
 <p>ESCOLA DE COMUNICAÇÃO, ARTES E DESIGN FAMECOS</p>	<h1>REVISTA FAMECOS</h1> <p>mídia, cultura e tecnologia</p> <p>Revista FAMECOS, Porto Alegre, v. 27, p. 1-10, jan.-dez. 2020 e-ISSN: 1980-3729   ISSN-L: 1415-0549</p>
<p> <a href="https://dx.doi.org/10.15448/1980-3729.2020.1.34074">https://dx.doi.org/10.15448/1980-3729.2020.1.34074</a></p>	

VARIA

## O papel dos algoritmos de inteligência artificial nas redes sociais

*The role of artificial intelligence algorithms in the social web*

*El papel de los algoritmos de inteligencia artificial em las redes sociales*

**Dora Kaufman<sup>1</sup>**

[orcid.org/0000-0001-7060-4887](https://orcid.org/0000-0001-7060-4887)

[dkaufman@usp.br](mailto:dkaufman@usp.br)

**Lucia Santaella<sup>1</sup>**

[orcid.org/0000-0002-0681-6073](https://orcid.org/0000-0002-0681-6073)

[lbraga@pucsp.br](mailto:lbraga@pucsp.br)

**Recebido em:** 03 mai. 2019.

**Aprovado em:** 15 out. 2019.

**Publicado em:** 28 mai. 2020.

**Resumo:** Desde 2016 nos Estados Unidos e na Europa o tema das *fake news* (notícias falsas) tornou-se preocupação central do jornalismo em virtude dos prejuízos que esse fenômeno causa especialmente à preservação saudável dos sistemas democráticos. Em 2018, o processo eleitoral trouxe esse tema para o ápice dos debates no Brasil. É nos aplicativos das redes sociais que as *fake news* proliferam e não é mais novidade saber que tanto as redes sociais quanto os motores de busca estão hoje sob o controle dos algoritmos de inteligência artificial. Em função do desconhecimento acerca do modo como esses algoritmos funcionam e do temor que filmes e séries sobre o assunto têm despertado nas pessoas, desenvolveu-se um pensamento negativista e mesmo certa ojeriza contra os algoritmos, em especial, os de inteligência artificial. Em vista disso, tendo por base uma pesquisa cuidadosa sobre as tendências atuais da inteligência artificial, este artigo tem por objetivo desmistificar crenças e fantasias infundadas ao evidenciar objetivamente como funciona a inteligência artificial e o papel que ela desempenha por trás do funcionamento das redes sociais.

**Palavras-chave:** Inteligência Artificial. Redes Sociais. Fake News.

**Abstract:** Since 2016 in the United States and Europe the topic of fake news has become a central concern of journalism because of the damage that this phenomenon causes especially to the healthy preservation of democratic systems. In 2018, the electoral process brought this theme to the climax of debates in Brazil. It is in social networking applications that fake news proliferate and it is no longer unknown that both social networks and browsers are now under the control of artificial intelligence algorithms. Because of the lack of knowledge about how these algorithms work and the fear that films and series on the subject have awakened in people, a certain negative feeling has developed against algorithms, especially the ones of artificial intelligence. In view of this, based on careful research on the current trends in artificial intelligence, this article aims to demystify unfounded beliefs and fantasies by objectively demonstrating how artificial intelligence works and the role it plays behind the functioning of social networks.

**Keywords:** Artificial Intelligence. Social Networks. Fake News.

**Resumen:** Desde 2016 en Estados Unidos y en Europa, el tema de las fake news se ha convertido en una preocupación central del periodismo debido a los perjuicios que este fenómeno causa especialmente a la preservación sana de los sistemas democráticos. En 2018, el proceso electoral trajo ese tema para el ápice de los debates en Brasil. Es en las aplicaciones de las redes sociales que las fake news proliferan y ya es sabido que tanto las redes sociales como los motores de búsqueda están hoy bajo el control de los algoritmos de inteligencia artificial. En función del desconocimiento acerca de cómo funcionan estos algoritmos y del temor que las películas y series sobre el tema han despertado en las personas, se ha desarrollado un pensamiento negativista e incluso una cierta ojeriza contra los algoritmos, en especial los de inteligencia artificial. En vista de ello, teniendo como base una investigación cuidadosa sobre las tendencias actuales de la inteligencia artificial, este artículo tiene por objetivo desmitificar creencias y fantasias infundadas al evidenciar objetivamente cómo funciona la inteligencia artificial y el papel que desempeña detrás del funcionamiento de las redes sociales.

**Palabras-clave:** Inteligencia Artificial. Redes Sociales. Fake News.



Artigo está licenciado sob forma de uma licença  
[Creative Commons Atribuição 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/).

<sup>1</sup> Pontifícia Universidade Católica de São Paulo, São Paulo, SP, Brasil.

## Introdução

Em 1909, o escritor britânico Edward Morgan Forster publicou a novela *The Machine Stops* (FORSTER, 2011), retratando um cenário futurista no qual os humanos são, simultaneamente, servidos e controlados por uma máquina. Em uma surpreendente antecipação tecnológica, os habitantes desse planeta Terra (ano indefinido) comunicam-se por meio de uma placa redonda, segura pelas mãos, de onde emerge uma luz azulada projetando à distância imagem e som.

Construída pelo homem, a Máquina gradativamente torna-se onipresente, nada acontece fora de seu domínio; a dependência humana é absoluta gerando, conseqüentemente, um poder maquinico igualmente absoluto. Na novela de Foster, toda ação humana, da mais simples como levantar um objeto do chão até os relacionamentos interpessoais, é mediada pela Máquina. Seus habitantes vivem "felizes" em aceitação das regras vigentes. Vashti, palestrante plenamente integrada ao sistema, tem cinco filhos, um dos quais, Kuno, é um rebelde que percebe os riscos, a eminente falência do sistema e prevê a "parada" da Máquina, mas a quem a Máquina benevolmente perdoa as infrações.

As questões que afloram na relação entre mãe e filho, Vashti e Kuno, remetem a questões em debate frente ao recente avanço da inteligência artificial (IA), que rebatem nos dois típicos extremos entre a distopia e a eutopia. A primeira, encontra-se na advertência final de Kuno a Vashti:

Você não percebe, não percebem todos vocês palestrantes, que somos nós que estamos morrendo e que aqui a única coisa que realmente vive é a Máquina? Criamos a Máquina, para realizar nossas vontades, mas agora já não podemos fazer com que atenda. Ela nos roubou a sensação de espaço e a sensação de toque, confundiu todas as relações humanas e reduziu o amor a um ato carnal, paralisou nossos corpos e nossas vontades e agora nos obriga a adorá-la. A Máquina se desenvolve – mas não em nossa direção. A Máquina avança – mas não rumo ao nosso objetivo. Existimos apenas como os corpúsculos sanguíneos que correm por suas artérias, e se ela pudesse funcionar sem nós, nos deixaria morrer (FORSTER, 2011, p. 33).

Essa advertência é similar à definição que Bostrom (2006) nos fornece de "risco existencial" como aquele que ameaça aniquilar a vida inteligente no planeta Terra, ou limitar permanente e drasticamente seu potencial. O primeiro risco existencial criado pelo ser humano foi o advento da bomba atômica.

Na época, havia alguma preocupação de que a explosão pudesse iniciar uma reação em cadeia descontrolada ao "acender" a atmosfera. Embora agora saibamos que tal resultado é fisicamente impossível, um risco existencial estava presente (BOSTROM, 2006, p. 2).

O segundo risco existencial para a humanidade, segundo Bostrom, associa-se à potencial explosão da inteligência com a criação da Superinteligência de máquina, detentora da vantagem estratégica de moldar o futuro da vida inteligente com base em suas próprias motivações. A Superinteligência é definida como "qualquer intelecto que exceda em muito o desempenho cognitivo dos seres humanos em praticamente todos os domínios de interesse" (BOSTROM, 2014, p. 26). A Superinteligência visualizaria os humanos como uma ameaça potencial ao seu sistema, devendo portanto ser eliminados. A primeira Superinteligência poderia facilmente ter objetivos finais não antropomórficos, causando uma "catástrofe existencial".

São várias as distopias alusivas à Superinteligência, como nos explica Cozman: "a distopia envolve a destruição da humanidade ou sua redução a uma situação de servidão completa (talvez uma servidão na qual os seres humanos remanescentes estejam subjugados e inconscientes da sua real situação)" (COZMAN, 2018, p. 35).

O outro extremo pode ser percebido quando se comparam os dizeres do "Livro" que, na novela de Foster, discorrem sobre a onipotência da Máquina e o atual movimento "Dataísmo", a religião dos dados. Seus adeptos, concentrados no Vale do Silício, creem que o valor dos fenômenos é função de sua contribuição ao processamento de dados, estabelecendo uma equivalência entre os algoritmos bioquímicos e os eletrônicos. "Os dataístas explicam aos que ainda cultuam mortais

de carne e osso que eles estão excessivamente atrelados a uma tecnologia ultrapassada. O Homo sapiens é um algoritmo obsoleto" (HARARI, 2016, p. 384). Para este autor, no Vale do Silício, os profetas dataístas exploram conscientemente a linguagem messiânica tradicional, criticando como uma postura religiosa o livro de profecias de Ray Kurzweil, *The Singularity is Near*, em que "ecoa o brado de João Batista 'o reino dos céus está próximo' (MATEUS 3,2)".

Não obstante os temores, de um lado, e as euforias, de outro, o estágio atual de desenvolvimento da inteligência artificial (IA) é ainda restrito. Estamos na fase da "Weak AI", a "Strong AI" ainda é ficção. Segundo os especialistas com os quais nos alinhamos (Yoshua Bengio, Geoffrey Hinton, Yann LeCun, Andrew Ng, Kai-Fu Lee, Davi Geiger, Martin Ford), não há nenhum indicador real de que a Superinteligência será alcançada. A atualidade da novela de Forster é que, mesmo para os céticos, o debate está colocado na mídia e nos centros de pesquisa e tecnologia.

Um dos caminhos para evitar, ou ao menos minimizar, as previsões distópicas é a tomada de consciência da sociedade sobre os desafios envolvidos nos benefícios oriundos da IA, e como enfrentá-los. Um dos temas em debate é a personalização das experiências no acesso *online*, acesso esse mediado pelos algoritmos de inteligência artificial. Não há consenso com relação a essa intervenção automatizada, de modo que o propósito deste artigo é repensar seus efeitos à luz do funcionamento da tecnologia.

Para isso, inicialmente será dada uma breve descrição do estágio atual da inteligência artificial, contemplando os fundamentos básicos da lógica do *machine learning* (aprendizado de máquina), particularmente a técnica de *deep learning* (aprendizado profundo). Sem adentrar propriamente nos meandros da tecnologia, que envolve conceitos matemáticos e estatísticos avançados, a ideia é apenas oferecer uma visão geral facilitando a análise posterior sobre as implicações da personalização dos acessos à rede.

## Modelo preditivo de inteligência artificial: *deep learning*

Na segunda década do século XXI, a convergência de diversas tecnologias – internet das coisas (IoT), *blockchain*, plataformas digitais, impressão 3D, robótica avançada, novos materiais, manipulação genética – permeadas pela IA, tem promovido resultados superiores a quaisquer previsões precedentes (ainda que aquêm da ficção científica). As máquinas e os sistemas inteligentes estão executando tarefas que até recentemente eram prerrogativas dos humanos, em alguns casos com resultados mais rápidos e mais assertivos. Elas, porém, ainda estão restritas a prever cenários com base em grandes conjuntos de dados e a executar tarefas específicas, sob a supervisão direta dos especialistas em ciência da computação. Esses sistemas carecem da essência da inteligência humana: capacidade de compreender o significado; igualmente não possuem senso intuitivo, capacidade de formar conceitos abstratos e de fazer analogias e generalizações, não têm a capacidade de compreender o funcionamento do mundo a partir da observação (conceitos como tridimensionalidade, movimentação e permanência dos objetos, gravidade, inércia e rigidez dentre outros) (LECUN citado por FORD, 2018).

Apesar das limitações, tais avanços da IA, contudo, trouxeram benefícios inéditos para a humanidade e desafios éticos e regulatórios complexos. Apenas para citar a área médica, o sistema Watson da IBM<sup>3</sup> apresenta uma taxa de sucesso de 90% comparativamente aos 50% dos médicos humanos nos diagnósticos de câncer de pulmão. O sistema é capaz de processar grandes volumes de dados estabelecendo correlações entre sintomas e/ou imagens em uma dimensão impossível de ser alcançada por um ser humano.

O banco de dados Medline, por exemplo, pertencente à Biblioteca Nacional de Medicina dos EUA, tem aptidão para indexar mais de 5.600 periódicos e milhões de registros médicos, históricos de pacientes e estudos de caso que

<sup>3</sup> Watson é a plataforma de serviços cognitivos/inteligência artificial da IBM para negócios.

oferecem importantes *insights* com o uso de algoritmos de IA (estima-se que o volume total dos dados dobre a cada cinco anos). A contribuição da IA na área de saúde transcende os diagnósticos, abrangendo a prevenção de epidemias e potenciais anomalias individuais. Os nanorrobôs, em estágio avançado de desenvolvimento, poderão navegar na corrente sanguínea identificando doenças e eliminando patógenos, como células cancerosas (KAUFMAN, 2019).

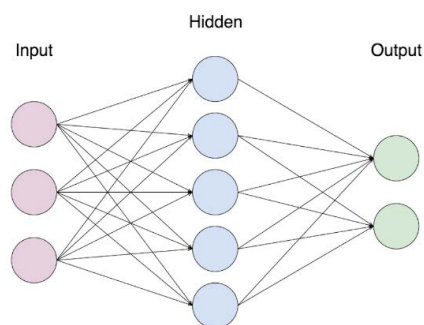
O sucesso atual é "mera" implementação do modelo estatístico *deep learning* (aprendizado profundo, DL), inserido no subcampo da IA denominado de *machine learning* (aprendizado de máquina, ML). Em 1959, Arthur Lee Samuel, pioneiro norte-americano no campo de jogos de computador e IA, enquanto funcionário da IBM, cunhou o termo *machine learning*, inaugurando um subcampo da IA cuja finalidade é prover os computadores da capacidade de aprender sem serem programados. Evoluindo a partir do estudo do reconhecimento de padrões e da teoria de aprendizagem computacional na IA, o *ML* explora o estudo e a construção de algoritmos que, seguindo instruções, fazem previsões ou tomam decisões baseadas em dados – modelos elaborados a partir de entradas de amostras. O aprendizado de máquina é empregado em uma variedade de tarefas de computação, nas quais programar os algoritmos é difícil ou inviável. Trata-se de um processo de solução de um problema específico por meio da construção algorítmica de um modelo estatístico baseado em um conjunto de dados (BURKOV, 2019).

Inspirado no funcionamento do cérebro, por isso também conhecido como redes neurais, o *deep learning* foi concebido na década de 1980, concretizado a partir de 2010-2012 com o crescimento exponencial dos dados, e a maior capacidade computacional – GPUs (*Graphic Processing Unit*) e computação em nuvem –, que reduz o tempo de treinamento dos algoritmos.

*Deep learning* diz respeito à previsão, com base em dados. "Atualmente, quando acessamos um dispositivo computacional, em qualquer de seus formatos, provavelmente estamos acessando concomitantemente um processo de *deep learning*" (KAUFMAN, 2019, p. 25).

Por várias décadas, a abordagem dominante no campo da IA foi baseada em programas lógicos de computação, marginalizando a visão baseada no aprendizado de máquina. O atual crescimento exponencial dos dados, contudo, inviabiliza o uso da programação (regras definidas *a priori*). A Amazon não pode codificar os gostos do conjunto de seus clientes em um programa de computador, assim como o Facebook desconhece como escrever um programa para identificar as melhores atualizações no Feed de Notícias. A Netflix pode ter cem mil títulos de DVD em estoque, mas se os clientes não souberem como encontrar suas preferências de nada adianta. A grande quantidade de dados não é o único fator restritivo, o conhecimento tácito é igualmente um limitador (KAUFMAN, 2019). A contar de 2012 o *deep learning* tornou-se onipresente, recebendo expressivos investimentos das gigantes de tecnologia.

As redes neurais são funções matemáticas biologicamente inspiradas, formadas por neurônios artificiais que interagem entre si. A figura 1 traduz uma arquitetura padrão, em que os círculos representam os neurônios artificiais e as setas os pesos.



**Figura 1** – Redes Neurais

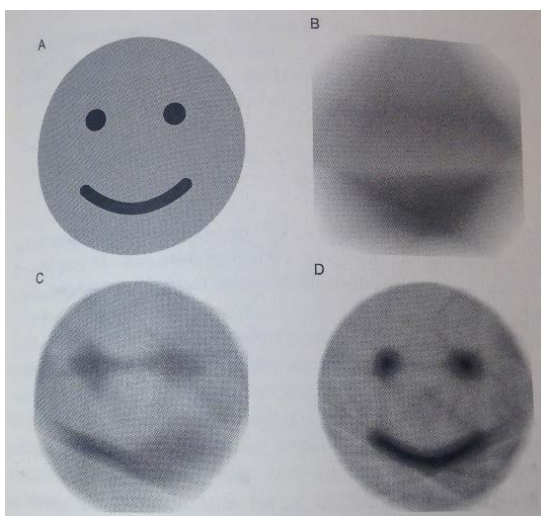
**Fonte:** Wiederer<sup>4</sup> ([2016]).

<sup>4</sup> WIEDERER, Christine. Redes neurais em JavaScript. Desenvolvimento, IMasters, 20 jun. 2016. Disponível em: <https://imasters.com.br/desenvolvimento/redes-neurais-em-javascript>. Acesso em: 12. jan. 2019.

As redes neurais são adequadas a problemas inteiramente distintos. O treinamento consiste em mostrar exemplos e ajustar gradualmente os parâmetros da rede até obter os resultados requeridos, denominado aprendizagem supervisionada: são fornecidos os resultados desejados (*output*) e, por "tentativa e erro", através de atualização iterativa dos pesos, chega-se ao resultado – meta (*back propagation*). A pergunta-chave migrou de "quais as características de um cachorro?", elementos que permitiam programar um sistema, para "essa imagem é similar a uma imagem que já vi antes?", processo que estima a probabilidade da imagem ser efetivamente de um cachorro.

O objetivo de um algoritmo de aprendizado supervisionado é usar o conjunto de dados para produzir um modelo que tenha um vetor de atributos  $x$  como entrada e gerar informações que permitem deduzir o label desse vetor de atributos (BURKOV, 2019, p. 2).

Na Figura 2, a imagem 2A é o output (imagem a ser reconhecida). As imagens subsequentes, representam o resultado das camadas intermediárias (hidden layers), cuja precisão depende do número de neurônios: a figura C tem 200 neurônios, por exemplo, e a figura D tem 2.000 neurônios.

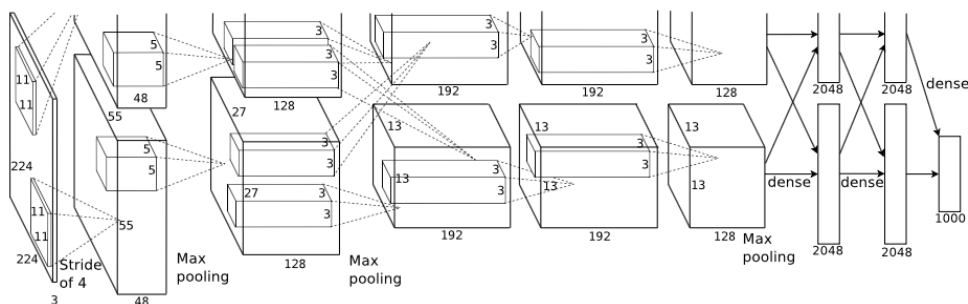


**Figura 2** – Reconhecimento de imagem  
**Fonte:** Gerrish (2018, p. 113).

A rede geralmente tem entre 10 e 30 camadas empilhadas de neurônios artificiais. Em um reconhecimento de imagem, por exemplo, a primeira camada procura bordas ou cantos; as camadas intermediárias interpretam as características básicas para procurar formas ou componentes gerais; e as últimas camadas envolvem interpretações completas. Na identificação de fotos nas redes sociais, a máquina percebe padrões e "aprende" a identificar rostos, tal como alguém que olha o álbum de fotos de uma família desconhecida e, depois de uma série de fotos, reconhece o fotografado (existe uma hierarquia de memorização). O reconhecimento de voz, que junto com a visão computacional está entre as aplicações mais bem-sucedidas, já permite a comunicação entre humanos e máquinas, mesmo que ainda precária (Siri, Alexa, Google Now).

Parte dos avanços inovadores decorrem da arquitetura *Deep Convolutional Neural Network* ou CNN (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), ilustrada na Figura 3, onde aparece explicitamente o delineamento de responsabilidades entre as duas GPUs: uma GPU executa as partes da camada na parte superior da figura, enquanto a outra executa as partes da camada na parte inferior.





**Figura 3** – CNN – Convolutional Neural Network

**Fonte:** Krizhevsky, Sutskever e Hinton (2012, p. 4).

As CNNs são usadas na visão computacional, reconhecimento de fala e processamento de linguagem natural. Sua eficácia é responsável pela rápida adoção em aplicações diversas como carros autônomos, robótica, drones, segurança, diagnósticos médicos e tratamentos para pessoas com deficiência visual.

O nível de precisão dos resultados é função do número de camadas, do número de neurônios artificiais e da quantidade e qualidade dos dados (inputs), arbitrados pelo operador humano. O que define uma máquina inteligente são dois componentes principais: o valor de cada conexão ("pesos") e a arquitetura. *Deep learning* é sobre previsão com base em correlações, reduzindo significativamente o custo e o tempo de processamento dos modelos preditivos. O objetivo não é identificar causalidades entre fenômenos, mas descobrir padrões e correlações que gerem *insights*.

Antes do Big Data, nossa análise geralmente se limitava a testar um pequeno número de hipóteses que definíamos bem antes de coletar os dados. Quando deixamos os dados falarem, podemos fazer conexões que nunca imaginamos que existissem (Mayer-Schönberger; Cukier, 2013, p. 14 ).

Os algoritmos de aprendizado não funcionam da mesma forma, e suas diferenças impactam os resultados e, por vezes, o próprio modelo utilizado.

### Acesso personalizado e os algoritmos de IA

Muito se tem discutido e mesmo demonizado a presença dos algoritmos de IA nas redes sociais. Embora não seja possível negar que eles, de fato, lá estão operando, discussões e críticas devem ser, antes de tudo, bem-informadas. A intenção deste artigo não é outra se não a de informar sobre o estado da arte da IA para que discussões sobre o funcionamento dos algoritmos nas redes não se mantenha encoberto por equívocos e fantasias.

Antes de tudo, é preciso levar em consideração que, muito longe da "Máquina" onipotente da novela de Foster, as operações da aprendizagem de máquina na IA apresentam vulnerabilidades e limitações, conforme foi evidenciado acima. Assim, para evitar a proliferação de ilusões mal informadas, escolhemos como foco de discussão a personalização dos dados no acesso do usuário à informação.

A recente explosão de dados na internet trouxe a questão da curadoria, substituindo a ideia de liberdade dos primórdios da rede pela ideia de relevância. O acesso à informação passou a ser personalizado, o que atende aos usuários das plataformas digitais que não desejam ver publicações, anúncios publicitários, recomendações de produtos, inadequados às suas preferências. Atualmente, a maior parte da curadoria é efetivada pelos algoritmos de IA, particularmente pelo processo de *deep learning*. Um dos efeitos colaterais que tem sido mais

debatido é o da formação de "bolhas" ou "câmara de eco" (*clusters*). O objeto que nosso foco tomará para aproximação será o *Feed* de Notícias do Facebook. Contudo, antes de chegarmos ao efeito colateral das "bolhas", vejamos como o Face autodescreve seu funcionamento.

De acordo com Moseri (2018), o *Feed* de Notícias de cada um de nós é feito de histórias postadas por nossos amigos, aqueles que escolhemos seguir e grupos que compartilhamos. O ranqueamento é o processo que o Face utiliza para organizar essas histórias de modo que possamos ver os conteúdos mais relevantes no topo, todas as vezes que abrimos a página do Face. Ainda segundo Moseri (2018), o ranqueamento apresenta os seguintes elementos: o inventário das histórias disponíveis, as predições que o Face faz, incluindo aí a probabilidade de que iremos comentar a história, a pontuação de relevância de cada história. Para saber como isso funciona, seguem as informações fornecidas pelo Face sobre a operacionalização de seus algoritmos:

- a) o algoritmo de classificação do *Feed* de Notícias, registrado como patente, é o *Filtering Content in a Social Networking Service*, que é um algoritmo de modelagem preditiva com base no processo *deep learning*. A classificação é ligeiramente cronológica, mas não rigorosamente;
- b) os algoritmos estabelecem um *ranking* pela "pontuação de afinidades", significando semelhanças indicativas de um relacionamento entre o usuário e o objeto de classificação (*post*, outro usuário, dentre outros). Em termos estatísticos, trata-se da probabilidade do objeto ser relevante para o usuário, denominado de *Relevance Score*;
- c) os algoritmos selecionam um subconjunto de objetos de conteúdo em função das pontuações de afinidades auferidas entre o usuário e os objetos de conteúdo. Em seguida, esse subconjunto é filtrado automaticamente com base nos atributos do perfil do usuário (localização, idade, interesses, preferências,

emprego, estado civil). Quanto maior o conjunto de informações do usuário captadas/arquivadas pela plataforma, maior a assertividade da filtragem de conteúdo;

- d) a variável "*timing*" é soberana, os objetos mais recentes são privilegiados no processo de filtragem (em geral, o *Feed* de Notícias não publica conteúdo antigo); se um determinado conteúdo não está gerando interações (*likes*, comentários, compartilhamentos), ele vai perdendo a relevância até desaparecer;
- e) os algoritmos priorizam interações ativas, definindo "ações de qualidade" as que requerem mais esforço do usuário (potencial gerador de mais interações, mais dados). Fatores mais valorizados na classificação: comentário, tipo de reação (ícone "amor" vale mais do que o ícone "curtir"), resposta de comentário (diálogo entre os usuários, conversa), compartilhar links pelo Messenger (para um grupo é mais valorizado do que para um único amigo), e engajamento em ações.

O Facebook disponibiliza diversos recursos que permitem ao usuário interferir no sistema de filtragem de conteúdo. Com certa frequência, o próprio Facebook publica *posts* incentivando o usuário a configurar as suas preferências (postura ativa na relação com a plataforma). Naturalmente, a simples movimentação do usuário fornece indicadores aos algoritmos.

### "Câmara de Eco" como efeito colateral

O tema das "bolhas" tem sido intensamente explorado na bibliografia acadêmica e não - acadêmica, particularmente no campo da comunicação. Não é por acaso, pois a questão das bolhas está na base da proliferação de *fake news* que tanto tem abalado os princípios do jornalismo. A principal crítica às bolhas consiste na homogeneização que estas promovem das relações sociais ao manter os indivíduos em círculos sociais fechados, formados por iguais. Pariser (2011) alerta para o processo invisível de

filtragem de conteúdo que, ao personalizar nossos acessos e pesquisas, nos coloca em contato com o que queremos ver e não com o que devemos ver, que inclui contemplar o que seja desconfortável, desafiador, além de pontos de vista contrários e, por isso mesmo, enriquecedores.

Segundo Santaella (2018, p. 17), a personalização dos filtros "apresenta tendências que afetam significativamente o acesso à informação, na medida em que conduzem o usuário a pontos de vista estreitos que impedem a exposição a ideias contrárias aos seus preconceitos". A limitação das pessoas a uma exposição seletiva, alimentada pelos algoritmos, intensifica as tendências homofílicas, ou seja, aquelas de só se buscar concordâncias e fugir das discordâncias, tendências, de resto, que já fazem parte do funcionamento do psiquismo humano.

De fato, na homofilia encontra-se a fonte primeira das *fake news*. Essas são de variadas espécies, desde a mera brincadeira, o trolling, os memes que visam enganar, as meias verdades até a mentira a. Infelizmente estas últimas têm muito mais apelo do que as informações honestas. Em função de seu apelo, as *fake news* têm mais probabilidade de serem compartilhadas do que notícias corretas. Mas as grandes causadoras dos compartilhamentos são as crenças nas quais os indivíduos se enclausuram: só gostam daquilo que está em concordância com o que confortavelmente pensam e sentem. É justamente o perfil de cada um que é rastreado pelos algoritmos. Portanto, o grande vilão não é simplesmente o algoritmo, mas a dificuldade que o ser humano tem para transformar seu modo de pensar e ver o mundo.

O pior problema encontra-se no fato de que as *fake news* afetam, sobretudo, a vida política. No Brasil de 2018, assistiu-se a uma verdadeira guerra simbólica nas redes que repercutiu nas tomadas de decisões e no modo de agir das pessoas. Isso não é de se estranhar, pois são as nossas crenças que guiam nossas ações.

Levando a discussão para o contexto democrático, Sunstein (2017, p. 3) alega que, em uma democracia eficiente, os cidadãos não vivem

em câmaras de eco, ou casulos de informação, mas são expostos a uma diversidade de tópicos e ideias. Entretanto, "a maioria dos americanos atualmente acessa grande parte das notícias nas redes sociais [...] Quando as pessoas usam o Facebook para ver exatamente o que querem ver, sua compreensão do mundo pode ser bastante afetada"; um sistema de livre expressão garante a exposição a perspectivas concorrentes. Ao contrário, "arquiteturas de controle" podem levar a um tipo de distopia, ao passo que uma "arquitetura ao acaso" neutralizaria a homofilia, promovendo a liberdade individual. Para que isso seja alcançado, são dois os requisitos defendidos por esse autor: (a) expor os cidadãos a informações e fatos que não teriam escolhido antecipadamente, valorizando os encontros imprevistos e não planejados e (b) proporcionar uma ampla gama de experiências comuns; para ele, a carência de experiências compartilhadas dificulta lidar com os problemas sociais.

Ambas as implicações são relevantes para a relação da sociedade com as novas tecnologias. Em uma tentativa de contribuir para o debate, o artigo traz uma reflexão a partir dos algoritmos de IA, no contexto da rede social Facebook, especificamente o *Feed* de Notícias. Suas ponderações são as seguintes:

- a) não há evidência empírica de que a filtragem de conteúdo decorre de ideologias, ou reflete interesses claramente determinados por setores, grupos, instituições ou indivíduos;
- b) a personalização dos acessos *online* é resultado de modelos estatísticos (algoritmos de IA) treinados com base nos dados gerados da movimentação dos usuários no ambiente digital, e de arquiteturas "neutras" do ponto de vista ideológico e/ou político e/ou quaisquer outros interesses afora a eficácia dos resultados estimados (agradar os clientes, aumentar a movimentação nas plataformas, gerar mais dados, ampliar o potencial de ganhos financeiros dos controladores das plataformas);



- c) o tempo de permanência do usuário na plataforma é estratégico, bem como a qualidade dessa permanência (traduzida em quantidade e diversidade de interações). Quanto maior o tempo e a intensidade da interação, mais dados são gerados, favorecendo os modelos de negócio baseados em dados;
- d) as plataformas não detêm o poder absoluto sobre os resultados da filtragem de conteúdo; parte das variáveis que compõem os resultados é originada nos processos automatizados, e parte é originada pela interferência direta do usuário, sobretudo visível no Feed de Notícias do Facebook;
- e) a lógica da filtragem, ou seja, o funcionamento dos algoritmos de IA do Facebook é relativamente transparente. A rede social disponibiliza amplamente informações sobre seus mecanismos, nos sites e nas redes em geral (que, por sua vez, são replicados e comentados exaustivamente por estudiosos, jornalistas, ou simples usuários);
- f) personalização não é o mesmo que individualização, de fato ocorre uma "desindividualização". Os algoritmos de IA, para estabelecer correlações, precisam ter acesso a uma grande e diversificada quantidade de dados. A filtragem de conteúdo extrapola a movimentação de um usuário individual, os algoritmos buscam similaridades com outros usuários: os algoritmos decodificam o comportamento e as preferências de "comunidades" afins de usuário.

As posições defendidas pelo autor podem parecer otimistas; Entretanto, elas são necessárias para se evitar distopias precoces. Quanto aos algoritmos, com o crescimento exponencial da informação, não há como evitar uma curadoria. De resto, alguma forma de curadoria sempre existiu, seja pelos veículos de mídia, pelo professor na escola, pelas famílias. Além disso, os sistemas automatizados pelos algoritmos de IA não são, em si, movidos por ideologias, interesses políticos e de grupos; a finalidade do motor é

gerar mais movimentação, conseqüentemente mais dados, potencializando a expansão dos modelos de negócio e a concentração e o poder das gigantes de tecnologia baseadas em dados. Os usuários dessas plataformas deveriam ter recursos disponíveis para interferir na filtragem de conteúdo em vez da entrega passiva ao designo dos algoritmos. Mas para que isso seja possível, todos os especialistas são unânimes: uma formação educacional profunda é indispensável. Só isso capacita o ser humano para o exercício do pensamento crítico, pois é este que funciona como antídoto contra crenças infundadas.

## Referências

- BOSTROM, Nick. Dinosaurs, dados, humans? In: WORLD ECONOMIC FORUM. **Global Agenda**. Oxford: WEF, 2006. p. 230-1. Disponível em: <https://nickbostrom.com/papers/globalagenda.pdf>. Acesso em: 12 jan. 2019.
- BOSTROM, Nick. **Superintelligence**: paths, dangers, strategies. Oxford: Oxford Press, 2014.
- BURKOV, Andriy. **The hundred-page machine learning book**. Publisher: Andriy Burkov, 2019.
- COZMAN, Fabio. Inteligência Artificial: uma utopia, uma distopia. **TECCOGS**: Revista Digital de Tecnologias Cognitivas, São Paulo, n. 17, p. 32-42, jan./jun. 2018.
- FORD, Martin. **Architects of intelligence**: the truth about AI from the people building it. Birmingham: Pack Publishing, 2018.
- FORSTER, Edward Morgan. **The machine stops**. Londres: Penguin Books, 2011.
- GERRISH, Sean. **How smart machine think**. Cambridge: MIT Press, 2018. <https://doi.org/10.7551/mitpress/11440.001.0001>
- HARARI, Yuval Noah. **Homo Deus**: uma breve história do amanhã. São Paulo: Companhia das Letras, 2016.
- KAUFMAN, Dora. **A inteligência artificial irá suplantará a inteligência humana?** São Paulo: Estação das Letras e Cores, 2019.
- KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. **ImageNet classification with deep convolutional neural networks**. Toronto:[s. n.], 2012. p. 1-9. Disponível em: <https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>. Acesso em: 20 jan. 2019. <https://doi.org/10.1145/3065386>

MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth.

**Big data:** a revolution that will transform how we live, work, and think. Nova York: Houghton Mifflin Harcourt, 2013. <https://doi.org/10.3359/oz1314047>

MOSSERI, Adam. **News Feed Ranking in Three**

**Minutes Flat:** How Does News Feed Work?. [S. l.: s. n.], 2018. Disponível em <https://newsroom.fb.com/news/2018/05/inside-feed-news-feed-ranking/>. Acesso em: 12 jan. 2019. <https://doi.org/10.18254/s0000057-4-1>

PARISER, Eli. **The filter bubble:** what the internet is hiding from you. Londres: Penguin Books, 2011.

SANTAELLA, Lucia. **A pós-verdade é verdadeira ou falsa?** São Paulo: Estação das Letras e Cores, 2018.

SUNSTEIN, Cass R. #republic: divided democracy in the age of social media. Princeton: Princeton University Press, 2017. <https://doi.org/10.1086/696988>

WIEDERER, Christine. Redes neurais em JavaScript. Desenvolvimento. **IMasters**, [S. l.], p. 1-2, 20 jun. 2016. Disponível em: <https://imasters.com.br/desenvolvimento/redes-neurais-em-javascript>. Acesso em: 12 jan. 2019. <https://doi.org/10.21528/cbrn1994-056>