris: Gallimard, 2004a.
- \_\_\_\_\_. *Sécurité, territoire, population.* Paris: Gallimard, 2004b.
- FULLER, Matthew (Ed.). *Software studies: A lexicon.* Cambridge: MIT Press, 2008.
- GIGERENZER, G. et al. *The empire of chance: How probability changed science and everyday life* Cambridge. Cambridge: Cambridge University Press, 1989.
- GOLUMBIA, David. *The cultural logic of computation.* Cambridge: Harvard University Press, 2009.
- GRAHAM, Paul. *A plan for Spam.* Disponível em <<http://www.paulgraham.com/spam.html>>. Acesso em: 22 abril 2018.
- HACKING, Ian. *The taming of chance.* Cambridge: Cambridge University Press, 1990.
- LAZER, David et al. The parable of Google Flu: traps in big data analysis. **Science**, v. 343, n. 6176, p. 1203-1205, 2014.
- LYOTARD, Jean-François. *The postmodern condition: A report on knowledge.* Manchester: Manchester University Press, 1984.
- MACKENZIE, Adrian. The production of prediction: What does machine learning want?. *European Journal of Cultural Studies*, n. 4-5, p. 429-445, 2015, 18 v.
- MACKENZIE, Adrian; MCNALLY, Ruth. Living multiples: How large-scale scientific data-mining pursues identity and differences. *Theory, Culture & Society*, n. 4, p. 72-91, 2013, 30 v.
- MARON, Melvin Earl. Automatic indexing: an experimental inquiry. *Journal of the ACM (JACM)*, v. 8, n. 3, p. 404-417, 1961.
- \_\_\_\_\_. An historical note on the origins of probabilistic indexing. **Information Processing & Management**, v. 44, n. 2, p. 971-972, 2008.
- \_\_\_\_\_; KUHNS, John L. On relevance, probabilistic indexing and information retrieval. *Journal of the ACM (JACM)*, n. 3, p. 216-244, 1960, 7 v.
- MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. **Big Data – A Revolution That Will Transform How We Live, Think and Work.** Nova York: Houghton Mifflin Harcourt, 2013.
- MERTON, Robert K. The Matthew effect in science: The reward and communication systems of science are considered. *Science*, n. 3810, p. 56-63, 1968, 159 v.
- MOOERS, Calvin N. *The theory of digital handling of non-numerical information and its implications to machine economics.* Boston: Zator Co., 1950.
- NISSENBAUM, Helen. Values in technical design. In:

- MITCHAM, C. (Ed.) *Encyclopedia of Science, Technology and Society*. Nova York: Macmillan, 2005, p. 66-70.
- PARISER, Eli. *The filter bubble: What the Internet is hiding from you*. Nova York: Penguin, 2011.
- PETTY, William. *Several essays in political arithmetik*. Londres: D. Browne, J. Shuckburgh, J. Wiston, B. White, 1655.
- PORTER, Theodore M. *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton: Princeton University Press, 1996.
- RICH, Elaine. Users are individuals: individualizing user models. *International journal of man-machine studies*, n. 3, p. 199-214, 1983, 18 v.
- SALTON, G.; WONG, A.; YANG, C. S. A Vector Space Model for Automatic Indexing. *Communications of the ACM*. 1975, 18 v.
- SANDVIG, Christian et al. Data and discrimination: Converting critical concerns into productive inquiry. In: *The 64th Annual Meeting of the International Communication Association*. 2014.
- SHANNON, Claude Elwood. WEAVER, Warren. *The mathematical theory of communication*. Urbana: University of Illinois Press, 1949.
- SIMONDON, Gilbert. *Du mode d'existence des objets techniques*. Paris: Presses Universitaires de France, 1958.
- \_\_\_\_\_. *Sur la technique (1953-1983)*. Paris: Presses Universitaires de France, 2014.
- SOLOMONOFF, Raymond J. An inductive inference machine. In: *IRE Convention Record, Section on Information Theory*. p. 56-62, 1957.
- SPÄRCK JONES, Karen. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, n. 1, p. 11-21, 1972, 28 v.
- SWEENEY, Latanya. Discrimination in online ad delivery. *Queue*, n. 3, p. 10, 2013, 11 v.
- SWETZ, Frank J.; *Capitalism and arithmetic*. The new math of the 15th century. La Salle: Open Court, 1987.
- THOMPSON, Paul. Looking back: On relevance, probabilistic indexing and information retrieval. *Information Processing & Management*, n. 2, p. 963-970, 2008, 44 v.
- WING, Jeannette M. Computational thinking. *Communications of the ACM*, n. 3, p. 33-35, 2006, 49 v.
- WINNER, Langdon. *Autonomous technology: Techniques-out-of-control as a theme in political thought*. Cambridge: MIT Press, 1978.
- ZIMMER, Michael. *Web search studies: Multidisciplinary perspectives on web search engines*. In: *International handbook of internet research*. Dordrecht: Springer, 2010. p. 507-521.

**Bernhard Rieder** – Professor Associado de Novas Mídias e Cultura Digital na Universidade de Amsterdã e colaborador da Digital Methods Initiative. **E-mail:** willianfaraujo@gmail.com

# Por que empresas de mídia insistem que não são empresas de mídia, por que estão erradas e por que isso importa\*

## Why media companies insist they are not media companies, why they are wrong, and why it matters

### **Philip Napoli**

Duke University, DeWitt-Wallace Center  
Durham, NC, Estados Unidos da América

### **Robyn Caplan**

Rutgers University, Data & Society  
Newark, NJ, Estados Unidos da América

\* Artigo publicado originalmente na Revista Online First Monday, Volume 22, número 5. Disponível em: <http://firstmonday.org/ojs/index.php/fm/article/view/7051/6124>. Tradução: Tarcízio Silva

## Resumo

Uma posição comum entre plataformas de mídia social e agregadores de conteúdo é sua resistência em ser caracterizados como empresas de mídia. Ao contrário, empresas como Google, Facebook e Twitter tem constantemente insistido que deveriam ser pensadas como puramente empresas de tecnologia. Este artigo critica a opinião que estas plataformas são empresas de tecnologias ao invés de empresas de mídia, explora seus argumentos subjacentes e considera as implicações políticas, legais e regulatórias associadas com o aceite ou recusa desta opinião. Como o artigo ilustra, não é apenas uma distinção semântica, dado que o histórico de classificação precisa de tecnologias e serviços da comunicação possui profundas implicações em como estas tecnologias e serviços da comunicação são avaliadas por decisores políticos e tribunais.

## Abstract

A common position amongst social media platforms and online content aggregators is their resistance to being characterized as media companies. Rather, companies such as Google, Facebook, and Twitter have regularly insisted that they should be thought of purely as technology companies. This paper critiques the position that these platforms are technology companies rather than media companies, explores the underlying rationales, and considers the political, legal, and policy implications associated with accepting or rejecting this position. As this paper illustrates, this is no mere semantic distinction, given the history of the precise classification of communications technologies and services having profound ramifications for how these technologies and services are considered by policy-makers and the courts.

## Palavras-chave

Algoritmos. Mídias sociais. Políticas de mídia. Jornalismo.

## Keywords

Algorithms. Social media. Media policy. Journalism.

Os últimos anos testemunharam uma série de controvérsias emergir no campo das mídias sociais. Estas controvérsias brotaram em torno de acusações de vieses na construção de listas de *trending topics* no Facebook e Twitter (NUNEZ, 2016a, 2016b; THUNE, 2016); críticas às plataformas de mídia social por censurar fotos de relevância jornalística histórica (SCOTT & ISAAC, 2016); evidência de crescente proeminência de relatos de falsas notícias nos *feeds* de mídias sociais; e preocupações do possível impacto de falsas notícias disseminadas nas mídias sociais sobre os resultados de eleições nos Estados Unidos e Europa (BELL, 2016; ROBINSON, 2016).

Esta sequência de eventos focou a atenção pública na questão de se e como plataformas de mídias sociais estão funcionando como organizações de notícias; junto ao problema relacionado sobre se estas plataformas devem agir como organizações de notícias (ISAAC, 2017; JARVIS, 2016; MANJOO, 2016). Estas discussões estão tomando forma na medida em que dados indicam que as plataformas de mídia social se estabeleceram como os mecanismos prioritários pelos quais indivíduos acessam e consomem notícias (GOTTFRIED & SHEARER, 2016). A potencialmente problemática ironia desta situação é que estas empresas de mídia social que estão exercendo um papel crescentemente importante e influente no ecossistema de notícias e informação possuem longo histórico em insistir que não são organizações de notícias, ou mesmo empresas de mídia em absoluto. Na verdade estas empresas insistem que devem ser vistas apenas como empresas de tecnologia (ver, por ex, DIXON, 2014; D'ONFRO, 2016).

Diversos analistas apontaram esta crescente desconexão entre como estas empresas percebem a si mesmas - e como querem ser percebidas por outros - e como elas efetivamente funcionam no ecossistema atual de notícias e informação (CAPLAN, 2016; BELL, 2016; BOGOST, 2016; INGRAM, 2012; SNIP, 2016). O objetivo aqui é dar alguns

passos além nesta observação, através do exame crítico dos raciais articulados e motivações subjacentes desta posição e de considerar as amplas implicações de aceitar este argumento. Um objetivo chave deste artigo é ilustrar como este argumento se encaixa em esforços maiores e contínuos destas empresas em "discursivamente enquadradas seus serviços e tecnologias" (GILLESPIE, 2010, p. 348). Como Gillespie (2010, p. 348) nota, muitas empresas de mídia digital usam termos como "plataforma" estrategicamente, "para se posicionarem em favor tanto de lucros atuais quanto futuros, para acertar um ponto regulatório ótimo entre proteções legislativas que as beneficiem e obrigações que não, e para dispor um imaginário cultural dentro do qual seus serviços façam sentido". Nós vemos a auto-definição como empresas de tecnologia e a forte resistência a qualquer classificação como empresas de mídia como uma dimensão chave deste enquadramento discursivo, que possui importantes implicações para os panoramas regulatórios e legais que são aplicados a estas plataformas.

Esta crítica do argumento que plataformas de mídias sociais (e mais amplamente curadores de conteúdo digital) devem ser considerados puramente como empresas de tecnologia, e não como empresas de mídia será fundamentado em perspectivas econômicas, históricas e políticas de mídia, em um esforço de ilustrar os pontos significativos de continuidade, ao invés de distinção, entre tradicionais e novas plataformas para notícias e informação. Ao destrinchar e criticar este argumento, este artigo procura demonstrar a importância do que pode parecer, na superfície, apenas uma distinção semântica sem relevância. Como este artigo mostrará, há implicações legais e políticas fundamentais associadas a como estas empresas se definem e são definidas por decisores políticos e pelo público.

A seção inicial deste artigo apresenta e refuta os argumentos mais comumente usados por aquelas empresas buscando ser percebidas como empresas de tecnologia ao invés de empresas de mí-

dia. A segunda explora os principais racionais para essas empresas manterem essa posição. A terceira seção pondera as implicações amplas de aceitar os argumentos cada vez mais tênues que estas empresas são de tecnologia ao invés de mídia. A conclusão avalia a necessidade de novas abordagens definicionais que capturem melhor a crescente intersecção entre tecnologia e mídia.

### **Por que empresas de mídia insistem que não são empresas de mídia; e por que estão erradas**

Parece que no atual panorama de ambiente de negócios impulsionados por tecnologia, empresas veem cada vez mais suas identidades em termos das abordagens tecnológicas aos negócios, mais do que sobre os negócios específicos sobre os quais operam. Um dos exemplos mais proeminentes - e controversos - envolve a Uber. Uber tem mantido firmemente que é uma empresa de tecnologia ao invés de uma empresa de transporte porque, segundo um representante a Uber, "nós não transportamos bens ou pessoas - nossos parceiros transportam. Nós apenas facilitamos isto" (CUKIER, 2016, p. 1). As ramificações de aceitar ou rejeitar este argumento são profundas, porque se a Uber é percebida como uma empresa de transporte, então está sujeita a regulações sob as quais o segmento de transportes opera. Mas se, de outro lado, a Uber é percebida como uma empresa de tecnologia, as regulações para empresas de transporte simplesmente não se aplicam; o que pode representar uma importante fonte de vantagem competitiva para a Uber (ver, por ex., CARNEY, 2015). Uber certamente não está sozinha ao manter esta posição. Empresas automotivas, financeiras e de telecomunicações tem simplesmente defendido que devem ser vistas puramente como empresas de tecnologia (NAPOLI & CAPLAN, 2016).

De forma similar, parece que uma característica definidora das plataformas de mídia social e

curadoras de conteúdo digital é a consistência que mostram ao resistir ser caracterizadas como empresas de mídia enquanto insistem que devem ser pensadas puramente como empresas de tecnologia (ver, por ex., HELFT, 2008; MICKEY, 2013). Mesmo empresas que estão inequivocamente no negócio de produção de conteúdo midiático, como Gawker e Vox, levaram seus CEOs afirmar que são empresas de tecnologia ao invés de empresas de mídia (ver, por ex., SICHA, 2011; TJAARDSTRA, 2015).

Antes de olhar para os argumentos específicos apresentados por estas empresas, é importante reconhecer que a terminologia guarda-chuva "empresa de mídia" em torno do qual estes argumentos têm sido estruturados é um tanto problemática. Mídia não é um constructo monolítico e singular (apesar da tendência de tratá-las como tal no discurso que analisaremos). Primeiro, de um ponto de vista regulatório e político, mídia eletrônica como televisão e rádio (e, com a imposição das regulações de neutralidade da Internet, também a Internet) são consideradas fundamentalmente diferentes de outras mídias, geralmente recebendo níveis menores de proteção da Primeira Emenda e operando sob requisitos regulatórios mais explícitos para servir ao interesse público amplo (NAPOLI, 2001). Esta é uma importante distinção que geralmente não encontra lugar nos argumentos discutidos abaixo. Também amplamente ausente é a importante (mesmo que eventualmente difusa) distinção entre entretenimento e mídia noticiosa. De um ponto de vista regulatório e político, a produção, distribuição e consumo de notícias tem sido um grande ponto e motivação e foco para legisladores (NAPOLI, 2015); apesar que certamente preocupações sobre a dimensão cultural da mídia de entretenimento motivaram regulações e elaboração de políticas também (NAPOLI, 2008). Como será ilustrado abaixo, estes pontos de distinção amiúde importam ao pensarmos sobre o argumento "empresa de tecnologia, não empresa de mídia".

Agora nos voltamos para os principais argumentos apresentados em apoio a posição de que

plataformas de mídia social e de curadoria de conteúdo digital são empresas de tecnologia ao invés de empresas de mídia. Em cada caso, usaremos perspectivas econômicas, históricas e políticas sobre os negócios de mídia para ilustrar as fraquezas de cada argumento.

### “Nós não produzimos conteúdo”

O argumento mais proeminente entre empresas de mídia digital em apoio a posição de que não são empresas de mídia é que muitas dessas empresas não produzem conteúdo original; apenas facilitariam a distribuição de conteúdo criado pelos seus usuários. Algumas das empresas que usam este argumento tem subsequentemente realizado integração vertical na criação de conteúdo, seguindo um padrão predizível na história da mídia, na qual os distribuidores de conteúdo buscam as vantagens estratégicas e econômicas de também ser criadores de conteúdo (ver NAPOLI, 2016). Então, por exemplo, em 2016 o YouTube começou a produzir séries originais em seu serviço de assinatura; enquanto (como será discutido na Conclusão), Facebook recentemente começou incursões em criação de conteúdo de vídeo original (ETHERINGTON, 2016; PEREZ & SHIEBER, 2017; DOUGHERTY & STEEL, 2015). Contudo, dada a anterior - e provável futura - persistência destes argumento entre distribuidores de conteúdo, é importante que receba escrutínio detalhado.

O executivo da Google Eric Schmidt (entre outros executivos da empresa) usou este argumento com frequência ao longo dos anos, declarando que “Nós não fazemos nosso próprio conteúdo. Nós encontramos mais rapidamente para você o conteúdo de alguém” (SULLIVAN, 2006, p. 1). O Facebook também tem sido um proponente deste argumento há bastante tempo. Como o Vice-Presidente de Global Marketing Solutions declarou, “Na verdade nos definimos como uma empresa de tecnologia. .... Empresas de mídia são conhecidas pelo conteúdo que criam” (FIERGEMAN, 2016, p. 1). Mark Zuckerberg,

CEO do Facebook, manteve sua posição ao longo dos anos (ver, por ex., FIVEASH, 2016); e tão recentemente quanto novembro de 2016, em sua resposta às acusações sobre o papel da disseminação de *fake news* nas eleições dos EUA (ROBERTS, 2016). Dick Costolo, então CEO do Twitter, afirmou que

Eu penso em nós como uma empresa de tecnologia porque eu penso que o futuro da empresa está em construir uma plataforma extensível que permite desenvolvedores terceiros e empresas incrementar valor ao Twitter de um modo que seja benéfico ao Twitter e benéfico a nossos usuários. Eu não preciso ou quero estar no negócio de conteúdo (BILTON, 2012, p. 1)

Steve Jobs da Apple também enfatizou este ponto de modo similar no lançamento do iTunes. Incomodado pela insinuação de um entrevistador da *Esquire* que a Apple estava se tornando uma empresa de mídia, ele respondeu “Nós não somos uma empresa de mídia. Nós não possuímos mídia. Nós não possuímos música. Nós não possuímos filmes ou televisão. Nós não somos uma empresa de mídia. Nós somos apenas a Apple” (LANGER, 2003, p. 2). A questão foi suficientemente sensível para Jobs que ele finalizou a entrevista abruptamente naquele momento.

Este argumento obviamente convida à discussão dos tipos de atividades que são centrais a nosso entendimento sobre o que empresas de mídia fazem. Tradicionalmente, a organização industrial da mídia tem sido descrita em termos de três fundamentais - mas raramente mutuamente exclusivos - atividades: 1) produção (exemplificada pelos criadores de conteúdo como agências de notícias e estúdios de televisão); 2) distribuição (o processo de levar conteúdo dos produtores aos consumidores); 3) exibição (o processo de prover conteúdo diretamente às audiências) (ver, por ex., PICARD, 2011).

Digitalização e convergência de mídia significam que estes processos, em alguns casos, foram

mesclados, na medida em que conteúdo pode ser distribuído diretamente para o usuário final. O declínio de exibidores tradicionais como livrarias e lojas de música em face da digitalização do conteúdo e distribuição representa um exemplo. Seria difícil discutir que exibidores tradicionais como salas de cinema, livrarias ou lojas de disco devem ser pensadas como empresas de mídia. Porém, os substitutos destas entidades, como Amazon, iTunes e Netflix são fundamentalmente diferentes, ao utilizar a infraestrutura de mídia eletrônica para distribuir conteúdo a audiências em um formato público de maneira que mais se aproxima de curadores e distribuidores tradicionais de conteúdo como redes de televisão e sistemas de conteúdo a cabo. Este é de vários modos também o caso de plataformas de mídia social como Twitter, Snapchat e Facebook, que agora servem como meios centrais de distribuição para muitos produtores de conteúdo (de agências de notícias a indivíduos, passando pelo NFL), enquanto também servem como o ponto final de acesso aos usuários, que acessam o conteúdo através dessas plataformas via diversos dispositivos. E com serviços como Facebook Live (que permite videodifusão direta e imediata pela plataforma do Facebook) e Facebook Instant Articles (através da qual Facebook hospeda diretamente conteúdo produzido por organizações de notícias; ver CONSTINE, 2015), o processo de produção, distribuição e exibição está cada vez mais integrado.

Nós enfatizamos a evolução destes processos porque distribuição é uma característica definidora da mídia, tanto quanto é a criação de conteúdo, como um crescente grupo de pesquisadores de mídia deixa claro (ver, por ex., CURTIN et al, 2014). Criação/posse de conteúdo nunca serviu como um ponto de distinção ao se definir uma empresa de mídia da perspectiva dos encarregados com regulação do segmento de mídia. Considere, por exemplo, que tanto a televisão a cabo e indústrias via satélite foram construídas inteiramente no fundamento de servir exclusivamente (ao menos inicialmente)

como distribuidoras de conteúdo de mídia. Este fato nunca serviu como um mecanismo para manter estas empresas além dos limites da autoridade da Comissão Federal de Comunicações dos EUA (FCC - Federal Communications Commission). No caso da televisão à cabo, a FCC estendeu sua autoridade regulatória ao cabo em parte ao classificar cabo como subordinado a difusão por *broadcasting* (*United States v. Southwestern Cable Co.*, 1968<sup>1</sup>). Isto é, o fato que sistemas de cabo serviram como um meio central de distribuir e acessar conteúdo (televisão *broadcast*) que caía sob a autoridade regulatória da FCC agiu como um mecanismo para a expansão daquela autoridade regulatória ao cabo.

Consequentemente, as empresas nos setores de cabo e satélite foram sujeitas a regulações e propriedade e certas obrigações de interesse público, de acordo com a abordagem da FCC em torno de regulação de mídia eletrônica (ver, por ex., FEDERAL COMMUNICATIONS COMMISSION, 2016a). Políticas fundamentais de mídia concernentes a diversidade, competição e localismo (ver NAPOLI, 2001) tem caracterizado a regulação desses mercados, independente de se, ou a qual extensão, estas empresas se engajam em criação de conteúdo. E enquanto os mecanismos de distribuição de conteúdo no ambiente digital certamente são diferentes daqueles empregados pelos meios tradicionais, como a executiva de mídia Elizabeth Spiers perguntou, "Alguém vai me explicar como distribuição digital de seu conteúdo torna sua empresa primariamente uma empresa de tecnologia?" (BENTON, 2014, p. 1). O argumento, então, que criação/posse de conteúdo significativamente separa estas empresas de "tecnologia" do setor de mídia reflete um entendimento de mídia inocente ou pouco informado ou, ainda, um esforço intencional e estratégico de redefinir os parâmetros de uma empresa de mídia nos discursos políticos e de negócio. Em qualquer caso, as implicações são problemáticas, como discutiremos em mais detalhes abaixo.

1 Disponível em: <<https://supreme.justia.com/cases/federal/us/392/157/>>

## “Nós somos cientistas da computação”

A segunda linha proeminente de argumentação empregue pelas plataformas de mídia social e curadores de conteúdo digital foca na composição de sua equipe profissional. Especificamente, representantes daquelas empresas frequentemente apontam para suas próprias áreas de formação e histórico profissional e de seus empregados em apoio do argumento de que são empresas de tecnologia e não empresas de mídia. Eric Schmidt, da Google, enfatizou que a Google é uma empresa de tecnologia “porque é liderada por três cientistas da computação” (KRAMER, 2006, p. 1). De forma similar, o CRO da Cheezburger Ben Huh enfatizou a proporção relativa de funcionários desenvolvedores em prol da afirmação de que a empresa é de tecnologia ao invés de mídia (KOETSIER, 2012).

Aqui também vemos argumentos pobres em fundamentação lógica ou histórica. Considere, por exemplo, que no momento de sua introdução não havia maravilha tecnológica maior que a radiodifusão. Esta representou um salto tecnológico nos meios de comunicação de uma magnitude que provavelmente só pôde ser igualada pelo advento da Internet. Consequentemente, radiodifusão era a província dos tecnologistas e engenheiros em seu tempo, dado sua relativa complexidade comparada a outros meios de comunicação disponíveis, como jornais impressos. Expertise tecnológica estava no centro das primeiras empresas de radiodifusão como RCA e a Marconi Company (WOOLLEY, 2017). Tecnologia de satélite representou outro avanço dramático; outra vez requerendo profissionais com um grande nível de expertise técnica.

Avanços tecnológicos - e a expertise tecnológica associada - tem sido fundamentais para o setor de mídia ao menos desde o advento da imprensa. Como Deuze (2007, p. 73) nota, é importante que reconheçamos “o papel central que as tecnologias exercem no trabalho de mídia”. Argumentar que a orientação tecnológica da equipe ou

liderança de uma empresa representa um fundamento lógico para evitar que seja pensada como uma empresa de mídia reflete ou um conhecimento baixo da história e evolução da mídia ou um esforço estratégico de estreitar a definição de mídia relativamente a seus parâmetros tradicionais. Qualquer que seja a razão, a premissa subjacente de que tecnologia e mídia são empreendimentos separados que colhem de habilidades e capacidades profissionais completamente diferentes não tem fundamento na história da mídia.

## “Sem interferência editorial humana”

Um argumento menos explícito que estas plataformas de mídia digital oferecem, mas certamente ligado ao último argumento, é a frequente contenda que o modo pelo qual o conteúdo vai pra “superfície” não é feito por decisões humanas, mas por algoritmos e tecnologias orientadas a dados que filtram, categorizam e classificam informação já presente no sistema, apenas refletindo o que os usuários querem. Plataformas como Google e Facebook frequentemente enfatizaram a falta de intervenção humana em seus processos de curadoria de conteúdo, e tem sido hesitantes em reconhecer intervenção humana quando acontece (ver, por ex., TRIELLI et al, 2016). Facebook, no início da controvérsia das Trending News, na qual um relatório insinuou que editores humanos empregados pelo facebook estavam suprimindo notícias conservadoras na lista da plataforma, focou em subestimar ao máximo o papel que a intervenção humana direta exerce na operação da plataforma (ver, por ex., FIVEASH, 2016; ISAAC, 2017), chegando ao ponto de eliminar os cargos de jornalistas e editores que supervisionavam o módulo de Trending depois da controvérsia (THIELMAN, 2016).

Esta falta percebida/declarada ausência de envolvimento humano editorial direto é, de vários modos, fundamental para a lógica de perceber estas plataformas como empresas de tecnologia ao invés de empresas de mídia. De fato, como Gilles-

pie (2010) demonstrou, o termo plataforma em si foi estrategicamente empregado como um meio de moldar estes serviços estritamente como facilitadores de criação e disseminação de conteúdo neutros, direcionados a tecnologia. Mark Zuckerberg, do Facebook, tem afirmado que a plataforma simplesmente provê ferramentas aos usuários para ajudá-los a se engajar na criação e curadoria de conteúdo (FIVEASH, 2016), uma posição que parece ignorar - ou ao menos descaracterizar - o papel que os algoritmos da plataforma exercem na priorização e filtragem de conteúdo para os usuários. A ausência afirmada de envolvimento editorial humano ajuda a levar adiante esta percepção de distância da, e/ neutralidade no, processo de seleção de conteúdo - um modelo que é presumivelmente diferente do tipo de critério editoriais diretos (e humanos) que tem sido uma característica definidora das empresas tradicionais de mídia.

Porém, simplesmente porque os mecanismos para exercer avaliação editorial - para *gatekeeping* - mudaram, não significa que a identidade fundamental dos gatekeepers deve ser reformulada. Representantes das plataformas argumentariam que a natureza de suas interações com os usuários é fundamentalmente diferente da mídia tradicional, uma vez que usuários de mídias sociais exercem um papel muito mais autônomo na determinação de conteúdo que recebem. Isto é, os usuários, em colaboração com suas redes sociais, em última instância ditam o conteúdo que consomem, com a plataforma servindo como facilitadora neutra. Seja um reflexo de inocência ou má fé, esta posição simplesmente não é mais sustentável (GILLESPIE, 2018).

Para começar, a mídia sempre tentou, até certo ponto, dar às audiências o que elas desejam. Quanto a isto, um Facebook, Twitter ou Google são pouco diferentes de qualquer grupo de imprensa, televisão ou portais de notícia desesperadamente procurando descobrir o quê as audiências querem e prover exatamente isto (NADLER, 2016). A verdadeira diferença é que plataformas de mídias digitais sim-

plesmente representam mecanismos mais eficientes e efetivos para aquele fim, dada a maior quantidade, escopo e profundidade dos dados de usuários que são capazes de trabalhar, que é uma função da maior interatividade associada a estas novas plataformas. O relacionamento entre plataformas de mídias sociais e usuários representa então o próximo passo na progressão contínua do que Napoli (2011) denominou de "racionalização da compreensão da audiência" (NAPOLI, 2011, p. 26).

Podemos olhar para desenvolvimentos recentes, como o controverso *rebranding*<sup>2</sup> da empresa de mídia tradicional Tribune Publishing, transformada na empresa de tecnologia Tronc (ver SATELL, 2016), como um indicador de quão permeável é o limite que separa empresas de tecnologias de empresas de mídia. A metamorfose da Tribune em Tronc parece apoiar-se quase inteiramente numa declarada confiança maior em aprendizado de máquina e algoritmos para melhor servir os interesses da audiência e facilitar a automação de produção de conteúdo (SATELL, 2016). O ceticismo disseminado (e até escárnio) que este *rebranding* sofreu (ver, por ex., SATELL, 2016), em relação a se isso representa ou não algo realmente inovador, ajuda a ilustrar que a crescente confiança na tecnologia para servir aos interesses da audiência não serve como uma distinção significativa entre empresas de tecnologia e empresas de mídia.

Além disso, a concepção de que algoritmos operam de um modo completamente neutro e objetivo, livre de inclinações que caracterizam as decisões editoriais humanas diretas, tem sido efetivamente desacreditada até o momento. Algoritmos, mesmo que automatizados, são usados para classificar, filtrar e priorizar conteúdo baseado em valores internos ao sistema e preferências e ações dos usuários (DEVITO, 2016). É sabido, neste ponto, que engenheiros e outros atores organizacionais

2 Nota do tradutor: *rebranding* representa uma estratégia de *marketing* baseada na revisão/mudança de elementos constituintes de uma marca, como nome, logotipo, etc.

devem tomar inúmeras decisões no design e desenvolvimento dos algoritmos. Por meio das decisões e relações, decisões subjetivas e vieses incorporado aos sistemas, o viés algorítmico tem se tornado um fenômeno largamente pesquisado e compreendido (ver, por ex., GILLESPIE, 2017; KITCHIN, 2017; PASQUALE, 2015).

Além do mais, independentemente dos mecanismos de *gatekeeping* empregados, plataformas como Facebook, Google e Twitter encontram-se diante da necessidade de estabelecer uma série de questões sobre políticas editoriais que os posiciona firmemente junto a organizações tradicionais de mídia como jornais e radiodifusores (GILLESPIE, 2018). Questões relacionadas a proteger os usuários de conteúdo ofensivo, provocativo ou adulto são centrais à operação destas plataformas (ver, por ex., PIERSON & DAVE, 2016). Essa tensão foi bem ilustrada pela controvérsia em torno da decisão do Facebook em censurar um post de um jornalista norueguês, que continha a famosa foto de uma garota vietnamesa fugindo, nua, de um ataque com napalm durante a Guerra do Vietnã (SCOTT & ISAAC, 2016). Depois dos protestos, Facebook rapidamente reverteu a decisão baseado na importância histórica da foto (SCOTT & ISAAC, 2016). Não obstante, a controvérsia sublinhou a realidade do alcance das decisões de política editorial, que não são diferentes daquelas costumeiras em imprensa tradicional, e que são essencialmente parte do âmago das plataformas de mídias sociais, mesmo quando o processo decisório é inicialmente tomado por algoritmos, como no caso da foto no Vietnã.

E, como a controvérsia em torno da lista de Trending do Facebook ilustrou, a intervenção editorial humana exerce um papel frequentemente mais significativo do que é admitido no processo de curadoria de conteúdo (FIEGERMAN, 2016). A Snapchat emprega uma equipe editorial de cerca de 75 pessoas e também

está usando esses trabalhadores para treinar os algoritmos e tecnologias que eventualmente automatizarão a seleção e organização das notícias para 100 milhões de usuários diariamente (SLOANE, 2016). Twitter também emprega um “time de curadoria” para peneirar conteúdo e destacar eventos e tendências na sua seção “Moments” - uma seção similar a um blog no site, na qual editores do Twitter combinam tweets e fotos publicadas no site em uma narrativa, especialmente em eventos noticiáveis que estejam propagando-se no site (TWITTER, 2016). As diretrizes que estes times usam não são diferentes de diretrizes editoriais usadas por muitas empresas de conteúdo - destacam tanto sugestões de conteúdo para promover precisão, e limitar vieses, como diretrizes quanto ao tipo de títulos que o time deve escrever ou às miniaturas a escolher.

Como deve estar claro, esse cenário aponta que as empresas operam não apenas como organizações de mídia no sentido amplo, mas também especificamente como organizações de notícias, dada a amplitude na qual se engajam em *gatekeeping* e decisões editoriais sobre o fluxo de notícias e informação.

### **A centralidade da publicidade**

Por fim, vale notar que todos os argumentos oferecidos por estas plataformas ruem sob o fato de que a fonte primária de receita destas plataformas é publicidade. Estar no negócio de fornecer conteúdo a audiências, enquanto vende essas audiências a publicitários é definitivamente característico do setor de mídia (INGRAM, 2012; WOLFF, 2012). Quando publicitários exploram como alcançar melhor potenciais consumidores, as forças e fraquezas das plataformas de mídia digital como Facebook, Twitter e Google são avaliadas junto a opções mais tradicionais de mídia como televisão a cabo e aberta, publicações impressas e online.

## Motivações subjacentes

A discussão até agora focou nos argumentos apresentados por estas empresas de mídia digital para serem consideradas empresas de tecnologia ao invés de empresas de mídia. O próximo passo é olhar além destes argumentos expressos e considerar algumas razões subjacentes que explicam por que esses argumentos são apresentados tão consistentemente.

### Apelo à comunidade de investidores

Talvez a razão mais óbvia para plataformas de mídias sociais e curadores de conteúdo digital preferirem ser vistos como empresas de tecnologia ao invés de empresas de mídia é o fato de que o rótulo de empresa de tecnologia traz consigo o potencial de avaliações melhores pela comunidade de investidores (ver, por ex., FOX, 2014). Como o investidor Chris Dixon deixa claro, classificação tem implicações significativas no mundo dos investimentos. Como ele declarou, “Uma das coisas mais importantes que você deve fazer em iniciativas de investimento em empresas de estágio avançado é pensar rigorosamente como as empresas são categorizadas” (KAFKA, 2014, p. 1). E, como tem sido frequentemente observado, a comunidade de investidores geralmente vê maior potencial de receita no setor de tecnologia do que no setor de mídia (ver, por ex., BOND, 2007; ROBERTS, 2016); e parece que se mantém uma distinção forte entre estes dois setores apesar das evidências em contrário cada vez mais numerosas. Pintar-se como uma empresa de tecnologia ao invés de uma empresa de mídia é “o que capitalistas de risco [querem] ouvir” (MORRISSEY, 2016, p. 1). Como a executiva de mídia Elizabeth Spiers notou, “investidores institucionais não querem serviços de mídia” (BENTON, 2014, p. 1).

Em certa medida, esse fenômeno pode refletir uma perspectiva cultivada pelas trajetórias discrepantes dos pioneiros digitais Google e Yahoo, e talvez também está relacionado a choques de cultura que foram centrais para a morte da AOL-Time War-

ner. Yahoo “chutou os engenheiros para o lado” (FOREMSKI, 2014, p. 1) bem cedo na sua história em um esforço de evoluir para uma empresa de mídia propriamente dita. Em seu pico, Yahoo até considerou comprar a NBC em um esforço de integrar-se melhor à mídia tradicional. Estes esforços para evoluir de um mecanismo de busca ou “portal” (uma designação muito popular à época) para uma empresa mais claramente de mídia falhou espetacularmente em comparação com abordagem estrategicamente mais tecnocrática da Google.

A fusão AOL-Time Warner tornou-se um estudo de caso icônico sobre como culturas corporativas incompatíveis afundaram o tipo de integração velha mídia-nova mídia que muitos observadores viam como um resultado desejável - ou mesmo inevitável - da convergência digital (MCGRATH, 2015). O primeiro aprendizado desse desastre foi que o “pessoal de tecnologia” e o “pessoal de mídia” não trabalham particularmente bem juntos e - talvez mais importante - não parecem capazes de compreender as idiosincrasias dos negócios um dos outros (MCGRATH, 2015). No fim, AOL foi excluída do conglomerado e o nome completamente apagado da identidade da organização.

Os caminhos percorridos por Yahoo e AOL-Tim e Warner provavelmente ajudaram a reforçar a lógica da estratégia de manutenção ao menos da ilusão de distância e distinção do setor de mídia. O resultado final, em qualquer caso, é que plataformas de mídias sociais e curadores de conteúdo digital encaram um forte incentivo a se mostrar de modo que ressoe mais fortemente as percepções e prioridades da comunidade de investidores, que continua a manter uma artificial distinção estrita entre empresas de tecnologia e empresas de mídia.

### Motivações legais e políticas

Também é importante considerar as motivações legais e políticas de se pensar como uma empresa de tecnologia ao invés de empresa de mídia. Como ilustrado no exemplo do Uber, o modo pelo

qual uma empresa é classificada pode ter implicações legais e políticas relevantes. Esse certamente tem sido o caso do setor de mídia também.

Por exemplo, ao decidir a constitucionalidade do *Communications Decency Act* (que aplicou um modelo regulatório da comunicação broadcast à Internet), a Suprema Corte dos EUA debateu se trataria a Internet de forma similar ao telefone, mídia impressa ou televisão/rádio (STEIN, 1997). Dada a longa tradição nos EUA de aplicar regimes regulatórios completamente diferentes de acordo com suas características tecnológicas, o argumento de qual analogia adotar - se alguma delas - para a Internet tem implicações legais e políticas de grande alcance. Mais recentemente, a decisão da FCC de impor regulações de neutralidade de rede em provedores de serviços de Internet (ISPs) dependeu da re-classificação dos provedores de serviços de Internet como provedores de serviços de telecomunicações (similar a empresas de telefonia) ao invés de serviços de provedores de informação (como serviços de hospedagem Web) (U.S. FEDERAL COMMUNICATIONS COMMISSION, 2015; RINEHART, 2015). A autoridade regulatória da FCC sobre os provedores de serviços de telecomunicações é muito maior que sua autoridade sobre provedores de serviço de informação (novamente, diferentes tecnologias e serviços frequentemente operam sob diferentes modelos regulatórios); ao ponto que o regime de neutralidade da rede imposto pela FCC seria impossível caso os provedores de serviços de internet fossem classificados como de informação (PATEL, 2014). A iminente reversão de regulações de neutralidade na rede no governo Trump provavelmente envolverá outra rodada de re-classificação.

Problemas similares de classificação estão começando a confrontar curadores de conteúdo digital tais como as plataformas de mídias sociais. Isso é bem ilustrado por alguns desafios legais que confrontaram o Twitter. Em um caso envolvendo uma intimação governamental por informação (incluindo tweets) de um usuário individual do

serviço, a corte sugeriu que o Twitter é análogo a "gritar na janela"<sup>3</sup>. Em outra intimação ao Twitter sobre informação de IP de alguns de seus usuários (que estavam associados à organização Wikileaks), a corte usou uma diferente classificação, comparando o Twitter, e os endereços IP usados para conexão ao site, a usar um telefone.

Uma questão fundamental subjacente a estes esforços de classificação envolve definir se, ou até que ponto, o Twitter possui propriedade e autoridade editorial sobre o conteúdo que circula em sua plataforma. Ao encarar esses problemas de classificação, observadores do mercado notam que "Tem sido adequado ao Twitter posar como uma empresa de tecnologia quando se trata de potencial regulatório e responsabilidades legais" (RANA, 2012, p. 1). Por exemplo, em resposta a uma ordem da justiça por informação sobre um usuário que foi preso durante um protesto Occupy Wall Street, o Twitter adotou a posição legal de que não possui propriedade de tweets individuais (RANA, 2012), uma posição que parece operar em conflito com as várias formas de arbítrio editorial que o Twitter tem performado em relação ao conteúdo na sua plataforma (ver, por ex., GUYNN, 2016; INGRAM, 2012).

A realidade é que plataformas como Twitter possuem ampla margem de manobra em arbítrio editorial sobre o conteúdo que hospedam sem desencadear categorização como editor e contrair as responsabilidades legais que são delas decorrentes. Essa é a essência da Seção 230 do Ato de Telecomunicações de 1996. A Seção 230 prevê plataformas como mecanismos de busca, provedores de Internet e sites de mídias sociais com proteção ampla de responsabilidade legal por discursos de terceiros que porventura hospedem, selecionem ou distribuam. Desse ponto de vista, plataformas de mídias sociais e curadores de conteúdo digital estão protegidos de muitas obrigações negativas (isto é, obrigações de policiar/

3 Decisão judicial disponível em: <https://www.aclu.org/files/assets/owsharristwitterdec63012.pdf>

proteger contra a circulação dos vários tipos de discurso) associadas a empresas de mídia; ou, nas palavras da Seção 230, “provedores de conteúdo informacional”. Muito do policiamento do conteúdo que estas plataformas se engajam, deve ser enfatizado, é na verdade voluntário (sob o que é referido como “Provisão do Bom Samaritano”), não diferente do modo como a indústria de televisão a cabo tem restringido linguagem chula e nudez em suas programações, apesar disto estar fora dos limites das regulações de indecência da FCC.

Porém, em muitos setores do mercado de mídia, há também uma longa tradição de *obrigações afirmativas* (isto é, obrigações de prover certos tipos de discursos). Empresas de mídia eletrônica historicamente operaram sob um conjunto específico de responsabilidades sociais direcionadas pelo governo, que estas empresas têm contestado como onerosas e caras. Nos EUA, estas responsabilidades sociais tomaram a forma de várias “obrigações de interesse público” (NAPOLI, 2015). Mesmo as empresas de mídia eletrônica primariamente focadas em distribuição de conteúdo (como sistemas de cabo, provedores de internet e provedores de serviços de satélite) têm sido sujeitas a obrigações como prover os assinantes com acesso a níveis mínimos de programação pública, educacional e governamental; prover acesso de banda larga a escolas e bibliotecas e prover a candidatos políticos a capacidade de anunciar nessas plataformas com taxas reduzidas (U.S. FEDERAL COMMUNICATIONS COMMISSION, 2016a). Em alguns casos, essas obrigações de interesse público tornaram-se particularmente agressivas, como no caso da Fairness Doctrine (Doutrina de Equidade), que, durante os anos 70 e parte dos anos 80, exigiu que difusores (dado a seu lugar privilegiado) provessem quantidades equivalentes de cobertura a diferentes perspectivas sobre temas controversos de importância pública (RUANE, 2011).

Mesmo regulações estruturais e governamentais motivadas economicamente tem sido historica-

mente agressivamente mais impostas no setor de mídias do que em outros setores, dado preocupações sobre o relacionamento entre competição nos mercados de mídia e o funcionamento efetivo dos “mercados de ideias” (NAPOLI, 2001). Várias regulações de propriedade persistem no setor de mídia eletrônica, apesar do acirramento de competição facilitado pela Internet e a diminuição de barreiras para novos entrantes em vários mercados de mídia (U.S. FEDERAL COMMUNICATIONS COMMISSION, 2016b). Fusões no setor de mídia eletrônica passam por uma avaliação própria de “interesse público” com escrutínio acima e além da média das demais fusões em termos de seus impactos no ambiente competitivo. Nesses casos, preocupações totalmente não relacionadas ao efeito proposto da fusão na competição podem entrar em jogo no processo de aprovação, no qual várias condições de interesse público são impostas para aprovação da fusão (SALLET, 2014).

O ponto chave na avaliação de como o interesse público é aplicado no contexto da regulação de mídia eletrônica é demonstrar que classificação como uma empresa de mídia - ao menos no reino da mídia eletrônica - historicamente tem significado observação governamental mais intensiva, na forma de obrigações afirmativas de servir ao interesse público e regulação mais rigorosa em áreas como concentração de propriedade. Dada essas cargas regulatórias maiores que caem sobre o setor de mídia eletrônica, é possível entender por que plataformas de mídia digital se esforçaram para estabelecer identidade organizacional que os deixe de fora deste modelo regulatório.

Imagine, por exemplo, se Facebook tivesse de operar sob uma Doutrina de Equidade para mídias sociais. De vários modos, poderia se olhar as controvérsias em torno da presumida supressão de notícias conservadoras ou a circulação de notícias falsas durante a eleição presidencial de 2016 como um tipo de fagulha que teria iniciado esse tipo de discussão. Ou, ainda, seria possível olhar para uma

plataforma de mídias digitais como Facebook, com sua posição dominante no mercado, com interesse crescente sobre a possibilidade de uma análise de competição mais apropriada a mídia sendo aplicada a sua posição cada vez mais dominante tanto no mercado econômico quanto no mercado de ideias (ver THOMPSON, 2016). Dados recentes mostram a extensão que o duopólio Facebook/Google desenvolveu em publicidade online e móvel (SLEFO, 2017). O fato um pouco surpreendente de que essas descobertas ainda não provocaram um debate amplo no discurso político dos EUA pode ser talvez atribuído ao sucesso, até o momento, da retórica empresa-de-tecnologia-não-empresa-de-mídia.

Como essa discussão esperançosamente deixou clara, a questão de se plataformas tais como Facebook e Twitter são empresas de mídia ou empresas de tecnologia não é só uma matéria de semântica mas também parte de uma estratégia discursiva maior. O sucesso desta estratégia tem implicações legais e políticas significantes, e é aquela na qual as motivações legais, econômicas e políticas das plataformas de mídia digital são (des)percebidas como empresas de tecnologia ao invés de empresas de mídia são bastante.

## Por que importa

Como pesquisadores de políticas da comunicação notaram, *palavras importam* (LENTZ, 2003). Os termos específicos empregados no discurso e documentos que formatam e refletem decisões políticas tem profundas consequências e devem ser empregados estrategicamente; muitas vezes em um esforço de definir os contornos de um tema ou para influenciar quais *stakeholders* caem dentro dos parâmetros de uma proposta específica de política pública (ver, por ex., LENTZ, 2011; STREETER, 1987). Dentro dos contextos da comunicação e mídia, uma estratégia discursiva específica entre muitos *stakeholders* tem sido empregar um foco fortemente tecnológico, tratando novas tec-

nologias de comunicação e mídia como agentes autônomos, e/ou estritamente definindo o terreno regulatório puramente em termos de questões e preocupações técnicas complexas, excluindo preocupações sociais maiores (NAPOLI, 2009; STREETER, 1987). Pesquisadores de governança da Internet, por exemplo, tem observado uma tendência entre muitos *stakeholders* de tentar despojar várias questões de governança de Internet de suas expressões políticas e culturais mais amplas, procurando enquadrar essas questões em seus termos estritamente técnicos, um processo que em última análise marginaliza do foco primariamente político ou cultural problemas de governança concernentes aos *stakeholders* (ver, por ex., RABOY & LANDRY, 2005). Esta tecnocratização do discurso político serve como um importante pano de fundo para o argumento "empresa-de-tecnologia-não-empresa-de-mídia" que sendo avaliado aqui.

No fim das contas, o enquadramento de plataformas de mídias sociais e curadores de conteúdo digital puramente como empresas de tecnologia marginaliza as cada vez mais proeminentes dimensões políticas e culturais de suas operações, que crescem mais pronunciadamente na medida em que estas plataformas se tornam *gatekeepers* centrais de notícias e informação no ecossistema contemporâneo de mídia. Nessas situações, onde há uma desconexão entre função e enquadramento, nós temos um "discurso [que] serve para moldar uma instituição que falha em descrevê-la" (STREETER, 1987, p. 176). Esse enquadramento tecnocêntrico pode contribuir para que estas plataformas operem sobretudo fora de enquadramentos legais e regulatórios que foram estabelecidos para organizações de mídia eletrônica; enquadramentos que foram estabelecidos sobretudo devido às significativas dimensões políticas e culturais de suas operações.

Essa situação torna-se particularmente problemática quando essas plataformas inevitavelmente evoluem de forma que tornam a distinção empresa-de-tecnologia-não-empresa-de-mídia cada vez

mais inválida. Por exemplo, como apontado acima, assim como gerações prévias de distribuidores de conteúdo, como sistemas de cabo e satélite e, mais recentemente, YouTube e Netflix (KOKALITCHEVA, 2016; NAPOLI, 2016), Facebook começou a integrar-se verticalmente com a criação de conteúdo. A empresa iniciou esforços para produzir programação de vídeo original (KAFKA, 2014). E, no despertar das massivas críticas sobre o papel do Facebook na disseminação de notícias falsas durante as eleições de 2016, a empresa iniciou o Facebook Journalism Project. Entre os objetivos do projeto estão "construir laços mais profundos com publicadores ao colaborar com recursos e ferramentas de publicação" (ISAAC, 2017, p. 1). Essa maior integração nas dinâmicas da produção e distribuição é, talvez, o primeiro passo na direção a uma integração vertical e genuína na área de notícias; mas é no mínimo uma mudança na direção organizacional na medida em que a referência ao jornalismo debilita quaisquer argumentos contra ser categorizada como uma empresa de mídia (BARAM, 2017).

Talvez não seja surpreendente, então, que o CEO do Facebook Mark Zuckerberg começou a suavizar um pouco sua posição quanto a identidade organizacional do Facebook. Em dezembro de 2016 ele descreveu Facebook como "não uma empresa de mídia *tradicional*" (CONSTINE, 2016, online, ênfase adicionada pelos autores). A posição cambiante de Zuckerberg é remanescente da do Google, que depois de anos insistindo que não é uma empresa de mídia, declarou em seu formulário SEC 10-K<sup>4</sup> que "Nós começamos como uma empresa de tecnologia que evoluiu para uma empresa de software, tecnologia, Internet, publicidade e mídia ao mesmo tempo" (BIA/KELSEY, 2005, p. 1). Em ambos os casos, a posição é de manter pontos significativos de

4 Nota do tradutor: relatório que sumariza a performance financeira de empresas com mais de 10 milhões em ativos nos EUA e também inclui informações de histórico, estrutura organizacional e outras sobre a empresa.

distinção e separação com empresas de mídia tradicional. Apesar disso, quando estas empresas reconhecerem que são empresas de mídia de qualquer tipo, muitos benefícios associados a não ser percebidas, ou tratadas como, empresas de mídia já terão sido colhidos. De fato, parte do modo pelo qual essas empresas conseguem evoluir para organizações de mídia dominantes é por contornar os quadros regulatórios e legais que poderiam ter inibido suas rápidas expansões e/ou imposto mais responsabilidades legais e/ou custos de responsabilidade pelo caminho. E a estratégia discursiva debatida aqui provavelmente tem seu papel em facilitar essa evasão.

É importante enfatizar que nós não procuramos aqui construir recomendações para intervenções legais ou regulatórias específicas. Nosso objetivo foi simplesmente esclarecer as falhas lógicas e motivações subjacentes de uma estratégia discursiva que dissocia plataformas de mídias sociais dos quadros normativos, legais e regulatórios que tradicionalmente têm sido aplicados a instituições que tradicionalmente performam as funções que estas plataformas agora performam. Por esta razão, nós tentamos rejeitar essa caracterização puramente orientada a tecnologia das plataformas de mídias sociais e curadores de conteúdo digital, na esperança que o discurso político e de governança sobre elas melhor reflita o fato de que os pontos de continuidade entre mídias novas e tradicionais são mais fortes que suas caracterizações sugerem.

## Conclusão

Através da história da mídia, as tecnologias e serviços têm evoluído para possibilitar funções que tornam-se muito diferentes daquelas para as quais foram criadas. Rádio foi criada para facilitar comunicação navio a costa. O videocassete foi criado para facilitar gravação de vídeo em casa e

seus criadores não faziam ideia do mercado massivo de aluguel e vendas de vídeo que iria surgir. A Internet foi criada para facilitar comunicação entre governo e pesquisadores acadêmicos. Facebook foi criado para ajudar estudantes universitários identificar colegas atraentes.

A velocidade com a qual as funcionalidades de tecnologias da comunicação podem evoluir levam a um estado persistente de retardamento entre percepção e realidade. Quando a funcionalidade em questão envolve informar e influenciar o processo democrático, essa defasagem perceptual pode ter ramificações perigosas e profundas. A eleição de 2016 parece ter desencadeado um pouco de aceleração neste processo, com usuários, jornalistas, legisladores, acadêmicos e até mesmo os executivos dessas plataformas parecem estar desenvolvendo um entendimento mais claro do papel cada vez mais influente que estas plataformas exercem no processo democrático. Se esta clareza maior levará a normas e estruturas de governança que melhor sirvam ao interesse público é algo a ser descoberto e deve ser um foco de pesquisa daqui pra frente.

Finalmente, tanto quanto este artigo defendeu que plataformas de mídias sociais e curadores de conteúdo digital enquadram-se muito bem dentro dos parâmetros de organizações de mídia, alguém poderia facilmente argumentar que é necessário e apropriado ao nosso entendimento de empresas de mídia também evoluir para englobar a estrutura e operação destas plataformas. Abordagens recentes de definição de mídia parecem refletir essa perspectiva (ver, por ex., HESS, 2014). Deuze (2007, p. 9), por exemplo, enfatiza que "Indústrias de mídia produzem conteúdo, de fato, mas também investem em plataformas para conectividade". Pelo menos, talvez seja hora de delinear mais rigorosamente os contornos e parâmetros da empresa híbrida de tecnologia/mídia contemporânea, e começar a articular se e como essas empresas devem caber em enquadramentos

legais, regulatórios e jornalísticos existentes; ou se devem ser concebidos novos ou modificados enquadramentos que reflitam sua natureza híbrida.

## Referências

BARAM, M. Facebook Finally Admits It's A Giant Media Company—Almost. *Fast Company*. 2017. Disponível em <<https://www.fastcompany.com/3067148/facebook-finally-admits-its-a-giant-media-company-almost>>. Acesso em: 03 abril 2018.

BELL, E. Facebook Facebook is eating the world. *Columbia Journalism Review*. 2016. Disponível em <[https://www.cjr.org/analysis/facebook\\_and\\_media.php](https://www.cjr.org/analysis/facebook_and_media.php)>. Acesso em: 03 abril 2018.

BENTON, J. Elizabeth Spiers on BuzzFeed and other "tech" companies: "You're Still A Media Company". *NiemanLab*. 2014. Disponível em <<http://www.niemanlab.org/2014/08/elizabeth-spiers-on-buzzfeed-and-other-tech-companies-youre-still-a-media-company/>>. Acesso em: 03 abril 2018.

BHATTACHARYYA, M. Digital Ads to Overtake Traditional Ads in U.S. Local Markets By 2018. *AdAge*. 2016. Disponível em <<http://adage.com/article/cmo-strategy/local-ads-digital-2018-bia-kelsey/306468/>>. Acesso em: 03 abril 2018.

BIA/KELSEY. Google 10-K: 'We're a Media Company' and Other Tidbits. 2005. Disponível em <<http://blog.biakelsey.com/index.php/2005/04/04/google-10-k-were-a-media-company-and-other-tidbits/>>. Acesso em: 03 abril 2018.

BILTON, N. Is Twitter a Media or Technology Company?. *The New York Times*. 2012. Disponível em <<https://bits.blogs.nytimes.com/2012/07/25/is-twitter-a-media-or-technology-company/>>. Acesso em: 03 abril 2018.

BOGOST, I. Facebook Is Not a Technology Company. *The Atlantic*. 2016. Disponível em <<https://www>>.

theatlantic.com/technology/archive/2016/08/facebook-is-not-a-technology-company/494183/>. Acesso em: 03 abril 2018.

BOND, P. Fresh Facebook draws attention. *The Hollywood Reporter*. 2007. Disponível em <<https://www.hollywoodreporter.com/news/fresh-facebook-draws-attention-151237>>. Acesso em: 03 abril 2018.

CAPLAN, R. Like it or Not, Facebook Is Now a Media Company. *The New York Times*. 2016. Disponível em <<https://www.nytimes.com/roomfordebate/2016/05/17/is-facebook-saving-journalism-or-ruining-it/like-it-or-not-facebook-is-now-a-media-company>>. Acesso em: 03 abril 2018.

CARNEY, M. As Uber fights to maintain its technology company classification in India, the rest of the world watches. *Pando*. 2015. Disponível em <<https://pando.com/2015/02/20/as-uber-fights-to-maintain-its-technology-company-classification-in-india-the-rest-of-the-world-watches/>>. Acesso em: 03 abril 2018.

CONSTANTINE, J. Facebook Starts Hosting Publishers' "Instant Articles". *TechCrunch*. 2015. Disponível em <<https://techcrunch.com/2015/05/12/facebook-instant-articles/>>. Acesso em: 03 abril 2018.

\_\_\_\_\_. Zuckerberg implies Facebook is a media company, just "not a traditional media company". *TechCrunch*. 2016. Disponível em <<https://techcrunch.com/2016/12/21/fbonc/>>. Acesso em: 03 abril 2018.

CUKIER, J. Why is Uber considered a technology company and a lot of people place it among tech giants like Google, Facebook, etc?. *Quora*. 2016. Disponível em <<https://www.quora.com/Why-is-Uber-considered-a-technology-company-and-a-lot-of-people-place-it-among-tech-giants-like-Google-Facebook-etc>>. Acesso em: 03 abril 2018.

, M.; HOLT, J.; SANSON, K.. Introduction: Making of a Revolution. In: *Distribution Revolution: Conversations about the Digital Future of Film and Television*. University of California Press, 2014. p. 1-17.

DEUZE, M. *Media work*. Polity, 2007.

DEVITO, Michael A. From editors to algorithms: A values-based approach to understanding story selection in the Facebook news feed. *Digital Journalism*, v. 5, n. 6, p. 753-773, 2016.

DIXON, C. BuzzFeed. *cdixon blog*. 2014. Disponível em <<http://cdixon.org/2014/08/10/buzzfeed/>>. Acesso em: 03 abril 2018.

D'ONFRO, J.. Facebook is telling the world it's not a media company, but it might be too late. *Business Insider*. 2016. Disponível em <<http://www.businessinsider.com/mark-zuckerberg-on-facebook-being-a-media-company-2016-8>>. Acesso em: 03 abril 2018.

DOUGHERTY, C; STEEL, E. YouTube Introduces YouTube Red, a Subscription Service. *The New York Times*. 2015. Disponível em <<https://www.nytimes.com/2015/10/22/technology/youtube-introduces-youtube-red-a-subscription-service.html>>. Acesso em: 03 abril 2018.

ETHERINGTON, D. Facebook exploring creation of its own original video content. *TechCrunch*. 2016. Disponível em <<https://techcrunch.com/2016/12/14/facebook-exploring-creation-of-its-own-original-video-content/>>. Acesso em: 03 abril 2018.

FEDERAL COMMUNICATIONS COMMISSION. In the matter of protecting and promoting the open Internet. 2015. Disponível em: <[https://apps.fcc.gov/edocs\\_public/attachmatch/FCC-15-24A1\\_Rcd.pdf](https://apps.fcc.gov/edocs_public/attachmatch/FCC-15-24A1_Rcd.pdf)>. Acesso em: 11 abril 2018.

\_\_\_\_\_. Cable television. 2016a. Disponível em: <<https://www.fcc.gov/media/engineering/cable-television>>. Acesso em: 11 abril 2018.

\_\_\_\_. Fact sheet: Updating media ownership rules in the public interest. Disponível em: <[http://transition.fcc.gov/Daily\\_Releases/Daily\\_Business/2016/db0627/DOC-340033A1.pdf](http://transition.fcc.gov/Daily_Releases/Daily_Business/2016/db0627/DOC-340033A1.pdf)>. Acesso em: 11 abril 2018.

FIEGERMAN, S. Dear Facebook, you're a media company now. Start acting like one. *Mashable*. 2016. Disponível em <<https://mashable.com/2016/05/15/facebook-media-company/#64EvZwrH0aqi>>. Acesso em: 03 abril 2018.

FIVEASH, K. "We're a tech company, we're not a media company," says Facebook founder. *ArsTechnica*. 2016. Disponível em <<https://arstechnica.com/tech-policy/2016/08/germany-facebook-edit-hateful-posts-zuckerberg-says-not-media-empire/>>. Acesso em: 03 abril 2018.

FOREMSKI, T. Stuff you didn't know about Google---a report from the 2003 Googleplex Xmas party. *SiliconValleyWatcher*. 2004. Disponível em <<https://www.siliconvalleywatcher.com/stuff-you-didnt-know-about-google---a-report-from-the-2003-googleplex-xmas-party/>>. Acesso em: 03 abril 2018.

FOX, J. Why It's Good to Be a "Technology Company". *Harvard Business Review*. 2014. Disponível em <<https://hbr.org/2014/08/why-its-good-to-be-a-technology-company>>. Acesso em: 03 abril 2018.

GILLESPIE, T. The politics of 'platforms'. *New media & society*, v. 12, n. 3, p. 347-364, 2010.

\_\_\_\_. Algorithmically recognizable: Santorum's Google problem, and Google's Santorum problem. *Information, Communication & Society*, v. 20, n. 1, p. 63-80, 2017.

\_\_\_\_. Regulation of and by platforms. In: BURGESS, J.; MARWICK, A.; POELL, T. (Ed.). *The Sage Handbook of Social Media*. SAGE, 2018.

GOTTFRIED, J.; SHEARER, E. News Use Across Social Media Platforms 2016. *Pew Research Center*. 2017. Disponível em: <<http://www.journalism.org/2016/05/26/>

[news-use-across-social-media-platforms-2016/](http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/)>. Acesso em: 03 abril 2018.

GUYNN, J. Twitter suspends alt-right accounts. *USA Today*. 2016. Disponível em <<https://www.usatoday.com/story/tech/news/2016/11/15/twitter-suspends-alt-right-accounts/93943194/>>. Acesso em: 03 abril 2018.

HELFT, M. Is Google a Media Company?. *The New York Times*. Disponível em <<https://www.nytimes.com/2008/08/11/technology/11google.html>>. Acesso em: 03 abril 2018.

HESS, Thomas. What is a media company? A re-conceptualization for the online world. *International Journal on Media Management*, v. 16, n. 1, p. 3-8, 2014.

INGRAM, M. Facebook's biggest problem is that it's a media company. *Gigaom*. 2012. Disponível em <<https://gigaom.com/2012/05/16/facebooks-biggest-problem-is-that-its-a-media-company/>>. Acesso em: 03 abril 2018.

ISAAC, M. Facebook, facing bias claims, shows how editors and algorithms guide news. *The New York Times*. 2016. Disponível em <<http://www.nytimes.com/2016/05/13/technology/facebook-guidelines-trending-topics.html>>. Acesso em: 03 abril 2018.

\_\_\_\_. Facebook, Nodding to Its Role in Media, Starts a Journalism Project. *The New York Times*. 2017. Disponível em <<https://www.nytimes.com/2017/01/11/technology/facebook-journalism-project.html>>. Acesso em: 03 abril 2018.

JARVIS, J. Facebook Needs an Editor. *Whither news?*. 2016. Disponível em <<https://medium.com/whither-news/facebook-needs-an-editor-b-388f789a9c3>>. Acesso em: 03 abril 2018.

KAFKA, P. How Jonah Peretti and BuzzFeed Got to \$850 Million. *RECODE*. 2014. Disponível em

- <<https://www.recode.net/2014/8/11/11629746/how-jonah-peretti-and-buzzfeed-got-to-850-million>>. Acesso em: 03 abril 2018.
- KITCHIN, Rob. Thinking critically about and researching algorithms. *Information, Communication & Society*, v. 20, n. 1, p. 14-29, 2017.
- KOETSIER, J. Exclusive: Cheezburger will take page out of Reddit's playbook, allow users to create own subsites. *VentureBeat*. 2012. Disponível em <<https://venturebeat.com/2012/08/22/exclusive-cheezburger-will-take-page-out-of-reddits-playbook-allow-users-to-create-own-subsites/>>. Acesso em: 03 abril 2018.
- KOKALITCHEVA, K. YouTube's First Lineup of Original Content Is Here. *Fortune*. 2016. Disponível em <<http://fortune.com/2016/02/10/youtube-originals-release/>>. Acesso em: 03 abril 2018.
- KRAMER, S. Google Is Still A Tech Company. Really. *Gigaom*. 2006. Disponível em <<https://gigaom.com/2006/06/12/google-is-still-a-tech-company-really/>>. Acesso em: 03 abril 2018.
- LANGER, A. Is Steve Jobs the God of Music?. *Esquire*. 2003. Disponível em <<https://www.esquire.com/news-politics/a111177/steve-jobs-esquire-interview-0703/>>. Acesso em: 03 abril 2018.
- LENTZ, Becky. Excavating historicity in the US network neutrality debate: An interpretive perspective on policy change. *Communication, Culture & Critique*, v. 6, n. 4, p. 568-597, 2013.
- MANSELL, R; RABOY, M (Ed.). *The handbook of global media and communication policy*. John Wiley & Sons, 2011.
- MANJOO, F. Facebook's bias is built-in, and bears watching. *The New York Times*. 2016. Disponível em <<https://www.nytimes.com/2016/05/12/technology/facebooks-bias-is-built-in-and-bears-watching.htm>>. Acesso em: 03 abril 2018.
- MCGRATH, R. 15 years later, lessons from the failed AOL-Time Warner merger. *Fortune*. 2015. Disponível em <<http://fortune.com/2015/01/10/15-years-later-lessons-from-the-failed-aol-time-warner-merger/>>. Acesso em: 03 abril 2018.
- MICKEY, B. Are You a Technology Company or a Media Company?. *Folio*. 2013. Disponível em <<http://www.foliomag.com/are-you-technology-company-or-media-company/>>. Acesso em: 03 abril 2018.
- MORRISSEY, B. BuzzFeed's Dao Nguyen: 'We don't think of ourselves as an island'. *DIGIDAY*. 2016. Disponível em <<https://digiday.com/podcast/buzzfeed-dao-nguyen-digiday-podcast/>>. Acesso em: 03 abril 2018.
- NADLER, A. M. *Making the news popular: Mobilizing US news audiences*. University of Illinois Press, 2016.
- NAPOLI, Philip M. *Foundations of communications policy: Principles and process in the regulation of electronic media*. Hampton Pr, 2001.
- \_\_\_\_\_. Bridging cultural policy and media policy. *The journal of arts management, law, and society*, v. 37, n. 4, p. 311-332, 2008.
- \_\_\_\_\_. *Audience evolution: New technologies and the transformation of media audiences*. Columbia University Press, 2011.
- \_\_\_\_\_. Social media and the public interest: Governance of news platforms in the realm of individual and algorithmic gatekeepers. *Telecommunications Policy*, v. 39, n. 9, p. 751-760, 2015.
- \_\_\_\_\_. Requiem for the long tail: Towards a political economy of content aggregation and fragmentation. *International Journal of Media & Cultural Politics*, v. 12, n. 3, p. 341-356, 2016.
- \_\_\_\_\_; CAPLAN, R. When Media Companies Insist They're Not Media Companies and Why It Matters for Communications Policy. SSRN. 2016. Disponível em <<https://papers.ssrn.com/sol3/papers>.

cfm?abstract\_id=2750148>. Acesso em: 03 abril 2018.

ELETRONIC FRONTIER FOUNDATION. New York v. Harris. 2012. Disponível em <<https://www.eff.org/cases/new-york-v-harris>>. Acesso em: 03 abril 2018.

NUNEZ, M. Former Facebook Workers: We Routinely Suppressed Conservative News. *Gizmodo*. 2016a. Disponível em <<https://gizmodo.com/former-facebook-workers-we-routinely-suppressed-conser-1775461006>>. Acesso em: 03 abril 2018.

\_\_\_\_\_. Want to know what Facebook really thinks of journalists? Here's what happened when it hired some. *Gizmodo*. 2016b. Disponível em <<http://gizmodo.com/want-to-know-what-facebook-really-thinks-of-journalists-1773916117>>. Acesso em: 03 abril 2018.

PASQUALE, Frank. *The black box society: The secret algorithms that control money and information*. Harvard University Press, 2015.

PATEL, N. The wrong words: How the FCC lost net neutrality and could kill the Internet. *The Verge*. 2014. Disponível em <<http://www.theverge.com/2014/1/15/5311948/net-neutrality-and-the-death-of-the-internet>>. Acesso em: 03 abril 2018.

PICARD, Robert G. *The economics and financing of media companies*. Fordham Univ Press, 2011.

PEREZ, S.; SHIEBER, J. YouTube Unveils YouTube TV, its live TV streaming service. *TechCrunch*. 2017. Disponível em <<https://techcrunch.com/2017/02/28/youtube-launches-youtube-tv-its-live-tv-streaming-service/>>. Acesso em: 03 abril 2018.

PIERSON, D. PARESH, D. If Facebook promotes propaganda, can it be a neutral news platform?. *Los Angeles Times*. Disponível em <<http://www.latimes.com/business/la-fi-tn-eu-tech-20160531-snap-story.html>>. Acesso em: 03 abril 2018.

, M; LANDRY, N. *Civil society, communication, and global governance: Issues from the World Summit on the Information Society*. Peter Lang, 2005.

RANA, S. Is Twitter a technology platform, a media company — or both?. *Tech2*. 2012. Disponível em <<http://tech.firstpost.com/news-analysis/is-twitter-a-technology-platform-a-media-company-or-both-29998.html>>. Acesso em: 03 abril 2018.

RINEHART, W. A semantic network analysis of the network neutrality debate. *SSRN*. 2015. Disponível em <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2587849](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2587849)>. Acesso em: 03 abril 2018.

ROBINSON, J. Merkel vows to regulate 'fake news' after blaming social media for the rise of populist politics and 'making it harder to form people's opinions than it was 25 years ago'. *Daily Mail*. 2016. Disponível em <<http://www.dailymail.co.uk/news/article-3964366/Merkel-vows-regulate-fake-news-blaming-social-media-rise-populist-politics-making-harder-form-people-s-opinions-25-years-ago.html>>. Acesso em: 03 abril 2018.

ROBERTS, J. Why Facebook won't admit it's a media company. *Fortune*. 2016. Disponível em <<http://fortune.com/2016/11/14/facebook-zuckerberg-media/>>. Acesso em: 03 abril 2018.

RUANE, Kathleen Ann. Fairness doctrine: History and constitutional issues. *Congressional Research Service*. 2011. Disponível em <<http://randomswede.org/pdf/fairness-doctrine-decline-R40009.pdf>>. Acesso em: 03 abril 2018.

SALLET, J. FCC transaction review: Competition and the public interest. *Federal Communications Commission*. 2014. Disponível em <<https://www.fcc.gov/news-events/blog/2014/08/12/fcc-transaction-review-competition-and-public-interest>>. Acesso em: 03 abril 2018.

SATELL, Greg. Tronc's data delusion. *Harvard Business Review*, v. 23, 2016.

SCOTT, M.; ISAAC, M. Facebook restores iconic Vietnam War photo it censored for nudity. *The New York Times*. 2016. Disponível em <<http://www.nytimes.com/2016/09/10/technology/facebook-vietnam-war-photo-nudity.html>>. Acesso em: 03 abril 2018.

SICHA, C. Inside Gawker Media's first company-wide meeting. *The Awl*. 2011. Disponível em <<https://theawl.com/inside-gawker-medias-first-company-wide-meeting-8abf673bf61#.umax6qp27>>. Acesso em: 03 abril 2018.

SLEFO, G. Here are some stats that put the whole 'duopoly of Google and Facebook' into perspective. *Advertising Age*. 2017. Disponível em <<http://adage.com/article/digital/google-facebook-digital-duopoly-stats-emarketer/308282/>>. Acesso em: 03 abril 2018.

SLOANE, G. Publishers and brands, get ready for the Snapchat algorithm. *Digiday*. 2016. Disponível em <<http://digiday.com/platforms/publishers-brands-get-ready-snapchat-algorithm/>>. Acesso em: 03 abril 2018.

SNIP, I. Facebook is now a media company. Bad? No. Different? Yes. *Medium*. 2016. Disponível em <<https://medium.com/topishare/facebook-is-now-a-media-company-bad-no-different-yes-bc0ca7f1b779>>. Acesso em: 03 abril 2018.

STEIN, B. Why Wait-A Discussion of Analogy and Judicial Standards for the Internet in Light of the Supreme Court's *Reno v. ACLU* Opinion. • *Louis ULLJ*, v. 42, p. 1471, 1997.

STREETER, T. The cable fable revisited: Discourse, policy, and the making of cable television. *Critical Studies in Media Communication*, v. 4, n. 2, p. 174-200, 1987.

SULLIVAN, D. Schmidt: Google still a tech company despite the billboards. *Search Engine Watch*. 2006. Disponível em <<https://searchenginewatch.com/>

[sew/news/2058565/schmidt-google-still-a-tech-company-despite-the-billboards](http://sew/news/2058565/schmidt-google-still-a-tech-company-despite-the-billboards)>. Acesso em: 03 abril 2018.

THIELMAN, S. Facebook fires trending team, and algorithm without humans goes crazy. *The Guardian*. Disponível em: <<https://www.theguardian.com/technology/2016/aug/29/facebook-fires-trending-topics-team-algorithm>>. Acesso em: 03 abril 2018.

THOMPSON, D. Facebook and fear. *The Atlantic*. 2016. Disponível em <<http://www.theatlantic.com/technology/archive/2016/05/the-facebook-future/482145/>>. Acesso em: 03 abril 2018.

THUNE, J. Letter to Mark Zuckerberg, Chairman and Chief Executive Officer, Facebook, Inc.. *Committee on Commerce, Science and Transportation*. 2016. Disponível em <<https://pt.scribd.com/doc/312153730/Sen-John-Thune-s-letter-to-Mark-Zuckerberg>>. Acesso em: 03 abril 2018.

TJAARDSTRA, N. Vox Media: Tech company first, publisher second. *World Association of Newspapers and News Publishers*. 2015. Disponível em <<http://blog.wan-ifra.org/2015/04/20/vox-media-tech-company-first-publisher-second>>. Acesso em: 03 abril 2018.

TRIELLI, D.; MUSSENDEN, S.; STARK, J.; DIAKOPOULOS, N. Googling politics: How the Google issue guide on candidates is biased. *Slate*. 2016. Disponível em <[http://www.slate.com/articles/technology/future\\_tense/2016/06/how\\_the\\_google\\_issue\\_guide\\_on\\_candidates\\_is\\_biased.html](http://www.slate.com/articles/technology/future_tense/2016/06/how_the_google_issue_guide_on_candidates_is_biased.html)>. Acesso em: 03 abril 2018.

TWITTER. Diretrizes e princípios dos Moments do Twitter. 2016. Disponível em <<https://help.twitter.com/pt/rules-and-policies/twitter-moments-guidelines-and-principles>>. Acesso em: 03 abril 2018.

WOOLLEY, S. *The Network: The Battle for the Airwaves and the Birth of the Communications Age*. Findaway World LLC, 2016.

**Philip Napoli** – Professor de políticas públicas na Duke University e membro do DeWitt-Wallace Center para mídia e democracia.

**Robyn Caplan** – Pesquisadora no instituto de pesquisa Data & Society e doutoranda em comunicação e informação na Rutgers University.

E-mail de contato: [willianfaraujo@gmail.com](mailto:willianfaraujo@gmail.com)



Entrevista

# Algoritmos como um devir: uma entrevista com Taina Bucher

## **Carlos d'Andréa**

Universidade Federal de Minas Gerais.  
Professor do Programa de Pós-Graduação em Estudos de Comunicação  
Belo Horizonte, MG, Brasil

## **Amanda Jurno**

Universidade Federal de Minas Gerais.  
Doutoranda pelo Pós-Graduação em Estudos de Comunicação  
Belo Horizonte, MG, Brasil

Perguntas, muitas perguntas instigantes. Parece que é usando pontos de interrogação que Taina Bucher olha para o “mundo empírico” e tenta entendê-lo. Ao interrogar os algoritmos e as diferentes formas de conhecimentos a eles associadas, nesta entrevista a pesquisadora nos convida a olhar com profundidade para esses objetos sociotécnicos contemporâneos. Jornalismo, a importância de programar (ou não) e a genealogia do termo “caixa-preta” são alguns tópicos abordados por Taina Bucher em diálogo com autores como Ross Ashby, Alfred N. Whitehead e Annemarie Mol. Através desta agradável conversa que se deu na cidade de Tartu, na Estônia, durante a Conferência da *Association of Internet Researchers - AoIR 2017*, somos convidados a assumir um modo de “devir” ao pesquisar os algoritmos, as práticas e as políticas associadas a eles. Taina Bucher é professora adjunta do Centre for Communication and Computing, na *University of Copenhagen* (Dinamarca) e autora do livro *“IF... THEN: Algorithmic power and politics”*, lançado em maio de 2018 pela Oxford University Press. Sua tese de doutorado (*Programmed Sociality: A Software Studies Perspective on Social Networking Sites*) foi vencedora do prêmio anual de teses da AoIR em 2013. Em sua pesquisa, Taina Bucher foca nas mídias sociais e no poder dos algoritmos na vida cotidiana, na interseção entre os estudos de *software*, STS e as teorias das novas mídias. Para mais informações, visite o site [tainabucher.com](http://tainabucher.com).

**Parágrafo:** O principal objetivo da pesquisa que você apresentou aqui, na conferência da AoIR, é a formulação de um sentido epistemológico dos algoritmos. Como você está enfrentando os desafios teóricos deste projeto de pesquisa?

**Bucher:** Há uma questão que realmente me interessa: “o que significa conhecer algo que é ‘invisível’, por definição distribuído, que não é uma ‘coisa’?”

Eu sempre estive interessada na questão: Como você conhece algo que não pode acessar ou que você não vê? O que você precisa saber para conhecê-lo um pouco mais? Durante meu doutorado, havia muitos debates sobre a necessidade de se saber, ou não, programar para dizer algo sobre uma tecnologia. Acho que não é a discussão mais proveitosa a longo prazo, e não acho que alguém necessariamente precise saber programar para poder dizer algo sobre o assunto. Definitivamente ajuda se você souber programar, mas essa não é a única maneira de conhecer objetos computacionais como os algoritmos. Eu também encontrei uma hierarquia intelectual entre “as pessoas que sabem” e “as pessoas que não sabem” programar. Eu realmente não gosto dessas afirmações sobre quem sabe mais, ou de forma mais autoritária. Para mim, foi muito útil aprender a programar, mas isso não significa que, por causa disso, eu necessariamente saiba como o Facebook funciona. Se realmente quisermos seguir esse argumento, também precisamos saber como o hardware funciona. Nós precisaríamos saber sobre estruturas de dados, sobre bancos de dados e sobre muitas outras coisas. Meus interesses estão nas questões epistemológicas: “o que significa conhecer? O que poderíamos estar perdendo em nossas tentativas de conhecer coisas específicas? Quais estão os limites do conhecimento? Onde estão as fronteiras?” Em um capítulo do meu livro (*“IF... THEN: Algorithmic power and politics”*. Oxford University Press, 2018<sup>1</sup>), tentei “desempacotar” a imagem da “caixa-preta”. Porque sempre que se entra em discussões sobre algoritmos ou afirmações semelhantes de conhecimento, as pessoas dizem: “oh, os algoritmos são uma ‘caixa-preta’; portanto, não podemos conhecê-los; ou, portanto, é difícil conhecê-los”. É claro que é difícil, mas para mim, novamente, o interesse está em tentar desempacotá-los. O que realmente significa dizer que algo “é uma caixa-preta”?

1 Link para o livro: <https://global.oup.com/academic/product/iftthen-9780190493035?lang=en&cc=us>

**Parágrafo:** Você está fazendo uma espécie de genealogia do conceito de caixa-preta, talvez “desencaixapretando a caixa-preta”? Quais seriam os principais autores que já discutiram isso anteriormente?

**Bucher:** Este conceito vem da Cibernética, do início da Ciência da Computação. As pessoas tinham que tornar suas tecnologias “caixas-pretas”, caso contrário seus inimigos na Segunda Guerra Mundial seriam capazes de decodificar seus sistemas. Então, a tecnologia crucial era na verdade “encaixapretada” para não se perder a guerra. Como muitas outras tecnologias computacionais, o termo tem uma genealogia militar. Na Ciência da Computação, tornar a complexidade mais palpável é uma necessidade. Mas então esta se torna uma metáfora para tudo, tudo é uma “caixa-preta”... atualmente, “caixa-preta” é apenas um sinônimo para tudo que é desconhecido, tudo que pode ser ou deveria ser escondido, como segredos comerciais. Para mim foi muito útil ler o artigo do Peter Galison “*The ontology of the enemy: Norbert Wiener and the cybernetic vision*”<sup>2</sup>. E, quanto à cibernética, W. Ross Ashby escreve diretamente sobre a “caixa-preta”<sup>3</sup>. E então, claro, Bruno Latour escreve sobre “caixa-preta” e “encaixapretamento” como um conceito. Para Latour, tudo é uma espécie de “caixa-preta”, porque tudo esconde a sua própria fabricação, sua ontologia em rede<sup>4</sup>. Ele usa uma antiga visão heideggeriana sobre “se a tecnologia quebra, então você vê como ela funciona”. Para ele, é aí que o conceito de “caixa-preta” é útil. Alexander Galloway também escreveu sobre a história e a epistemologia da “caixa-preta” de uma maneira muito útil. Assim, Galison, Ashby, Latour e Galloway seriam um bom começo para se pensar sobre a genealogia do conceito.

2 Artigo disponível em [www.jstor.org/stable/1343893](http://www.jstor.org/stable/1343893). Accessed 9 mar. 2018.

3 Por exemplo, ver o capítulo 6 do seu livro “*An Introduction to cybernetics*”, publicado pela primeira vez em 1956.

4 Entre outros, ver o livro do Latour “*Ciência em Ação*”, publicado pela primeira vez em 1987 pela Harvard University Press, em inglês.

**Parágrafo:** Como você está articulando autores como Annemarie Mol e Whitehead para estudar os algoritmos?

**Bucher:** O que é útil para mim está, novamente, amarrado à noção de uma “caixa-preta”. Porque se você olhar para algo por uma perspectiva relacional, não há caixa e ela não é nunca só preta. Assim, torna-se um pouco difícil ou impossível operar com a afirmação de que “algoritmos são ‘caixas-pretas’”. Com Whitehead<sup>5</sup>, acho útil pensar em um algoritmo nunca como uma ‘coisa’, mas como uma fotografia temporária da “coisa” naquele momento. Então, ele diria que, seja lá o que for a “coisa”, ela está “se tornando”. Eu estive pensando sobre o porquê de, agora, tantas pessoas estarem interessadas em Whitehead de novo, e acho que é porque a filosofia do processo ressoa com coisas que estão tão mudando o tempo todo. Um algoritmo ou sistema de aprendizado de máquina é, por definição, não uma “coisa”, é sempre um devir (em inglês, *‘a becoming’*). Para mim, essa noção de devir é bastante útil, porque é também um ponto metodológico que, seja o que for que estudemos, trata-se de algo em seu devir. Então, não estamos fazendo afirmações sobre o algoritmo do Facebook ou o algoritmo do YouTube porque, seja para o que for que estivermos olhando, será uma atualização de diferentes tipos de relações que se juntam ali para então moldar essa realidade. Para mim, em termos de ontologia, essa tem sido uma maneira útil de pensar sobre o que são os algoritmos. E então, é claro, isso se encaixa em questões epistemológicas. Sobre Annemarie Mol, fiquei bastante fascinada pelo seu livro “*The Body Multiple*”<sup>6</sup>. Há milhões de coisas para se dizer sobre ele... Ela escreve de uma forma muito interessante, o livro tenta romper com limites acadêmicos muito

5 Alfred North Whitehead (1861-1947) foi um matemático, lógico e filósofo britânico mais conhecido pelo seu trabalho sobre lógica matemática e filosofia da ciência.

6 Annemarie Mol é professora de Antropologia do Corpo na Universidade de Amsterdã e uma importante pesquisadora no campo de TAR/STS. Seu livro “*The Body Multiple: Ontology in Medical Practice*” foi lançado em 2003 (Duke University Press).

rígidos e se envolver com o mundo através, também, da escrita. Por ela ser uma antropóloga, tem essa inserção a longo prazo em um espaço onde pode contar aquela história de uma maneira muito convincente. Sobre seu pensamento, tem sido muito útil pensar no corpo, e especificamente nas doenças, como tendo diferentes realidades, múltiplas realidades. O que estou achando inspirador é essa ideia do múltiplo; até que ponto nós estamos estudando a mesma “coisa”, os mesmos fenômenos? Então, eu também estou pensando no algoritmo como múltiplo. O que é o algoritmo do *Facebook*? É uma “coisa” ou é apenas um nome? Isso significa coisas diferentes? Na verdade, o algoritmo é tecnicamente bastante distribuído, mas também é muito distribuído socialmente. Ele vive diferentes tipos de vidas. E nós precisamos ser bem específicos sobre qual tipo de vida em contexto que estamos falando. É aí que o trabalho dela está sendo bastante inspirador, mas também os estudos de John Law, especialmente o livro sobre as histórias das aeronaves<sup>7</sup>. Ele constrói o mesmo tipo de argumento, ele conta histórias diferentes sobre esse único objeto e teoriza de uma forma muito semelhante como sendo, na verdade, múltiplos objetos.

**Parágrafo:** Você fala que seu foco não está em “onde” a agência está localizada, mas em “quando”. Isso está relacionado à abordagem que enfatiza a performance dos algoritmos, certo? E parece ter algo a ver com a necessidade de descobrir onde ou quem é o responsável quando algo errado acontece. Se a agência dos algoritmos é distribuída, talvez seja impossível descobrir quem fez o quê em relação a alguma coisa...

**Bucher:** Certamente, e essa é a razão pela qual eu acho a questão do “quando” particularmente

importante. Quando falamos sobre *accountability* algorítmica, transparência, discriminação, parcialidade, etc., nós estamos muito obcecados em encontrar a origem certa de culpa, a origem da ação. Nós ainda precisamos compreender as responsabilidades. Ao perguntar “quando” alguma coisa acontece ou “quando” uma agência é tornada mais ou menos disponível, não é para nos livrar de responsabilidades, mas para reconhecer que nem sempre importa “onde” a agência está, certo? Porque eu não acho que podemos necessariamente chegar à origem certa. O que importa é o tipo de trabalho envolvido para fazer com que essas origens certas apareçam como tais em primeiro lugar. Com carros autônomos, por exemplo, não tenho certeza se estamos realmente resolvendo o problema ao insistir em ir ao carro, ao fabricante do carro, à pessoa ou a um sistema... Pode ser mais construtivo e, na verdade, mais prático olhar para “quando” alguém ou alguma coisa é tida como mais ou menos provável como origem da ação. O tipo de trabalho que disponibiliza a origem das ações é político, assim, acho que podemos aprender muito a partir dessas práticas. Quem quer colocar a culpa em alguém? Por exemplo, olhe para o *Facebook* ou o *Google*: sempre que algo dá errado, o seu discurso é muito estrategicamente construído... Mas o que está em jogo quando o *Facebook* ou o *Google*, em uma dada situação, dizem “é apenas culpa dos algoritmos, nós não fizemos nada”? E, às vezes, é como: “Sim, você sabe, foi isso... Não culpem os nossos programadores”. Há uma ignorância formal estratégica envolvida, usa-se o não-conhecimento de forma muito estratégica. Isso que está em jogo, então, para mim, é bastante interessante perguntar “quando” essas origens ou esses elementos são mobilizados, e “para quem”. Eu também acho interessante pensarmos por que estamos incomodados em determinados momentos, quando não estamos em

<sup>7</sup> Referência ao sociólogo John Law, um dos principais proponentes da Teoria Ator-Rede, e seu livro “*Aircraft Stories: Decentering the Object in Technoscience*”, publicado em 2002 pela Duke University Press.

outros. Há coisas que talvez nós não estamos discutindo. Nós tendemos a falar sobre algumas das mesmas coisas. Então, para mim, é interessante perguntar “quando” falamos delas e “de que forma” elas tendem a emergir. Porque elas só importam às vezes. Eu estou interessada nesses “às vezes”.

**Parágrafo:** Você tem alguns trabalhos recentes sobre jornalismo e algoritmos. Hoje, é realmente desafiador para as instituições jornalísticas lidar com todos os tipos de mediações algorítmicas, porque elas enfrentam as mediações tradicionais de jornalismo. Como essas pesquisas empíricas dialogam com suas questões mais amplas?

**Bucher:** Eu estava pensando: “onde os algoritmos são importantes? Em que tipo de contextos eles desafiam organizações de trabalho existentes, seus modos de pensar, seus sistemas de valores?” Há, naturalmente, o setor financeiro onde os algoritmos são imensamente importantes, e há também o comércio, a biotecnologia e as ciências duras. Mas, vindo do campo de mídias e comunicação, eu também estava vendo que os algoritmos importam para o jornalismo e para as organizações jornalísticas. E, então, eu me interessei em olhar para como a tecnologia funciona. Qual é o papel dos algoritmos? Qual tipo de trabalho os algoritmos fazem? Eu queria olhar para a materialidade, para os usuários, para as práticas de trabalho e ver como eles importam. É bastante útil entrar em contextos muito concretos, sobre como as organizações jornalísticas lidam com algoritmos, métricas e análise de dados. Os leitores têm novas demandas e expectativas. As organizações jornalísticas tiveram que lidar com o fato de que as pessoas apenas usam as redes sociais. Elas se acostumaram

a ter informações mais personalizadas, mais em tempo real, elas se acostumaram a mais curadorias de informações e assim por diante. E, é claro, também há o dinheiro de publicidade para as organizações jornalísticas. Plataformas de mídia como o *Google*, o *Facebook* e o *Twitter* se tornaram concorrentes. As organizações jornalísticas não estão mais competindo com outros jornais nacionais, seus principais concorrentes são agora as empresas de tecnologia. Esta nova realidade existe há pelo menos cinco anos, mas esses tipos de preocupações e discussões estão apenas se intensificando. Eu achei que esse seria um momento oportuno para estudar a espécie de “impacto” desses novos tipos de sistemas sociotécnicos. Talvez em cinco anos isso seja esquecido ou as pessoas não consigam chegar ao centro da controvérsia. Mas agora ela está acontecendo, então é um excelente momento para ver quais são as preocupações. Eu visitei algumas organizações jornalísticas e conversei com jornalistas e editores, observei a sua rotina de trabalho. Falei com os programadores sobre essas mudanças e sobre o que eles acham que está acontecendo. Muitas dessas organizações também estão desenvolvendo seus próprios sistemas. Assim você não precisa necessariamente ir ao *Facebook* e estudar como o código é feito, mas você pode ir às organizações jornalísticas. Estou falando apenas sobre o contexto escandinavo, e eu me concentrei um pouco mais em olhar o desenvolvimento interno dos algoritmos e dos sistemas. A maior preocupação é realmente em como concorrer sem perder sua integridade jornalística. Como você pode se orientar por uma lógica algorítmica no *Facebook*, sem perder a credibilidade, os valores jornalísticos? Existe um equilíbrio entre o que realmente funciona e como isso é negocia-

do no dia a dia. Eu encontrei muitas pessoas debatendo-se com a melhor forma de se orientar estrategicamente pela lógica de outros e em como ser mais esperto para usá-la em benefício próprio, mantendo aquele valor jornalístico. Isso é muito interessante para mim. Como isso é negociado? Isso é possível? Que tipo de compromissos as organizações jornalísticas assumem nesse jogo? Elas perdem? Elas encontram novas formas mais inteligentes e bastante produtivas para lidar com isso?





# Artigos Livres

# **Relações públicas e comunicação organizacional digital: empresas fronteiriças e suas estratégias\***

## **Public relations and organizational digital communication: companies on the border and their strategies**

### **Henrique Esper**

Universidade Federal do Pampa  
Bagé, RS, Brasil

### **Cristóvão Domingos de Almeida**

Universidade Federal do Pampa.  
Bagé, RS, Brasil

\*A base deste texto foi apresentada na Divisão Temática Comunicação Organizacional e Relações Públicas, da Intercom Júnior – XIII Jornada de Iniciação Científica em Comunicação, evento componente do 40º Congresso Brasileiro de Ciências da Comunicação, 4 a 9 de setembro em Curitiba/PR.

## Resumo

O objetivo deste trabalho é compreender a comunicação organizacional e as estratégias digitais da Associação Comercial e Industrial de Prestação de Serviços e Agropecuária de São Borja (ACISB), localizada na fronteira oeste do estado do Rio Grande do Sul. Com foco nos usos e apropriações, o estudo se fundamenta nos conceitos de relações públicas, comunicação organizacional, redes sociais digitais, estratégias de comunicação e organização como espaço de fortalecimento das atividades comerciais. A partir de registros das publicações na página no facebook e entrevista, analisamos as estratégias e as formas de aproximação com os públicos, demonstrando a necessidade que as empresas têm nos dias atuais de se comunicar com seus públicos pelos meios digitais. Evidencia-se que as redes sociais podem servir como potencialidade de desenvolvimento do empreendimento e de aproximação dos diferentes públicos.

## Abstract

This paper aims to understand the digital communication strategies of the Associação Comercial e Industrial de Prestação de Serviços de São Borja - ACISB, located on the western border of the state of Rio Grande do Sul. With the focus on uses and appropriations, the study works with the concepts of public relations, organizational communication, digital social networks, communication strategies and organization as a space for the strength of commercial activities. Based on comments and information published on the Facebook page and interview, we analyze the strategies and the ways of approaching the public, indicating the necessity for companies in the present day to communicate with their audiences through digital media. It makes clear that social networks can serve as a potentiality for the development of the business and the approximation of different audiences.

## Palavras-chave

Relações públicas. Comunicação organizacional. Estratégias. Redes sociais. ACISB.

## Keywords

Public relations. Organizational communication. Strategies. Social networks. ACISB.

## Introdução

O artigo se apoia no argumento que nos dias atuais as redes sociais são parceiras das instituições, pois a partir desses espaços digitais as organizações conseguem manter diálogo com seus públicos, aproximações e estabelecer processos de fidelização.

A internet, na Era das Relações que engloba tecnologias, criatividade, responsabilidade social, conexões transpessoais e respeito à diversidade cultural (MORAES, 2007), tem produzido mudanças nos hábitos, costumes, comportamentos, colaborando, de algum modo, para auxiliar na identificação da identidade humana. Nesse sentido, as ferramentas tecnológicas, cada vez mais presente no nosso cotidiano, flexibilizam os processos materiais, físicos, simbólicos, numa dinâmica que requer intuição, pensar crítico, garantia dos princípios éticos, afetivos e diálogo com as situações reais e práticas, ou seja, na era das relações, o desenvolvimento tecnológico, requer, centralidade na comunicação.

Este estudo traz as questões da comunicação organizacional digital no município de São Borja, com intuito de registro, para dar visibilidade às estratégias de comunicação digital da Associação Comercial e Industrial de Prestação de Serviços e Agropecuária (ACISB), que diretamente auxilia administradores locais, apresentando dados para reflexão, projetos empresariais, modelo de gestão, legislação, informações sobre a concorrência, ou seja, circula conteúdos que mostram situações-problema e apresentam os cenários promissores para se desenvolver no pampa gaúcho.

A abordagem metodológica deste estudo é de cunho qualitativo. Na coleta das informações, utilizamos registros das publicações, entrevistas semidirigidas com a equipe responsável da área de relacionamento e realizamos o mapeamento da presença da associação no Facebook. Já no tratamento dos dados, trabalhamos com a análise focada no “grau da conexão de uma rede ego” (RECUERO,

2011, p. 71) que se traduz em: analisar a popularidade de uma página através de suas conexões com os seguidores da página, por curtidas, comentários dentre outras interações entre os usuários.

Evidencia-se que as redes sociais podem servir como potencialidade de desenvolver o empreendimento e aproximar os diferentes públicos, quando usadas a partir de estratégias tomadas para melhorar o desempenho, a circulação de informações aos associados, o desenvolvimento da compreensão entre gestores e usuários das redes sociais e novos conhecimentos empresariais. Essas dimensões auxiliam no surgimento de novas identidades, nova cultura organizacional e resultados satisfatórios às organizações.

## Redes sociais digitais e identidade

Consideramos aqui que as redes sociais vêm provocando mudanças socioculturais que envolvem: descrição, execução e reflexão. E é objeto de estudo de diversos autores, dentre eles, Castells (2008) qual afirma que a formação de redes gera flexibilidade. Isto é, tornam praticamente indistintas as fronteiras de participação e de envolvimento, individualizam as relações sociais de produção e provocam instabilidade estrutural do trabalho, do tempo e do espaço. Essas dimensões também recebem atenção de Santaella (2010) e Terra (2011), como ruptura das variáveis clássicas da comunicação e das bases materiais da vida (CASTELLS, 2008), possibilitando a circulação das mensagens, os registros por longos períodos e a presença em lugares onde não se está fisicamente, para além das fronteiras geográficas.

Fragoso, Recuero e Amaral (2011) falam das redes sociais como alcance de sabedoria objetiva, ou seja, as plataformas digitais contribuem com a evolução do conhecimento. Por sua vez, a tecnicização (RODRIGUES, 1990) da internet naturaliza as relações entre os indivíduos e as máquinas. Muito embora, é preciso sair da era material, técnica, que

se registra a ausência de cooperação, para aportar na era das relações, isto é, integração, cooperação, respeito, consciência e diálogo. Castells (2009) fala dos limites das redes globais, aborda o poder das influências transformadoras em sua capacidade maleável de construção e estagnação, sendo totalmente dependente de forças externas para que haja mudanças.

As redes não têm limites estabelecidos; elas são ilimitadas e têm várias bordas, e sua expansão ou contração depende da compatibilidade ou competição entre os interesses e valores programados em cada rede e os interesses e valores programados nas redes com os quais elas entram em contato em seu movimento de expansão. (CASTELLS, 2009, p.65)

As redes sociais digitais funcionam como o exemplo de aparelho, proposto por Flusser (1985). Para Flusser (1985):

Aparelhos são caixas pretas que simulam o pensamento humano, graças as teorias científicas, as quais, como o pensamento humano, permutam símbolos contidos em sua memória, em seu programa. São caixas pretas que brincam de pensar.(p. 17)

Assim como a rede de comunicação global que tem sua realidade em evolução e sua complementaridade guiada dependente dos produtos da ação humana.

Nesse cenário, Castells (2015) propõe olharmos a sociedade global em redes, uma vez que vivenciamos o avanço das expressões fortes de identidade coletiva que reforçam o modelo de globalização, fortalecendo pequenos grupos elitistas, fundamentados por princípios e valores tradicionais. Mas também, aquelas que estão na contra-mão da globalização e provocam os desafios, indo contra as separações, as rupturas, as dualidades e a homogeneização do pensamento. As fontes

históricas da formação de cada identidade são diversificadas e incorporam movimentos voltados à transformação das interações humanas.

Castells (2008) tece um panorama das principais identidades e as classificam em três grupos: identidade legitimadora, identidade resistência e identidade de projeto. Para o autor a identidade legitimadora é introduzida pelas instituições dominantes da sociedade. Esse processo é que dá origem a sociedade civil. Para Castells (2008), a identidade de resistência se traduz nas minorias, aos desfavorecidos, aos desvalorizados pela lógica de dominação, promovendo resistências ao que lhes são disponibilizados, se por um lado tentam a imposição de mensagens, por outro, tem-se a resignificação e o pensamento crítico. Por sua vez, a identidade de projeto busca mudanças da estrutura cultural, através da construção de projetos de vida, isto é, ela passa pela identidade de resistência, mas visa às lutas enquanto construção de uma nova identidade, com o propósito de redefinir o seu espaço, promover parcerias, compartilhar informações entre os diferentes grupos, ou seja, o sujeito com essa nova identidade se torna capaz de avaliar, reunir e organizar as informações.

Nessa perspectiva, entendemos que o usuário das redes sociais é o centro do processo decisório, com capacidade de julgar, examinar, comparar, divergir, mesmo diante da flexibilidade (CASTELLS, 2008), pois se deve potencializar as habilidades comunicacionais, as tomadas de decisões decorrente das suas vivências em grupo, no trabalho, na família, na escola, entre outros espaços sociais. Diante dessa abordagem, definimos, neste estudo que as redes sociais digitais fomentam a construção de novas identidades, marcadas pela facilidade de acesso, flexibilização de tarefas, processos democráticos, de cooperação, de parcerias, de diversificação de plataformas e dispositivos. A identidade empresarial surge com o intuito de gerar visibilidade, aproximação e interação com seus públicos (TERRA, 2011), apoiando-se nas redes sociais como uma via de mão dupla (GRUNIG e HUNT, 2003).

## Redes estratégicas

As empresas estrangeiras, nacionais, regionais e locais têm procurado cada vez mais as tecnologias que lhes ofereçam facilidade e agilidade para a realização de tarefas, seja nos processos mais simples como a limpeza diária de seu prédio, até os processos mais detalhados e complexos, como a comunicação com os diferentes públicos. Diante disso, Castells (2003), afirma que os empresários descobriram na internet um novo planeta, povoado por inovações tecnológicas extraordinárias. Ou seja, diversas opções de atuação, trabalho, cooperação, divulgação dentro das redes sociais digitais. A efetivação dessas ações depende das organizações para se mobilizar e produzir novas transformações. Para Corrêa (2009, p. 163)

A digitalização na sociedade contemporânea é um fato. Difícil pensarmos nosso cotidiano sem a presença de algum dispositivo que transforma em bits simples operações como olhar as horas ou assistir à televisão. A digitalização hoje ocorre em rede mundial, conectando computadores, dispositivos e, especialmente, pessoas. Seja na forma mais imperceptível (quando realizamos uma transação por meio de um cartão magnético, por exemplo), seja na mais explícita (quando nos conectamos à internet), a sociedade conectada se faz presente e natural.

O argumento de Corrêa (2009) nos ajuda a fazer a conexão com o contexto organizacional, havendo mútua compreensão, cooperação, parceria e adaptação da organização com as novas tecnologias. A naturalização também precisa estar presente, sendo agente de mudanças, tornando a socialização da mensagem um processo comum, claro, objetivo, sem dualidades e nem estagnações dos conteúdos. Os gestores estão cada vez mais à procura de tecnologias capazes de gerar resultados positivos para seus empreendimentos. Dentre as possibilidades ofertadas pela internet estão as redes sociais digitais, no meio de muitas utiliza-

des, elas se inserem no mercado para auxiliar nos fluxos comunicacionais e promover as conexões, interações, cooperações e parcerias quando as usam de forma estratégicas.

Nessa perspectiva, as plataformas presentes na internet são vistas como aliadas da comunicação organizacional, que une o todo com base em suas partes (BALDISSERA, 2008). Para Terra (2011, p. 22), que observa o panorama da mídia, unindo a comunicação organizacional digital, entende que:

[...] Seguindo a evolução da comunicação de massa, passando dos meios impressos aos eletrônicos e, mais recentemente aos digitais, a comunicação organizacional incorpora uma vasta lista de ferramentas que vão da intranet à televisão via satélite, agora a televisão digital. Aos blogs, microblogs, chats, podcasts, entre outras. A soma dessa ferramenta digital que informa, treina e motiva públicos ligados à organização é o que se denomina comunicação organizacional digital. (TERRA, 2011, p. 22)

Desse modo, mesmo as organizações afastadas dos grandes centros urbanos, como é o caso das Empresas localizadas na região de fronteira, que possuem seus próprios hábitos em sua cultura organizacional (SCHULER, 2009), começam a se preparar para trabalhar de forma estratégica com as ferramentas disponíveis no mundo digital. Isso exige criatividade, conscientização, abertura crítica e tomar decisões que potencializam a interação entre emissor e receptor, enquanto produtores dinâmicos do processo comunicacional. Entendemos aqui que a organização inserida no mundo digital é fundamental para sua sobrevivência, tendo em vista, que nos dias atuais o ambiente é “competitivo, incerto e exigente” (SILVA, SANTOS E SILVA, 2017, p. 167).

## Empresa Fronteiriça

O mapeamento da comunicação digital foi feito na cidade de São Borja, localizada na fron-

teira oeste do estado do Rio Grande do Sul, cidade interiorana, com 335 anos de história, cerca de 60 mil habitantes e conhecida por ser a terra natal de dois ex-presidentes, Getúlio Vargas e João Goulart.

A instituição escolhida para a coleta de dados é a Associação Comercial e Industrial de Prestação de Serviços e Agropecuária (ACISB), por sua grande visibilidade no município e forte atuação na área empresarial. A ACISB<sup>1</sup> promove vários serviços aos seus mais de trezentos associados, os quais têm como principal característica oferecer benefícios diferenciados aos usuários, como custo reduzido, maior prazo, vantagens adicionais e atendimento personalizado. Por seus próprios meios ou através de convênios e parcerias com outras empresas oferece serviços como envio e recebimento de fax, xerox, linha em comodato da operadora de telefonia celular, eventos de integração e serviços médicos através de convênios particulares.

Segundo a equipe de relacionamento da associação, a máquina de fax que está em manutenção, era disponibilizada aos usuários e associados. O envio e recebimento de fax, foi por muito tempo utilizado pelos administradores, mas com o desuso do equipamento, apenas empresas dependentes de terceiros, que aceitam somente esta forma de envio utilizam o equipamento, cerca de 2 ou 3 empresas, entre as 319 associadas. O xerox é mais utilizado para o público em geral, porque as empresas geralmente possuem o seu próprio equipamento.

Em sua sede, comporta uma extensão do Centro de Integração Empresa Escola - CIEE e uma Unidade da Parceiros Voluntários. Junto ao Centro Unificado de Fronteira, a chamada Aduana, entre São Borja / Brasil e San Tomé / Argentina, comporta diversas empresas que lidam com exportação e importação, a entidade mantém

uma Unidade de Emissão de Certificado de Origem, a qual disponibiliza seus serviços às empresas exportadoras.

Além disso, a ACISB conta também com um espaço destinado à realização de eventos, localizado no parque de exposições da cidade, o qual é disponibilizado aos associados e entidades parceiras. Além da representatividade e integração com o setor empresarial, a entidade promove e participa de ações sociais que ajuda no bem-estar e nas melhorias da qualidade de vida aos moradores do município. A Associação possui funcionários no regime CLT, estagiários, voluntários e equipe diretiva, composta por mais de vinte membros, entre: presidente, secretários, fiscais e diretores: Comercial, Empresarial, Jurídico, Administrativo, Patrimônio, Responsabilidade Social, Relações Internacionais e Produtos e Serviços. Com essa estrutura organizacional, a ACISB está investindo em ações estratégicas, tais como, o lançamento do aplicativo *SmartCity*, plataforma digital que aproxima a empresa do consumidor; Campanha "comprar em São Borja é bom demais", numa tentativa de conscientizar os clientes a investirem nos produtos locais, bem como o curso de Gestão da Comunicação Criativa, ações visam criar, refletir, examinar e fidelizar os consumidores e ampliar o número de associados.

## **Estratégias Comunicacionais na Empresa Fronteiriça**

Apoiados no pensamento de Rodrigues (1990):

A imposição de uma ordem axiológica, a sanção positiva da sua observância e a sanção negativa da sua violação fazem intervir um modo de funcionamento específico dos campos sociais, um modo de funcionamento estratégico. (p. 150)

Diante disso, distinguimos dois tipos de estratégias: pedagógica e terapêutica. A estratégia

<sup>1</sup> [www.acisb.com](http://www.acisb.com)

pedagógica é um processo contínuo de instituir a ordem, fazer com que a ação realizada seja correta e produtiva. A estratégia terapêutica consiste no restabelecimento da ordem violada, utilizando-se de práticas de incorporação inconsciente de padrões de culpa. Quando diversos campos atuam ao mesmo tempo, necessitam repartir e aumentar a força de sua intervenção para se destacar.

Para que isso ocorra, utiliza-se a cooperação e o conflito. A cooperação consiste em um ou mais campos para se intervir juntos, de forma amplificadora, para anular apenas um tipo de padrão de valores dominante, fazendo da neutralidade o valor da sua estrutura e do seu funcionamento. Ao contrário desta proposta está o conflito, baseado na busca pela ordem axiológica de dois ou mais campos na mesma esfera de experiência, ou seja, na busca pelo topo, de forma emergente, sair do patamar de subordinado pelos tradicionais para desfrutar do poder e ter acesso a mecanismos de subordinação. Através de práticas, rompendo os discursos e comportamentos tradicionais anteriores já naturalizados nestas esferas.

Os hábitos produzidos nas ações comunicacionais das empresas dizem muito sobre o posicionamento da gestão na procura por novas tecnologias que facilitam processos e também sobre como os públicos recebem as informações dos empreendimentos locais. Para Rodrigues (1990, p. 149) "as dimensões de um campo social funcionam, sobretudo, sincronicamente enquanto estratégias funcionais que se articulam, no presente, com os procedimentos táticos dos agentes e dos actores sociais". Utilizando o exemplo da via de mão dupla, o receptor das mensagens neste caso, os associados da ACISB, possuem seu modo de se comunicar e receber informações, o emissor, neste caso a ACISB, está atenta aos meios digitais, por ser a plataforma mais utilizada tanto pelos associados e parceiros quanto pelos clientes. Um dos diretores salienta essa mudança, ao nos informar que antes, a ACISB

enviava e-mail e uma quantia considerável davam retorno, com o ingresso das plataformas digitais "enviamos e-mail e tínhamos que ligar para falar da mensagem enviada. Agora criamos um grupo no whatsapp e a comunicação ficou mais rápida e acessível". Essa percepção e mudança também ocorreram nas ações estratégicas da instituição, ao utilizar o facebook, com foco nos objetivos e monitoramento das mensagens.

Apresentamos aqui as ações comunicacionais da ACISB como espaço de comunicação com os seus públicos através do site e do facebook. Essas estratégias oferecem acolhida, protagonismo e incentivo às empresas de São Borja na fidelização, engajamento e trocas comunicacionais entre os associados e os públicos em geral.

## **As redes sociais como fomentadora da identidade Empresarial**

Em busca do desenvolvimento regional empresarial e industrial, a associação executa de forma estratégica suas ações comunicacionais digitais, para obter resultados positivos e fomentar a nova identidade empresarial da região. Com isso, a ACISB utiliza o site e as redes sociais para informar, comunicar, divulgar e compartilhar as mensagens de interesse dos associados, da comunidade de São Borja e dos municípios do entorno: Itaqui, Santo Ângelo e Uruguaiana.

As plataformas digitais possibilitam que as estratégias ganhem outras dimensões: agilidade, rapidez na propagação, criatividade, economia, ou seja, a ACISB, ao utilizar a comunicação organizacional digital, por intermédio do site e das redes sociais digitais, supera a Era Material e ingressa de vez na Era das Relações (MORAES, 2008), isto é, se antes o uso de equipamentos tradicionais era o mais viável, atualmente, a integração, a cooperação, o diálogo e o desenvolvimento das relações passam pela facilidade em

identificar as fontes de informação e a clareza das mensagens divulgadas.

Segundo a equipe de relacionamento, a instituição utiliza: a página do facebook e o aplicativo para dispositivos móveis whatsApp. A página do facebook para divulgação de informações aos usuários da plataforma e o whatsApp, a partir de um grupo criado para os colaboradores, para facilitar a comunicação e descontração do ambiente de trabalho, fatores cruciais para o surgimento de tecnologias emergentes, não lineares, instantâneas, com qualidade, num menor custo, acessível, interativo e com potencial de integração (PINHO, 2002). Vale dizer, que o whatsApp atravessou as barreiras corporativas e está modificando a forma de se comunicar nas organizações, abrangendo um novo cenário para o consumo de mídia e de visibilidade pública (ALMEIDA, 2017). As ações desenvolvidas pela ACISB ingressam na dimensão proposta por Rocha (2009, p. 269) a qual afirma que "consumir, hoje, é consumir cultura midiaticamente mediada, digitalmente interligada, imaginariamente compartilhada, imagetivamente realizada".

Ao analisar o site da Associação, na tela principal, conforme a imagem 1, além das informações básicas, como contato, horário de funcionamento, local e a opção de associar-se, possui de forma clara e objetiva, acesso a links sobre informações históricas da ACISB e da cidade. Estratégias comunicacionais já realizadas pela associação, como páscoa solidária, campanha de natal, palestras, café semanal com os associados, entre outras atividades. Há também conteúdos com foco na capacitação dos gestores, soluções administrativas para empresas, como locação, plano de saúde, representatividade da ACISB e temas atuais, por exemplo, as informações sobre a mudanças nas Reformas trabalhistas, conteúdo que atinge diretamente às empresas associadas.



Figura 1: Tela principal do Site da ACISB.

Fonte: ACISB

Ainda na primeira página, ao rolar a tela, tem-se as últimas notícias, Parceiros/ Gestão/ Economia/ Associativismo, opinião de associados sobre determinados temas que estão sendo comentados e informados sobre o mercado financeiro, tais como: taxa de câmbio, preço do arroz, uma vez que São Borja é o terceiro maior produtor de arroz no Rio Grande do Sul, e cotação de commodities em geral. Ou seja, de maneira estratégica a equipe de relacionamento da ACISB utiliza a plataforma para realizar as trocas comunicacionais de interesse dos associados, de modo transparente, preciso e claro. Segundo a equipe, a ACISB não possui profissionais capacitados para a criação e manutenção do site. A opção da Instituição foi trabalhar com empresas terceirizadas para a produção de conteúdo e o gerenciamento do site.

Na rede social Facebook<sup>2</sup>, a associação se insere de forma dinâmica, com dados atualizados em sua página, possibilitando o acesso às informações da associação, aos usuários da plataforma.

<sup>2</sup> <https://www.facebook.com/acisbcanalrelacionamento/>



lidade e ao mesmo tempo interação entre eles, legitimando a ação de comunicação organizacional digital e com o intuito de fortalecer a identidade empresarial.



Figura 4: Estratégia compartilhada.  
(Fonte: facebook/acisb)

Notamos através destes exemplos, algumas das ações estratégicas que a ACISB possui para se comunicar com seu público e, essa comunicação gera trocas e compartilhamentos. Com isso, a associação se mostra presente e engajada no compromisso em realizar ações estratégicas de comunicação digital nas redes sociais, neste caso no facebook. O ato de se preocupar em transmitir suas mensagens através desta plataforma demonstra a sensibilidade da organização com o conhecimento contemporâneo e que potencializa as chances de acesso às informações, de aproximação e de fidelização dos públicos.

## Considerações finais

O presente artigo se propôs a compreender as estratégias comunicacionais da associação a partir dos registros das publicações na página do facebook e para aprofundar a análise, realizamos entrevista com a equipe de relacionamento, com o intuito de entender as estratégias e as formas de aproximação com os públicos enquanto ações criativas, de cooperação, integração e fortalecimento da imagem organizacional.

Apresentamos algumas dimensões no desenvolvimento das estratégias digitais, dentre elas, é a terceirização do site da ACISB, a busca contínua por novos usuários na página do facebook, de modo que se amplia a visibilidade das ações, a integração e a cooperação, bem como o uso do aplicativo whatsapp, garantindo a socialização das mensagens aos associados, com agilidade e interatividade. Percebemos que há uma mudança na comunicação organizacional, com uma visão estratégica, preocupando-se tanto com a manutenção das primeiras tecnologias instaladas na instituição (fax), quanto à busca permanente com o processo de inovação e acesso às tecnologias mais recentes que podem auxiliar no desenvolvimento da fonte de produção da informação, articulando os processos organizacionais, com alcance local, regional, nacional e internacional, já que as empresas São Borja/Brasil, San Tomé/Argentina, são portas de acesso ao Mercosul.

Com base na abordagem qualitativa, notamos a promoção de estratégias de comunicação para atrair, motivar e incentivar os públicos (TERRA, 2011). E também, a naturalização (CORRÊA, 2009) das novas tecnologias em sua cultura organizacional. Observamos ainda, que a associação compreende as dimensões e a importância das plataformas digitais (PINHO, 2002), fator importante e esclarecedor, em se tratando de uma organização fronteira que está em um local afastado dos grandes polos administrativos.

Compreendemos que o cenário da comunicação digital nas empresas de São Borja, na região da fronteira, é promissor e com potencial de desenvolvimento. Cabe às empresas locais, refletir, buscar e criar mecanismos para praticar o protagonismo, a flexibilidade, os valores, o dinamismo, a integração, a cooperação e o diálogo. Essas dimensões da cultura organizacional fazem com que os empreendimentos se atentem para as oportunidades que podem surgir através do uso das plataformas, especialmente, a partir da interatividade e da visibilidade dos empreendimentos, isso, por si só justifica o uso das redes sociais digitais como geradora de desenvolvimento e também de uma nova identidade empresarial.

## Referências

- ALMEIDA, Cristóvão Domingos. *Comunicação, consumo e democracia: desafios e possibilidades na imigração haitiana*. São Paulo: Revista Paulus, Fapcom, 2017, v.2, n.1, p. 99-110.
- BALDISSERA, Rudimar. *Comunicação Organizacional: uma reflexão possível a partir do paradigma da complexidade*. In: OLIVEIRA, Ivone de Lourdes; SOARES, Ana Thereza Nogueira. (Org.). *Interfaces e tendências da comunicação no contexto das organizações*. São Caetano do Sul: Difusão Editora, 2008, p. 149-77.
- CASTELLS, Manuel. *A Galáxia da Internet: reflexões sobre a internet, os negócios e a sociedade*. Rio de Janeiro: Zahar, 2003.
- \_\_\_\_\_. *A sociedade em rede*. São Paulo: Paz e Terra S/A, 2009.
- \_\_\_\_\_. *O Poder da Comunicação*. São Paulo/Rio de Janeiro: Paz e Terra, 2015.
- \_\_\_\_\_. *O poder da identidade*. São Paulo: Paz e Terra S/A, 2008.
- CORRÊA, Elizabeth Saad. **A comunicação digital nas organizações**: tendências e transformações. *Revista Organicom*, n. 10-11, 2009. Disponível em: <[http://www.eca.usp.br/departam/crp/cursos/posgrad/gestcorp/organicom/re\\_visita10-11/161.pdf](http://www.eca.usp.br/departam/crp/cursos/posgrad/gestcorp/organicom/re_visita10-11/161.pdf)> Acesso em: 3 abril 2017.
- FRAGOSO, Suely; RECUERO, Raquel; AMARAL, Adriana. *Métodos de pesquisa para internet*. Porto Alegre: Sulina, 2011.
- FLUSSER, Vilém. *Filosofia da caixa preta*. São Paulo: Hucitec, 1985.
- GRUNIG, James; HUNT, Todd. *Dirección de Relaciones Públicas*. Barcelona: Gestión, 2003.
- JUNIOR, Luiz Francisco Ananias; SANTOS, Célia. M. Retz Godoy. *A comunicação reticular e o WhatsApp como tendência no ambiente corporativo*. *Anais do 40º Congresso Brasileiro de Ciências da Comunicação*, Curitiba: Intercom, 4 a 9 set. 2017. Disponível em: <<http://portalintercom.org.br/anais/nacional2017/resumos/R12-0101-1.pdf>> Acesso em: 23 março 2018.
- MORAES, Cândida Maria. *O Paradigma Educacional Emergente*. 13. ed. Campinas: Papirus, 2008.
- PINHO, J. B., *Relações Públicas na internet: técnicas e estratégias para informar e influenciar públicos de interesse*. 2. ed. São Paulo: Summus, 2002.
- RECUERO, Raquel. *Redes Sociais na Internet*. 2. ed. Porto Alegre: Sulina, 2011.
- ROCHA, Rose de Melo. *É a partir das imagens que falamos de consumo: reflexões sobre fluxos visuais e comunicação midiática*. In: CASTRO, Gisela Grangeiro da Silva; BACCEGA, Maria Aparecida (Orgs.). *Comunicação e consumo nas culturas locais e global*. São Paulo: ESPM, 2009, p. 268-293.
- RODRIGUES, Adriano Duarte. *Estratégias da Comunicação*. Lisboa: Editorial Presença, 1990.

SANTAELLA, Lucia; LEMOS, Renata. *Redes digitais: a cognição conectiva do Twitter*. São Paulo: Paulus, 2010.

SCHULER, Maria. A cultura organizacional como manifestação da multidimensionalidade humana. In: KUNSCH, Margarida M. Krohling (Org.). *Comunicação organizacional: linguagem, gestão e perspectivas*. 2. ed. São Paulo: Sarai-va, 2009. p. 243-274.

SILVA, Marcelo; SANTOS, Ana K. N.; SILVA, Priscila, F. G. Identidade Organizacional e Mídias Sociais na pós-modernidade: O caso Deva Curl no Brasil. *Revista Comunicologia*. Brasília, UCB, n.1, p.163-181, jan./jun.2017. Disponível em: <<https://portalrevistas.ucb.br/index.php/RCEUCB/article/view/8100>>. Acesso em: 24 março 2018.

TERRA, Carolina Frazon. *Mídias Sociais... e agora?* São Caetano do Sul: Difusão Editora, 2011.

TERRA, Carolina Frazon. Relações Públicas na era dos Megafones Digitais. In: FARIAS, Luiz Alberto de (Org.). *Relações Públicas Estratégicas: técnicas, conceitos e instrumentos*. São Paulo: Summus, 2011.

**Henrique Esper** - Graduado em Relações Públicas pela Universidade Federal do Pampa (UNIPAMPA)

E-mail: henrique.esper@gmail.com

**Cristóvão Domingos de Almeida** - Pós-doutor em comunicação e professor da Universidade Federal do Pampa (UNIPAMPA). E-mail: cristovaoalmeida@gmail.com

**Recebido: 16 set. 2017**

**Aprovado: 03 abr. 2018**

# **Convergência midiática e cultura participativa: possíveis interações entre novas tecnologias e agentes sociais no campo da comunicação**

## **Media convergence and participatory culture: possible interactions between new technologies and social agents in the field of communication**

### **Antonio Francisco Magnoni**

Universidade Estadual Paulista Júlio de Mesquita Filho  
Programa de Pós-Graduação em Mídia e Tecnologia  
Baurú, SP, Brasil

### **Giovani Vieira Miranda**

Universidade Estadual Paulista Júlio de Mesquita Filho  
Programa de Pós-Graduação em Mídia e Tecnologia  
Baurú, SP, Brasil

## Resumo

Este artigo analisa as editorias colaborativas de dois portais da internet, G1 e A Tarde/UOL, com o objetivo de identificar a mudança na posição dos agentes sociais no campo da comunicação com o advento das novas tecnologias e no contexto do uso dos dispositivos móveis. Para realizar as análises, partiu-se dos conceitos de Castells (1996), Jenkins (2009) e Bourdieu (2004; 2008), e aplicou-se a metodologia de Análise de Conteúdo (AC), utilizando os métodos qualitativos de análise do *corpus*. Os resultados mostram que existe uma mudança na posição dos agentes sociais no campo da comunicação, de consumidores a produtores, de passivos a ativos, de isolados a conectados.

## Abstract

This article analyzes the collaborative publishing of two Internet portals: G1 and A Tarde/ UOL, to identify the change in the position of social agents in the field of communication with the advent of new technologies and in the context of mobile devices usage. In order to perform the analysis, concepts of Castells (1996), Jenkins (2009), and Bourdieu (2004, 2008) were used. The methodology of Content Analysis (CA) was applied using qualitative methods of corpus analysis. The results show that there is a change in the position of social agents in the communication area, from consumers to producers, from passive to active, from isolated to connected.

## Palavras-chave

Análise de conteúdo. Comunicação. Convergência midiática. Cultura participativa. Dispositivos móveis.

## Keywords

Content analysis. Communication. Media convergence. Participatory culture. Mobile devices.

## Introdução

As tecnologias produtivas e informacionais e também as mídias contemporâneas estão em um processo constante de convergência, que segundo Jenkins (2009, p. 29), não deve ser vislumbrado apenas no âmbito tecnológico, mas principalmente, em “uma transformação cultural, à medida que consumidores são incentivados a fazer conexões em meio a conteúdos midiáticos dispersos”.

De acordo com pesquisa nacional da Secretaria de Comunicação Social da Presidência da República-Secom (BRASIL, 2016), 58% dos brasileiros têm acesso à internet. A televisão ainda é a plataforma preferida para consumo de notícias pela maioria da população, mas, ao analisar esse consumo por faixa etária, é possível constatar que a internet é a plataforma preferida pelos brasileiros, com idades entre 12 e 15 anos, e entre as pessoas com formação universitária. Entre os usuários da mídia tradicional que também têm acesso aos dispositivos móveis, 20% disseram preferir ler jornais e revistas pelo celular. Já entre os usuários de tablets, este número sobe para 55%. Entre os usuários da rede web, os mecanismos de busca são citados como a principal fonte de notícias (55%), com as redes sociais logo atrás, em segundo lugar (51%), seguidos por: 37% portais, 35% sites de notícias, 25% sites de mídias impressas, 20% sites de emissoras e 7% de leitores de blogs (BRASIL, 2016).

A informação tornou-se um bem disputado e de valor crescente – as pessoas querem produzir e compartilhar conhecimento, em uma cultura cada vez mais participativa. Com a difusão de diversas categorias de informações pelas mídias sociais, a checagem e a veracidade dos fatos passaram a ser (re) discutidas e questionadas pelo público. A confiabilidade das fontes, boa apuração dos fatos e edição coerente das informações, com linguagem eficiente e atraente para público, era uma premissa inquestionável na era analógica, para que houvesse produção de bom jornalismo, verdadeiro e ético. Hoje, as anti-

gas exigências passaram a figurar como elementos estratégicos e vitais, em meio à difusão acelerada de informações pela internet, principalmente pelas redes sociais, canais interativos de difusão comunicativa, que facilitam a dispersão de informações inverídicas, de boatos, calúnias e de *fake news*.

Plataformas como Facebook, Twitter e Whatsapp favorecem a replicação de informações mal apuradas, de boatos e mentiras intencionais. Grande parte dos factóides é compartilhada por conhecidos, em quem os usuários confiam. A aparente proximidade dos influenciadores aumenta o verniz de legitimidade das falsas histórias. Um exemplo concreto são os algoritmos utilizados pelo Facebook, que direcionam para cada usuário, as informações que reforçam os seus pontos de vista repertoriais ou ideológicos; um processo seletivo nefasto que produz bolhas com indivíduos similares e isola as narrativas gerais, polariza os questionamentos à esquerda ou à direita, sem ocorrer uma diversidade mínima de visões conjunturais ou de argumentos. O resultado é o bloqueio de qualquer debate com alguma pluralidade e, sobre qualquer tema de interesse coletivo, ou que seja reivindicado por um nicho social.

Os veículos jornalísticos, que tradicionalmente são os responsáveis por checar os fatos, e construir narrativas com histórias baseadas na realidade, hoje enfrentam diversos obstáculos, tanto econômico, quanto tecnológico e profissional, para disputar espaço nas redes sociais. Em junho de 2017, o Facebook alterou o seu algoritmo para diminuir o alcance de postagens de sites noticiosos derivados de veículos da “velha mídia” e assim, poder privilegiar os assuntos interpessoais, com temáticas praticamente domésticas, que são compartilhadas rotineiramente entre amigos e familiares.

De acordo com um levantamento do Grupo de Pesquisas em Políticas Públicas para o Acesso à Informação (Gpopai)<sup>1</sup> da Universidade de São Paulo (USP), cerca de 12 milhões de pessoas difundem

1 <http://ibidem.org.br/tag/grupo-de-pesquisa-em-politicas-publicas-para-o-acesso-a-informacao-gpopai-da-usp/>

notícias falsas no Brasil – as chamadas *fake news* –, principalmente aquelas de conteúdo político. O número representa cerca de 6% da população brasileira. Entretanto, se considerarmos que cada usuário das redes sociais tem, em média, 200 conexões, os autores do levantamento acreditam que quase todos os brasileiros têm por meio da internet, o contato diário com algum tipo de notícias falsas. A conclusão dos analistas da USP é resultado de um trabalho de monitoramento de 500 páginas de conteúdo político falso ou distorcido, uma pesquisa empírica realizada durante o mês de junho de 2017.

Nesse contexto, o ambiente digital, anônimo e de trânsito livre, constitui um ambiente propício para as práticas colaborativas dos internautas. Partindo-se dessas ideias e do conceito de campo apontado por Bourdieu (2004; 2008), analisam-se as editorias colaborativas dos portais G1 e A Tarde/UOL, para entendimento de como ocorre o processo de participação dos internautas e compreensão sobre a mudança na posição dos agentes sociais no campo da comunicação.

## Convergência midiática

A digitalização tem afetado intensamente os espaços dos suportes culturais. A qualidade e a rapidez na transmissão de pacotes de dados dentro de um ambiente que permite a comunicação de forma anônima e a livre circulação de informações criaram um cenário propício para a criação e o compartilhamento de conteúdos – por qualquer pessoa; a qualquer tempo e lugar (SILVEIRA, 2010). As novas tecnologias têm afetado as relações humanas de tal maneira que tais relações já não conseguem mais ser completamente entendidas fora de seu diálogo com a tecnologia. Entenda-se aqui o termo “tecnologia” da mesma forma como coloca Castells (1996, p. 5), ao apontar que “a tecnologia não é somente a ciência e as máquinas: é também tecnologia social e organizativa”. Em outras palavras, a revolução tecnológica

está diretamente relacionada com as habilidades de uma sociedade para difundir e trocar informações, relacionando-as com o restante do mundo.

A convergência tecnológica resulta em um fluxo de conteúdos que transita pelas múltiplas plataformas de mídia. “Convergência é uma palavra que consegue definir transformações tecnológicas, mercadológicas, culturais e sociais” (JENKINS, 2009, p.29). Com essa definição do conceito, o autor elimina qualquer possibilidade de se vislumbrar a convergência como um processo puramente tecnológico, que apenas uniria funções diversas dentro de um mesmo aparelho. A convergência deve ser entendida como uma transformação cultural, uma vez que incentiva os consumidores a buscar informações em diversos meios e, a partir deles, criar conexões. É por essa razão que, para o autor, a convergência não ocorre apenas dentro dos aparelhos, mas, principalmente, dentro do cérebro de cada indivíduo e em suas relações sociais. Jenkins (2009, p. 30) acredita que “cada um de nós constrói a própria mitologia pessoal, a partir de pedaços e fragmentos de informações extraídos dos fluxos midiáticos [...] através dos quais compreendemos nossa vida cotidiana”.

Para chegar a essas conclusões, o autor analisa o fluxo de conteúdos que permeiam múltiplos suportes midiáticos, considerando o comportamento migratório do público, que oscila entre diferentes canais de mídia em busca de novas experiências, e entende que a cultura da convergência se apoia em três grandes pilares: (1) a convergência tecnológica dos meios de comunicação, (2) a cultura participativa e (3) a inteligência coletiva. A convergência de mídias, a cultura participativa e a inteligência coletiva favorecem as práticas de criação, compartilhamento e recombinação de conteúdos dentro do ambiente digital.

Para a sociedade idealizada por Castells (1996), a rede de computadores se constitui com uma morfologia própria, na qual a web torna-se símbolo do desenvolvimento comunicacional e espaço de liberdade comunicacional.

A internet é fundamentalmente um espaço social, cada vez mais amplo e diversificado a partir das tecnologias de acesso móvel a ela. Por isso a preservação da liberdade de expressão e comunicação na internet é a principal questão na liberdade de expressão em nosso mundo. (CASTELLS, 1996, p.227)

Esses processos possibilitados pelas novas tecnologias quebram de vez o modelo estratificado e unidirecional de emissor-mensagem-receptor, que já se mostrava demasiadamente simples para responder ao complexo processo de comunicação. Os atuais consumidores podem se tornar produtores, assim como o caminho contrário também é válido, e a informação segue fluxos diversos, sofrendo modificações ao longo do percurso. Isso se justifica, segundo Jenkins (2009), principalmente pela redefinição do papel e da posição do consumidor midiático:

Se os antigos consumidores eram tidos como passivos, os novos consumidores são ativos. Se os antigos consumidores eram previsíveis e ficavam onde mandavam que ficassem, os novos consumidores são migratórios, demonstrando uma declinante lealdade a redes ou os meios de comunicação. Se os antigos consumidores eram indivíduos isolados, os novos consumidores são mais conectados socialmente. Se o trabalho de consumidores de mídia já foi silencioso e invisível, os novos consumidores são agora barulhentos e públicos. (JENKINS, 2009, p. 47).

Essa mudança na posição do consumidor midiático ocorre, principalmente, porque a internet permite a qualquer pessoa criar conteúdos e formatos. Esses conceitos de Jenkins podem ser comparados às indicações de Benkler (2006), para quem a internet representa uma mudança radical nas antigas tendências de comunicação, justamente porque é o primeiro meio de comunicação que consegue expandir o seu alcance ao

descentralizar a estrutura de produção e distribuição de informações, cultura e conhecimento. A maior parte das tecnologias que compõem a internet é baseada em recombinações e está aberta, ou seja, não está sob o controle de patentes ou outras formas que bloqueariam o seu acesso, o que facilita o compartilhamento e a recombinação de conteúdos.

Silveira (2008), em uma interpretação similar à de Benkler, coloca que há uma série de práticas socioculturais que reconfiguraram as redes informacionais, transformando-as em um espaço comum.

Uma série de práticas socioculturais reconfiguraram as redes informacionais como um terreno comum — commons, no sentido anglo-saxônico — e incentivaram a produção de processos, repositórios e interfaces a partir do ciberespaço ou em seu redor, tais como, a música tecno, a Wikipedia, as redes sociais, a blogosfera, o jornalismo open source, o desenvolvimento de softwares livres, [...] as licenças *Creative Commons* e até o YouTube (SILVEIRA, 2008, p. 86).

Em resumo, o tipo de comunicação que prospera no ambiente digital, está relacionado à livre expressão. “É a transmissão de fonte aberta, a livre divulgação, a transmissão descentralizada, a interação fortuita, a comunicação propositada e a criação compartilhada que encontram sua expressão na internet” (CASTELLS, 2003, p. 165). O ambiente digital permite, portanto, uma nova cultura: a cultura participativa.

## **Cultura participativa**

As mídias digitais pós-massivas, por sua vez, se caracterizam principalmente pela capacidade de gerar comunicação e diálogo entre “consciências engajadas em romper o isolamento e em compartilhar uma atividade conjunta” (MARTINO, 2001).

Na avaliação plenamente otimista e “integrada” de alguns autores, essas mídias são marcadas por

um novo formato de consumo, de produção e circulação de informação, que tem como característica principal a liberação do polo da emissão, a conexão planetária de conteúdos e pessoas e, conseqüentemente, a reconfiguração do espaço comunicacional. (LEMOS, 2011, p. 2).

Assim, Lemos salienta que, neste início de século, estamos diante de um sistema comunicacional em que convivem mídias massivas e pós-massivas, o que possibilita o surgimento de uma esfera comunicacional em que a conversação pública se dá no seio mesmo da produção e das trocas informativas, entre atores individuais ou coletivos (LEMOS, 2011). Essa nova “esfera pública mediada” marca uma transição para mídias mais conversacionais, do que informativas, devido ao fato de que as trocas informativas interpessoais ocorrem mais próximas ao diálogo entre amigos e conhecidos, do que à emissão-recepção impessoal, típica das velhas mídias. A ampliação e o refinamento daquilo que é discutido nessa nova “esfera pública mediada”, pode resolver os problemas do engajamento político e levar a uma maior ação política e a uma ampliação da participação pública. “Podemos dizer que a nova prática do jornalismo é mais um exemplo que ilustra a ampliação da conversação aplicada a uma dimensão mais local, permitindo maior engajamento comunitário e político” (LEMOS, 2011, p. 13).

Diversas descrições e teorizações sobre hiperlocalismo advogam que a nova modalidade informativa, além de propiciar maior pluralidade de cobertura, pode gerar mais empregos locais para jornalistas, criar oportunidades de renovação dos processos e das práticas de produção de conteúdos e formatos jornalísticos e, de sustentação econômica das atividades noticiosas; também ajudam a reforçar a noção de identidades e de culturas locais, que reagem instintivamente às estratégias

de homogeneização dos hábitos e identidades culturais. Afinal, a cibercultura progride em um mundo extremamente tangido pelos interesses de uma globalização, que é primeiramente, financeira, e difundida pelas narrativas internacionais, que vendem incessantemente modelos sociais, individualistas, consumistas, e fortemente ideologizados.

Diante da mundialização dos valores da “civilização capitalista ocidental e democrática” as identidades locais, comunitárias e regionais têm se tornado mais importantes, ao mesmo tempo em que os instrumentos informacionais em tempo real e com alcance global parecem provocar uma homogeneização das culturas. Concomitantemente às influências do global, um novo interesse pelo local parece emergir.

De acordo com Hall (2000), a “globalização” afeta a identidade cultural e as sociedades modernas se baseiam na noção de descontinuidade, fragmentação e ruptura. Hall define a globalização como “processos, atuantes numa escala global, que atravessam fronteiras nacionais, integrando e conectando comunidades e organizações em novas combinações de espaço-tempo, tornado o mundo, em realidade e em experiência, mais interconectado” (HALL, 2000, p. 67). A globalização, para ele, deslocou as identidades culturais no final do século XX.

Segue-se que a nação não é apenas uma entidade política, mas algo que produz sentidos - *um sistema de representação cultural*. As pessoas não são apenas cidadãos/dãs legais de uma nação; elas participam da ideia da nação tal como representada em sua cultura nacional. Uma nação é uma comunidade simbólica [...] (HALL, 2000, p. 49).

As conseqüências da globalização para as identidades culturais podem ser resumidas em três pontos: (i) As identidades culturais nacionais estão se desintegrando como resultado do cresci-

mento da homogeneização cultural e do “pós-moderno global”; (ii) as identidades nacionais e outras identidades “locais” ou particularistas estão sendo reforçadas pela resistência à globalização; e (iii) as identidades nacionais estão em declínio, mas novas identidades – híbridas – estão tomando seu lugar.

Com a chegada dos meios digitais, o antigo formato de jornalismo, seja na produção da notícia ou na sua distribuição, mudou de maneira irreversível. A agilidade, a independência na busca de informações e a troca de conteúdo que ocorre pelos usuários das mídias digitais transformam o jeito antigo do jornalista fazer notícias, mas também transformam o jeito do público receber e interpretar as notícias. Devido ao seu caráter colaborativo, o jornalismo hiperlocal só é possível com a participação do público na produção das notícias. Contudo, experiências existentes já demonstraram que de nada adianta centenas de leitores enviarem conteúdo, uma vez que não haverá condições de analisar todo o material (CASTILHO, 2012). Alternativas viáveis existem e podem criar novas formas para dar fôlego às combatidas empresas jornalísticas, e aos pequenos jornalistas produtores de notícias locais.

Somado a políticas públicas eficientes, o jornalismo pós-massivo pode ajudar a promover e contribuir para que se mantenham traços históricos que distinguem os habitantes de determinado lugar e, assim, despertar a responsabilidade dos seus cidadãos (CANCLINI, 2005). No cenário convergente em que vivemos, a colaboração passa a ser fator determinante. Ao mesmo tempo em que jornalistas desenvolvem coberturas cada vez mais rápidas, o consumidor de informação é convidado a participar do relato, enviando todo tipo de conteúdo possível pelas redes sociais (CARVALHO; CARVALHO, 2014).

Por participação na rede deve-se considerar um processo mais complexo do que o simples retorno do consumidor a uma votação virtual, ou a

seleção de um item em uma lista predefinida, por exemplo, uma vez que essas ações caracterizam simplesmente uma interatividade. Jenkins (2009) explica a diferença entre os termos:

A interatividade refere-se ao modo como as novas tecnologias foram planejadas para responder ao feedback do consumidor. [...] A participação, por outro lado, é moldada pelos protocolos culturais e sociais. [...] A participação é mais limitada, menos controlada pelos produtores de mídia e mais controlada pelos consumidores de mídia (JENKINS, 2009, p.189).

A participação, portanto, pressupõe maior poder ao consumidor midiático, porque não é tão controlada pelos produtores de mídia. Benkler (2006) defende que a participação através da rede desenvolve as capacidades práticas dos indivíduos de três grandes maneiras: (1) melhora a sua capacidade de fazer mais por si e pelos outros; (2) impede que os indivíduos organizem suas relações através de modelos hierárquicos tradicionais; e (3) aumenta sua capacidade de participação em organizações formais que operam fora da esfera do mercado.

Avançando em Benkler, à luz de Rossini (2010, p. 211) compreende-se a vantagem que a internet oferece para criar um ambiente que instiga a participação, uma vez que a web confere aos cidadãos “novas oportunidades de participação locais, regionais e globais. Sua possibilidade de construção e apropriação de conhecimento, adequando-o para sua realidade local, são potencialmente infinitas”. Dessa forma, os cidadãos já não estão mais sujeitos somente à posição de leitores ou de observadores de conteúdo; a rede permite sua participação e redefine o lugar de criação de conteúdos. Criam-se, assim, novas oportunidades para que esses cidadãos possam construir e transmitir conhecimentos, criar e participar de canais de interação e de ativismo. Silveira (2010) compartilha desse pensamento ao indicar que os processos de comunicação criados pelas novas tecnologias reforçam o potencial comu-

nicador das localidades e permite a conexão delas com quaisquer pontos e nós da rede, comprimindo e anulando as distâncias e sintetiza a importância da participação efetiva na internet, ao colocar que a participação nas redes digitais é um exercício criativo de cidadania digital.

Os novos canais de interação surgidos com as novas tecnologias permitem, enfim, o desenvolvimento de uma cultura participativa. O termo é utilizado por Jenkins (2006, p. 378) que o define como uma "cultura em que os fãs e outros consumidores são convidados a participar ativamente da criação e da circulação de novos conteúdos". A cultura atual é mantida por meio da produção mútua e da troca recíproca de conhecimento que fluem pelos novos canais de interação criados pelas tecnologias atuais.

Conforme defende Jenkins (2009), o novo ambiente midiático promete um fluxo mais livre de ideias e conteúdos.

Consumidores estão aprendendo a utilizar as diferentes tecnologias para ter um controle mais complexo sobre o fluxo da mídia e para interagir com outros consumidores. As promessas desse novo ambiente midiático provocam expectativas de um fluxo mais livre de ideias e conteúdos. Inspirados por esses ideais, os consumidores estão lutando pelo direito de participar mais plenamente de sua cultura. (JENKINS, 2009, p.44)

Em suma, a produção de informação, conhecimento e cultura através das relações sociais, ao invés de relações de propriedade e de mercado, é o que cria as oportunidades para uma cultura mais crítica e uma república autônoma, engajada e melhor informada (BENKLER, 2006). Na cultura da convergência todos são participantes em maior ou menor grau e a cultura participativa modifica as antigas noções de passividade dos consumidores midiáticos, colocando-os em papel similar ao dos produtores de informação.

Com esses novos fluxos e essa nova cultura

surgidos a partir das novas tecnologias da informação e da comunicação, é preciso compreender como os agentes sociais alteram sua posição dentro do campo da comunicação. Para entender esse fenômeno, recorre-se aos conceitos propostos por Bourdieu (2004; 2008).

## **O campo da comunicação e a posição dos agentes locais**

O campo, seja ele literário, artístico, político, jurídico ou científico, é o universo no qual se inserem os agentes sociais e as instituições que produzem, reproduzem ou difundem a arte, a literatura ou a ciência (BOURDIEU, 2004). Esse espaço é relativamente autônomo, regido por leis próprias e no qual os agentes lutam entre si, mostrando suas forças para manter ou transformar o campo. De acordo com Bourdieu (2004) são as relações entre os agentes e a posição que ocupam dentro do campo que determinam o que eles podem ou não fazer:

É a estrutura das relações objetivas entre os agentes que determina o que eles podem e não podem fazer. É a posição que eles ocupam nessa estrutura que determina ou orienta suas tomadas de posição. Isso significa que só compreendemos, verdadeiramente, o que diz ou faz um agente engajado num campo se estamos em condições de nos referirmos à posição que ele ocupa nesse campo, se sabemos "de onde ele fala". (BOURDIEU, 2004, p. 24)

Por essa razão, a posição dos agentes é um fator importante para se entender o campo. Os agentes podem ocupar uma das seguintes posições: (1) a posição de dominantes e (2) a posição de dominados, ou pretendentes. Os dominantes se utilizam de estratégia de conservação e defesa das convenções e, por esta razão, estão localizados no centro do campo, num espaço de tradição. Já os dominados, ou pretendentes, estão localizados fora do campo, num espaço de vanguarda e se utilizam

de estratégias de subversão e transgressão das convenções para se inserir no campo.

De acordo com Bourdieu (2008) esses lugares são classificatórios e permitem que os agentes se situem no campo:

As estruturas objetivas do campo da produção estão na origem das categorias de percepção e apreciação que estruturam a percepção e a apreciação de seus produtos. É, assim, que pares antitéticos de pessoas ou de instituições, jornais, teatros, galerias, editoras, revistas, costureiros, podem funcionar como esquemas classificatórios que só existem e são significativos em suas relações mútuas, além de permitirem a identificação dos interlocutores e de se situarem com precisão. [...] Os produtores ou produtos que não estão no devido lugar, mas, como se diz, 'descolados', estão mais ou menos condenados ao fracasso: todas as homologias que garantem um público ajustado, críticos compreensivos, etc, àquele que encontrou seu lugar na estrutura funcionam, pelo contrário, contra aquele que se desviou de seu lugar natural. (BOURDIEU, 2008, p.56-57).

Quando há uma mudança na posição desses agentes, ou seja, ocorrem lutas para entrar no campo, são gerados movimentos de quebra no círculo da crença com o objetivo de fundar uma nova crença. Quando a crença ou o agente social alcança o centro do campo, ele se torna consagrado. A eficácia dos atos de consagração está no próprio campo, reproduzida com a ajuda dos agentes, as instituições e de suas lutas no campo.

No campo do jornalismo, as instituições consagradas, ou legitimadas, para a construção das notícias são as redações, e os agentes que ocupam o centro do campo são os jornalistas. Porém, com as alterações de fluxos de conteúdos permitidas pela internet, os consumidores se tornaram produtores de conteúdo e estão se inserindo no campo. Para entender como está ocorrendo a inserção desses agentes sociais no campo do jornalismo e verificar

se a participação dos internautas é ainda uma vanguarda ou se já se tornou uma tradição no campo, serão analisadas a seguir duas editorias colaborativas de portais da internet.

## Metodologia e análises

A metodologia escolhida para a aplicação da fundamentação teórica no objeto de estudo foi a Análise de Conteúdo (AC). Wilson Fonseca Junior (2009) acredita que a formação do campo comunicacional não pode ser compreendida sem se fazer referência à Análise de Conteúdo.

Desde sua presença nos primeiros trabalhos da communication research às recentes pesquisas sobre novas tecnologias, passando pelos estudos culturais e de recepção, esse método tem demonstrado grande capacidade de adaptação aos desafios emergentes da comunicação e de outros campos do conhecimento (FONSECA JUNIOR, 2009, p. 42).

Este item traz a análise do *corpus*, o processo de codificação e categorização, além da descrição quantitativa e qualitativa das análises.

### Análise do *corpus*

Os objetos de estudo que se estabelecem como *corpus* deste trabalho são as editorias colaborativas do portal G1, chamada de VC no G1, e do portal A Tarde/ UOL, chamada de Cidadão Repórter.

No caso do VC no G1, verifica-s a) Aparência: não há um portal que compile as notícias, mas apenas uma aba para colaboração; b) Elementos: as reportagens são inseridas no portal em meio às demais notícias tendo a tag 'VC no G1' como elemento diferenciador; c) Destaque da editoria na home do portal: não ocupa nenhum lugar de destaque na página principal, é colocada junto às outras editorias do portal; d) Canais de interação: os principais canais de interação são as redes sociais, como o perfil no Twitter <<https://twitter>.

com/vcnog1>; e) O que o internauta pode enviar: o portal recebe textos, vídeos e fotos dos internautas; f) Como ele pode participar: o internauta deve preencher um cadastro no site da globo.com. Depois disso, ele faz o login no canal de interação do G1, com o usuário e senha criados no cadastro. Para enviar a notícia, o internauta deve dar um título, selecionar o arquivo que deseja enviar, descrever comentários e indicar o local do acontecimento do fato; g) Informações complementares: na página de envio de arquivos, há informações aos internautas sobre o processo de participação. Há indicações, por exemplo, sobre o tempo em que o conteúdo enviado ficará disponível: vídeos ficam publicados durante seis meses; fotos e textos ficam disponíveis na rede por tempo indeterminado. As informações também deixam claro quais os tipos de conteúdos que são proibidos no portal e ainda oferece algumas dicas para melhorar a qualidade e facilitar o envio dos conteúdos; h) comentários nas matérias: há espaço para que outros internautas possam comentar as matérias ou compartilhá-las através das redes sociais e e-mail.

Já no caso do Cidadão Repórter, coloca-se a) Aparência: possui layout similar ao das outras editorias da porta; b) Elementos: as imagens e os vídeos são colocados no mesmo tamanho dentro da página. Apenas uma matéria ganha um destaque maior no início da tela. Todas as notícias possuem ícones que permitem compartilhá-las através das redes sociais; c) Destaque da editoria na home do portal: não ocupa nenhum lugar de destaque na página principal, é colocada junto às outras editorias do portal; d) Canais de interação: o principal canal de interação é o “Seja um Cidadão Repórter”, localizado no canto direito do meio da página. O texto incita a colaboração do internauta ao utilizar a expressão “Envie fotos e vídeos. Envie sugestão de pautas. E seja um cidadão repórter”; e) O que o internauta pode enviar: o portal recebe fotos, vídeos, áudios e

sugestão de pauta dos internautas; f) Como ele pode participar: não é necessário criar um cadastro no portal; basta que o internauta preencha algumas informações, como nome, e-mail e telefone, selecione o arquivo que deseja enviar e escreva as informações referentes ao arquivo; g) Informações complementares: abaixo do campo de envio, há um termo de cessão de direitos autorais, que esclarece ao internauta que, ao enviar fotos, vídeos e áudios, ele cederá ao portal o direito de edição, comercialização, reprodução, impressão e publicação para todos os canais do grupo A Tarde; h) Comentários nas matérias: há espaço para que outros internautas possam comentar as matérias ou compartilhá-las através das redes sociais e e-mail.

Para a codificação e categorização dos elementos textuais das editorias colaborativas, selecionou-se o método de análise categorial. A análise categorial trata do desmembramento do discurso em categorias, que serão posteriormente analisadas qualitativamente. Bardin (2003) define este método como espécie de gavetas ou rubricas significativas que permitem a classificação dos elementos de significação constitutivos da mensagem. Para identificar as principais características dos textos colaborativos publicados nesses portais, criou-se uma tabela com os principais pontos observados em quatro matérias, duas do VC no G1 e duas do Cidadão Repórter.

No VC no G1, foi selecionada a reportagem intitulada “Internauta registra neblina em região serrana do AP durante verão amazônico”, publicada em 29 de julho de 2017, e “Ruas em loteamento da Zona Norte de Macapá são alvos de reclamação de internauta”, publicada em 17 de agosto de 2017. Já no Cidadão Repórter, foi escolhida a reportagem “Moradores denunciam péssimo estado de trecho da BA-245”, publicada em 02 de maio de 2017, e “Dona de cadela denuncia maus tratos durante tosa”, publicada em 15 de janeiro de 2016.

Os textos estão assim caracterizados:

Tabela 1 - Categorias observadas nos textos do VC no G1

<b>Pessoa do discurso (jornalista ou internauta)</b>	Jornalista. Os internautas entraram como fonte na matéria.
<b>Créditos nas matérias</b>	Sim, foi colocado o nome dos internautas no texto e nas imagens.
<b>Fontes utilizadas</b>	Os próprios internautas que enviaram as matérias e fontes que representavam o "outro lado" da notícia.
<b>Fotos e vídeos</b>	As fotos e vídeos enviados pelos internautas entraram para ilustrar o texto.

Fonte: Tabela elaborada pelos autores.

Tabela 2 - Categorias observadas nos textos do Cidadão Repórter

<b>Pessoa do discurso (jornalista ou internauta)</b>	Jornalista. Os internautas entraram como fonte na matéria.
<b>Créditos nas matérias</b>	Uma das matérias dava o crédito do texto para "Da redação" e nos créditos da outra matéria havia o nome de um cidadão, mas não ficou claro se era jornalista ou internauta.
<b>Fontes utilizadas</b>	Os próprios internautas que enviaram as matérias e fontes que representavam o "outro lado" da notícia.
<b>Fotos e vídeos</b>	Foram utilizadas imagens marcadas como 'reprodução'

Fonte: Tabela elaborada pelos autores.

Observou-se que, inicialmente, as novas tecnologias colaboram com as práticas de participação, alterando os fluxos de produção das notícias e as duas páginas, através de textos e imagens e

ferramentas de interação, instigam a colaboração do internauta, convidando-o a ser um jornalista cidadão. Porém, o texto final, publicado nos portais, acaba sendo reescrito pelos agentes do centro do campo, os jornalistas. Aos internautas, que são chamados de "jornalista" e "repórter" nas editorias, resta ser apenas uma fonte da matéria. Mesmo que as matérias venham com créditos, fica bem claro neles a indicação de "internauta". Em outras palavras, os agentes dominantes do campo, os jornalistas, lutam para manter a tradição no campo, impedindo que os agentes de vanguarda, os internautas, consigam chegar ao centro dele.

Outro ponto a ser observado é que apesar de as notícias estarem sendo geradas e disseminadas a partir de diversos locais, as redações ainda podem ser consideradas instituições legitimadas de disseminação de informações dentro do campo da comunicação. Primeiro, porque é nela que trabalham os agentes de tradição do campo, os jornalistas. Segundo, as redações recebem e concentram todos os textos produzidos pelos internautas, alterando-os conforme suas necessidades. Isso fica claro, por exemplo, nos termos de cessão dos direitos autorais e nas políticas dos portais, que esclarecem que os textos enviados pelos internautas serão modificados, comercializados e publicados e, portanto, os direitos sobre eles passam a ser desses portais.

Além disso, os jornalistas são considerados agentes com mais credibilidade dentro do campo da comunicação e é por essa razão que as notícias devem mostrar que passaram por suas mãos. Uma das formas que as editorias colaborativas utilizam-se para provar isso é o uso, em todas as matérias, de fontes, entrevistas e formatação de texto comumente utilizado e consagrado no jornalismo.

Em síntese, os resultados das análises mostram que embora os agentes pretendentes - os internautas participativos - tentem se inserir no campo do jornalismo, os agentes dominantes - os jornalistas - ainda se mantêm como tradição no centro do campo. Esse agentes se utilizam de estratégias de conservação e

defesa das convenções para se manterem no centro do campo. Entre essas estratégias, a principal que pode ser citada é a reescrita da matéria, deixando os internautas apenas como fonte.

## Considerações finais

Analisar as editorias colaborativas dos portais G1 e A Tarde/ UOL foi o objetivo deste estudo. Para isso, foram feitas breves reflexões a respeito de conceitos que envolviam o objeto de estudo, como convergência midiática, uso de dispositivos móveis, cultura participativa e o campo da comunicação.

Observa-se que a digitalização contribuiu para o fornecimento de notícias de uma forma geral. A disponibilidade e a quantidade de notícias aumentaram, mas a maioria das fontes permanece a mesma. Os principais sites de notícias no Brasil são controlados pelos mesmos agentes que controlavam a mídia no período pré-internet, e, apesar de haver uma produção de conteúdo relevante por parte de blogueiros independentes e ONGs que possuem seus grupos e seguidores, os grandes grupos da mídia tradicional continuam a exercer forte controle sobre a forma como a opinião popular é moldada (MIZUKAMI, 2014).

É importante salientar que o termo “audiência ativa” utilizado por Downing (2002) reflete a realidade de cultura participativa frente às novas tecnologias digitais, em especial o uso dos dispositivos móveis. O papel do público não se restringe mais ao papel passivo de recepção de conteúdos, ao contrário, ele deseja participar ativamente e de modo colaborativo da rede, atuar em sua construção e também na ligação com o mundo técnico, profissional e conceitual do jornalismo e da mídia como um todo. A internet alterou – e continua alterando – gradativamente os hábitos coletivos de comunicação mediada, que antes ocorriam por processos mais físicos e a curta distância.

As possibilidades de se estabelecer relações remotas interpessoais e interculturais cresce-

ram com o desenvolvimento de tais tecnologias. Mesmo que elas ainda sejam distribuídas de forma desigual, são capazes de induzir mudanças e possibilidades comunicativas para movimentos e organizações. As redes trouxeram recursos de visibilidade e de empoderamento ao cidadão comum, organizações e coletivos.

De acordo com documento da United Nations (2014), há oportunidades claras para a melhoria participação dos cidadãos, a partir, por exemplo, dos dispositivos móveis. Existem também alguns problemas a serem considerados, como a exclusão digital e as poucas iniciativas de incentivo à participação. Estas oportunidades e desafios exigem estratégias eficazes para criar um ambiente propício para a participação, incluindo quadros legais e institucionais adequados, capacidade de desenvolvimento para a alfabetização de mídia digital para os cidadãos e a integração de recursos online e offline para a participação do público.

Com esses aspectos considerados, partiu-se para a análise de conteúdo do *corpus*, buscando-se fragmentar seu texto em categorias que facilitassem a apreensão dos resultados. O resultado das análises mostrou que os internautas, com o uso das novas tecnologias, deixam de ser simples consumidores de notícias para se transformarem em produtores de informação. Ocorre, dessa forma, uma mudança na posição desses agentes sociais no campo da comunicação, de consumidores a produtores, de passivos a ativos, de isolados a conectados. Entretanto, essa participação dos internautas ainda pode ser considerada de vanguarda, visto que suas notícias costumam ser modificadas pelos jornalistas e ainda ficam em editorias separadas e específicas, não se misturando as outras produções dos repórteres, que se inserem no campo ainda como tradição. Sendo assim, os internautas ainda são dominados em um campo onde os jornalistas são os dominantes.

Vale ressaltar que os trabalhos que abordam o campo da comunicação são importantes para com-

preender com maior clareza os diferentes fluxos e funções dos agentes sociais participantes do processo comunicacional. As transformações sociais, culturais e econômicas advindas das novas tecnologias são um terreno fértil para o desenvolvimento de estudos e devem ser explorados.

## Referências

- BARDIN, L. *Análise de Conteúdo*. Lisboa: Edições 70, 2003.
- BENKLER, Y. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven e Londres: Yale University Press, 2006. Disponível em: <[http://cyber.law.harvard.edu/wealth\\_of\\_networks/Download\\_PDFs\\_of\\_the\\_book](http://cyber.law.harvard.edu/wealth_of_networks/Download_PDFs_of_the_book)>. Acesso em: 27 de ago. 2017
- BOURDIEU, P. *Os usos sociais da ciência*. Por uma sociologia clínica do campo científico. São Paulo: Unesp, 2004.
- BOURDIEU, P. *A produção da crença*. Contribuição para uma economia dos bens simbólicos. Porto Alegre, Zouk, 2008.
- BRASIL. SECRETARIA DE COMUNICAÇÃO DA PRESIDÊNCIA DA REPÚBLICA *Pesquisa Brasileira de Mídia 2016: Hábitos de consumo de mídia pela população brasileira*. Disponível em: <<http://www.secom.gov.br/atuacao/pesquisa/lista-de-pesquisas-quantitativas-e-qualitativas-de-contratos-atuais/pesquisa-brasileira-de-midia-pbm-2016.pdf/view>>. Acesso em: 27 ago. 2017.
- CANCLINI, N.G. *Consumidores e cidadãos*. Rio de Janeiro: Editora UFRJ, 2005
- CARVALHO, J. Maurício; CARVALHO, A. M. G. Do hiperlocal aos insumos criativos: as mutações do jornalismo na contemporaneidade. In: CARVALHO, J. M; BRONOSKY. (Org.). *Jornalismo e convergência*. São Paulo: Cultura Acadêmica, 2014, p. 69-87.
- CASTELLS, M. Fluxos, Redes e Identidades: Uma Teoria Crítica da Sociedade Informacional. In: Castells, M. (Org.). *Novas perspectivas críticas em educação*. Porto Alegre: Artes Médicas, 1996.
- CASTILHO, C. Jornalismo Hiperlocal ganha adeptos na grande imprensa. Observatório da Imprensa, 2012. Disponível em <<http://www.observatorio-daimprensa.com.br/posts/view/jornalismo-hiperlocal-ganha-adeptos-na-grande-imprensa>> Acesso em 25. ago. 2017
- FONSECA JUNIOR, W. C. *Análise de Conteúdo*. In: DUARTE, Jorge; BARROS, Antonio (org). Métodos e técnicas de pesquisa em comunicação. São Paulo: Atlas, 2009.
- HALL, S. *A identidade cultural na pós-modernidade*. 4. ed. Rio de Janeiro: DP&A, 2000
- JENKINS, H. *Cultura da convergência*. 2. ed. São Paulo: Aleph, 2009.
- LEMOS, A. Nova esfera conversacional. In MARQUES, A. et al. Esfera pública, redes e jornalismo. Rio de Janeiro: Editora. E-papers, 2009.
- LEMOS, C.E.B; PEREIRA, R.M. Jornalismo hiperlocal no contexto multimídia: um relato da experiência do jornal-laboratório Contramão Online. XVI Congresso de Ciências da Comunicação na Região Sudeste. São Paulo, 2011.
- MARTINO, L. M. *Teoria das Mídias Digitais: linguagens, ambientes, redes*. Petrópolis: Vozes, 2014.
- JENKINS, H. *Cultura da convergência*. São Paulo: Aleph, 2009.
- MIZUKAMI, P. et al. *Mapeamento da mídia digital no Brasil*. Open Society Foundations, 2014 Disponível em: <<http://www.opensocietyfoundations.org/sites/default/files/mapping-digital-media-brazil-por-20141201.pdf>>. Acesso em: 10 set. 2017.

ROSSINI, C. Aprendizagem digital, recursos educacionais abertos e cidadania. In: *Cidadania e redes digitais*. São Paulo: Comitê Gestor da Internet no Brasil: Maracá - Educação e Tecnologias, 2010.

SILVEIRA, S. A. Cibercultura, commons e feudalismo informacional. *Revista Famecos*, Porto Alegre, n. 37, 2008.

UNITED NATIONS. *E-government Survey: E-government for the future we want*. Disponível em: <[http://unpan3.un.org/egovkb/Portals/egovkb/Documents/un/2014-Survey/E-Gov\\_Complete\\_Survey-2014.pdf](http://unpan3.un.org/egovkb/Portals/egovkb/Documents/un/2014-Survey/E-Gov_Complete_Survey-2014.pdf)> Acesso em: 14 ago. 2017.

**Antonio Francisco Magnoni** – Professor do curso de Jornalismo e do Programa de Pós Graduação em Mídia e Tecnologia (Mestrado Profissional e Doutorado) da Faculdade de Arquitetura, Artes e Comunicação da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP). **E-mail:** afmagnoni@gmail.com

**Giovani Vieira Miranda** – Mestrando em Comunicação Midiática e graduado em Comunicação Social – Jornalismo pela Faculdade de Arquitetura, Artes e Comunicação (FAAC) da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP). **E-mail:** giovani.vieira.miranda@gmail.com

**Recebido: 23 set. 2017**

**Aprovado: 14 maio 2018**