



O BERCO DO BIG DATA

A monumental abundância de dados, sua variedade e a velocidade com que trafegam no universo digital estão revolucionando a civilização

ANDRÉ PETRY, DE NOVA YORK

Há uma década, a Target, a gigantesca loja de departamentos com 1800 pontos de venda nos Estados Unidos, atribuiu um número a cada um dos seus milhões de clientes e passou a rastrear e armazenar todas as pegadas digitais deixadas por eles: produtos preferidos, hábitos de consumo, média de gastos, uso de cupons, cartão de fidelidade. Somou a isso dados demográficos de cada um deles, adquiridos em empresas do ramo: sexo, idade, profissão, local de moradia, estimativa de renda. Contratou estatísticos para analisar essas informações e montou um retrato preciso do padrão de consumo de cada cliente. Um dia aconteceu um incidente.

Um senhor entrou esbravejando numa loja da Target em Minnesota. Trazia nas mãos cupons de produtos para bebês. "Minha filha recebeu isto aqui pelo correio", reclamou o senhor para o gerente. "Ela é uma adolescente. Vocês estão querendo estimulá-la a engravidar?" O gerente conferiu a remessa dos cupons e, constrangido, pediu desculpas. Dias depois, com receio de perder o cliente, telefonou para ele a fim de desculpar-se outra vez. O pai da adolescente estava desconcertado do outro lado da linha: "Tive uma conversa com a minha filha. Fiquei sabendo de algumas coisas que estavam acontecendo dentro da minha

casa". Respirou fundo e completou: "Ela vai dar à luz em agosto...".

O caso, narrado por Charles Duhigg no livro *O Poder do Hábito*, virou um clássico do Big Data — o nome em inglês usado para definir a tectônica quantidade de dados e informações que produzimos no mundo digital. É uma tal profusão de dados que, em quinze minutos, a humanidade gera o triplo de informações disponíveis no acervo da Biblioteca do Congresso americano, a maior do mundo. A era do Big Data só se materializou com a confluência de alguns fatores. Caiu o custo de armazenar dados. Há duas décadas, estocar 1 gigabyte saía por 1.000 dólares. Hoje, custa 6 centavos. Os processadores tornaram-se velozes, os programas ficaram mais inteligentes e a quantidade de dados cresce exponencialmente. Tudo isso junto é o berço do Big Data.

A imensidão de informações é composta, numa pequena parte, de dados limpos, corretos, checados, como uma pesquisa do IBGE. São os "dados estruturados". Mas a grande novidade do Big Data, o elemento verdadeiramente novo e transformador, são os "dados não estruturados". Os dados sujos, incompletos, caóticos. É o clique do mouse para comprar um livro na Amazon, ou não comprar. É a nova foto no Facebook, um novo tuíte, mesmo criado de erros ortográficos. É um novo vídeo no YouTube. É a pesquisa no Google, o telefonema gravado "para sua segurança", a música que se escuta online, o livro que se lê num leitor eletrônico. É o e-mail que se envia para um amigo, e mesmo o e-mail que não chega a ser enviado. É o lixo on-line. Mas um lixo que vale ouro. Eis o que diz o matemático americano Peter Norvig, ex-diretor de tecnologia da informação da Nasa e atual diretor de pesquisa do Google: "Com o Big Data, gasta-se muito mais tempo coletando dados do que chegando



MICHAEL COGLIANTINI/2012/FROM THE HUMAN FACE OF BIG DATA



O HOMEM BIG DATA

O escritor americano A.J. Jacobs transformou sua rotina em um banco de informações. Seu corpo é cheio de aparelhos que medem a quantidade de passos que dá, a qualidade de seu sono, as calorias do que come e as que são gastas (passar aspirador em casa, 246 calorias por hora). Autor best-seller, ele leva a vida como um grande experimento: em *Um Ano Bíblico*, Jacobs seguiu todas as regras da Bíblia, incluindo mandamentos obscuros como não comer frutos de árvores com menos de cinco anos. A obsessão por registrar dados começou em 2009 e resultou em seu último livro, *Drop Dead Healthy* (em tradução livre, *Morto de Saúde*), no qual se propôs a ser o homem mais saudável possível.

Jacobs descobriu que calcular tudo que faz é o melhor caminho para ter o corpo em dia.

A contrapartida é o stress mental.



PETER MENZEL 2012/FROM THE HUMAN FACE OF BIG DATA

SOB CONTROLE

Espantado por não saber por que era tão cara a conta de eletricidade de sua casa, o americano Shwetak Patel, professor de engenharia na Universidade de Washington, criou um sistema que registra o consumo de energia de cada equipamento doméstico. Patel verificou que qualquer aparelho tem assinatura digital única e criou sensores para detectá-la. Espalhados pela casa e conectados a um tablet equipado com um software também desenvolvido por ele, os sensores permitem calcular o consumo de energia da TV de sua sala, da geladeira, da secadora, de cada lâmpada, e assim por diante. O professor descobriu, por exemplo, que os maiores gastadores eram a bomba do filtro da piscina e os aparelhos de DVD, ligados ininterruptamente. Com esse conhecimento, ficou mais fácil cortar despesas. Seguindo a lógica do mundo digital, Patel patenteou a invenção, fundou uma startup e já a vendeu a uma empresa maior.

A MULHER BIG DATA

Big Data poderia ser o apelido da nova-iorquina Hilary Mason. Cientista-chefe da empresa Bitly, ela transformou uma ferramenta que antes se limitava a reduzir o tamanho de URLs de sites (um endereço longo como veja.abril.com.br/multimidia/programa/help-desk-vira-abr.ai/12hTAwZ) em fenômeno do Big Data. Com ela, o Bitly passou a analisar como as pessoas navegam pelos sites. Hilary também estuda algoritmos usados em programas para criar máquinas que, baseadas em informações do Big Data que elas mesmas coletam, possam aprender sozinhas novas funções. A mulher Big Data ainda fundou uma organização inusitada: rastreia hackers e os ajuda a conseguir empregos legais. É uma forma de tirá-los da bandidagem virtual.

ERIN PATRICE O'BRIEN 2012/FROM THE HUMAN FACE OF BIG DATA

a resultados. Mais de 90% da informação armazenada sempre aparenta ser dispensável. O que percebemos é que o real diferencial está no que é considerado lixo. Grandes descobertas ocorrem quando olhamos com os olhos corretos o que foi descartado e assim vemos o que esses dados podem nos revelar do mundo. Esse é o segredo do Big Data".

A análise do oceano de informações pode revelar um padrão, uma correlação, um significado antes oculto e, quase sempre, ajuda a prever o futuro. Stephan Fihn, médico fascinado por estatística, coordena o sistema do Big Data da rede médico-hospitalar que atende 25 milhões de veteranos de guerra dos EUA. Sua base de dados é colossal: 80 bilhões de arquivos, 9 milhões de internações, 6 bilhões de exames de laboratório, 3 bilhões de bilhetes médicos — manuscritos, caóticos. Cruzando tudo, Fihn faz projeções sobre a saúde dos pacientes. "Temos previsões semanais", diz. "Qual a probabilidade de tal paciente ser internado em trinta dias? Qual o risco de falecer? Isso está revolucionando o modo como pensamos em saúde."

Em toda a história, a humanidade tomou decisões com base numa parte do todo — na amostragem, na pesquisa, na estimativa. Agora, pela primeira vez, podemos começar a tomar decisões com base no todo, com 100% de informação. Escreve Michael Malone, professor da Universidade Santa Clara, no Vale do Silício: "A humanidade viveu 10000 gerações fazendo escolhas impregnadas de incerteza. Eram grandes saltos no desconhecido". Na era do Big Data, deixaremos de viver como as 10000 gerações que nos antecederam. Para alguns pesquisadores, o Big Data é tão revolucionário na vida humana quanto a descoberta do fogo ou o início da agricultura.

O monumental impacto será sentido em todos os campos. É provável que tenhamos de superar nossa tendência mental a medir tudo com precisão e rigor. Na era do Big Data, isso é desnecessário e inútil. A vasta quantidade de dados diluiu a eventual má qualidade. Teremos de aprender a lidar com o caos do mundo real. Em vez de sairmos à rua para recolher dados perfeitos

O MUNDO EM TRÊS "Vs"

Volume, Variedade e Velocidade formam a unidade trina do Big Data, expressão ainda sem tradução adequada para o português. Ela identifica um conjunto de informações digitais tão grande que ultrapassa a capacidade de armazenamento e processamento de qualquer ferramenta tecnológica isoladamente

VOLUME

No princípio era o bit

Bit

Abreviação de dígito binário, a linguagem dos computadores: sequência de números 0 e 1 que guarda informações codificadas em HDs de computador

Byte

Equivale a **8 bits** e é o suficiente para armazenar **um caractere** de texto no PC



Kilobyte (1 000 bytes)

A informação contida em **uma página** de livro



Megabyte (1 000 000 de bytes)

Armazena **um quinto** de toda a obra de **William Shakespeare**

Gigabyte (1 000 000 000 de bytes)

Uma hora de vídeo em baixa resolução



Terabyte (1 000 000 000 000 de bytes)

385 terabytes guardam todo o catálogo da **Biblioteca do Congresso americano**, a maior do mundo



Petabyte (1 000 000 000 000 000 de bytes)

1,5 petabyte armazena todas as músicas já criadas pela humanidade

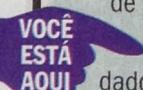


Exabyte (1 000 000 000 000 000 000 de bytes)

3 exabytes é tudo o que a humanidade conseguia guardar em 1986 — hoje produzimos quase o dobro disso em dois dias

Zettabyte (1 000 000 000 000 000 000 000 de bytes)

48 bilhões de iPads, que montam um muro de **17 metros** de altura com **6 400 quilômetros** de extensão (o raio da Terra)

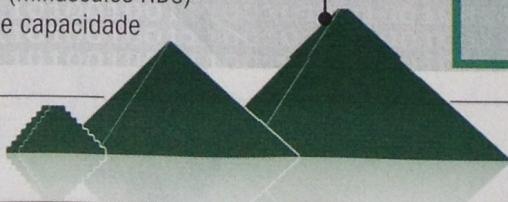


VOCÊ
ESTÁ
AQUI

1,8 zettabyte armazena todos os dados acumulados pela civilização em um ano

Yottabyte (1 000 000 000 000 000 000 000 000 de bytes)

Encheria a **Grande Pirâmide de Gizé** com memory cards (minúsculos HDs) de **64 gigabytes** de capacidade



VARIEDADE

De onde vem a informação

A cada dia

2,5 exabytes

de informação são produzidos pela humanidade



375 megabytes

de dados são acumulados por cada família



24 petabytes

são processados pelo site do Google

Google

43 petabytes

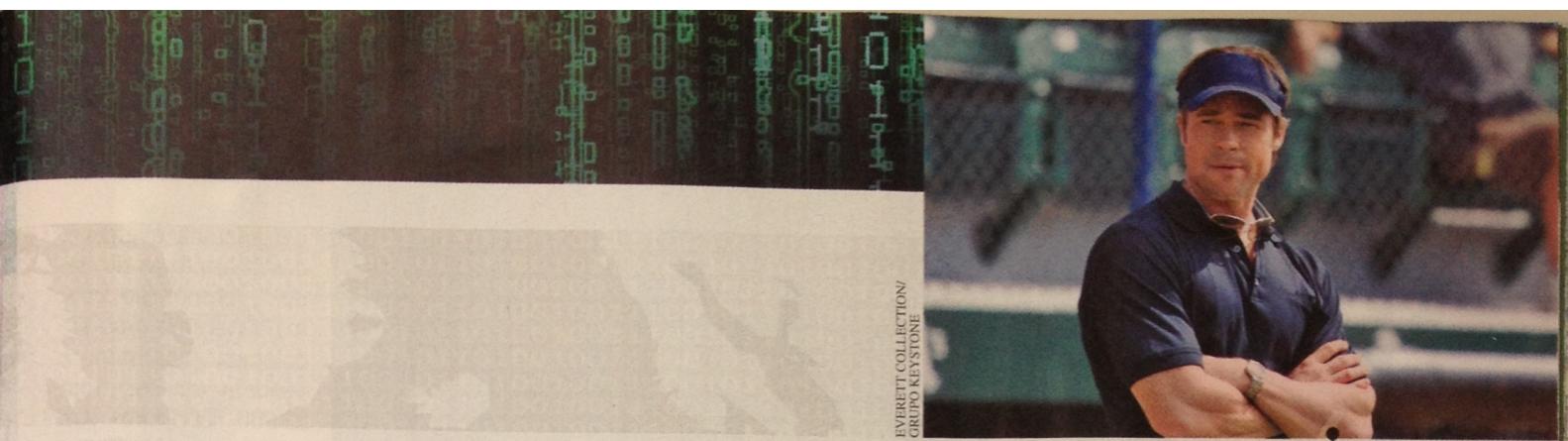
de dados são trocados por smartphones e tablets conectados à internet



10 petabytes

correspondem aos e-mails enviados





EVERETT COLLECTION/GRUPO KEYSTONE

VELOCIDADE

Flops

Unidade para calcular a velocidade de processamento de computadores: equivale à capacidade de realizar um cálculo matemático simples em um segundo

Kiloflops (1 000 flops)

Processamento de um **supercomputador** em 1951



Megaflops (1 000 000 de flops)

Velocidade alcançada por **computadores pessoais** no início dos anos 90



Gigaflops (1 000 000 000 de flops)

É nesta faixa que operam os **PCs** mais comuns, desses usados em casa



Teraflops (1 000 000 000 000 de flops)

Capacidade da próxima geração de videogames, o PlayStation 4 e o **Xbox 720**, que serão lançados neste ano



Petaflops (1 000 000 000 000 000 de flops)

Apenas supercomputadores chegam a esse patamar: o mais poderoso deles, o americano **Titan**, roda a 27 petaflops



VOCÊ
ESTÁ
AQUI

Exaflops (1 000 000 000 000 000 000 de flops)

Pelo mais conhecido parâmetro da computação, a Lei de Moore, que estima que a capacidade de processamento dobra a cada dois anos, um supercomputador deve alcançar essa velocidade na década de 2020

O Big Data na prática

Google

Permitir que centenas de milhões de usuários recebam resultados de suas buscas em meio segundo não é o grande desafio do Big Data no Google. O tesouro de verdade está nos zilhões de informações que os usuários fornecem aos computadores do Google. Sempre que alguém faz uma busca informa ao Google sua localização geográfica, seus hábitos de navegação, seu tempo de permanência em cada página da web e, assim, revela suas predileções. Cada clique é uma informação. Em breve, o movimento dos olhos pela tela também poderá ser capturado pelas webcams. Big Data é a associação de todos esses dados aparentemente desestruturados e conectá-los de forma que passem a fazer sentido e ter valor

amazon.com

De cada 100 pessoas que entram em um e-commerce como a Amazon, a maior loja on-line do planeta, apenas duas cumprem todas as etapas para efetivamente comprar algum item. Sobre esses 2% de usuários, a loja sabe quase tudo: nome, endereço, nacionalidade, e-mail, tipo de cartão de crédito que usa, produtos preferidos e hábitos de compra. Esses são dados estruturados. O desafio do Big Data é capturar dados dos 98% que não chegaram até o fim, mas foram deixando pelo caminho informações valiosas que vão ajudar a Amazon a entender as razões pelas quais a venda não foi concretizada. A partir dessas análises, a empresa poderá aprimorar seu site, diminuindo as etapas da escolha até a compra

MONEYBALL

No filme *Moneyball* (no Brasil, *O Homem que Mudou o Jogo*), transposição cinematográfica de um caso real, o personagem interpretado pelo ator Brad Pitt despreza as informações tradicionais de seus assessores sobre os jogadores do time de beisebol que ele dirige e passa a se guiar exclusivamente pelos números fornecidos por um economista e estatístico. Os assessores eram capazes de dizer quantas bolas cada jogador havia acertado em determinada temporada. A aplicação de Big Data do estatístico fornecia isso e muito mais: a porcentagem de acerto de cada jogador em cada uma das posições do campo ocupada por ele. Com os dados projetados sobre o desempenho futuro de cada atleta, o personagem de Brad Pitt conseguiu prever quantos pontos cada jogador faria nos jogos seguintes. Ele escalou o time assim. O método hoje é adotado por quase todas as grandes equipes profissionais de beisebol nos Estados Unidos

ShotSpotter

A maneira clássica da geometria para localizar com exatidão um ponto é a intersecção de três retas que se cruzam exatamente sobre ele. Uma aplicação de Big Data contra o crime adotada por dezenas de cidades americanas, o ShotSpotter, une a geometria euclidiana com as informações digitalizadas enviadas por milhares de sensores sonoros. Quando um tiro é disparado na cidade, seu som é captado pelos sensores e enviado a um computador que faz em milésimos de segundo a triangulação necessária para determinar exatamente o local do incidente

e ordenados, teremos de abraçar as imperfeições. A busca pela causa, essa obsessão humana para entender o mundo, pode ser substituída pela correlação. O Big Data oferece significado, correlação, mas não oferece a causa. Uma pesquisa da Universidade de Cambridge, com base em dados de 58 000 usuários do Facebook, descobriu uma relação inusitada: pessoas com alto QI têm apreço especial pela voz do ator Morgan Freeman. Por quê? Ninguém sabe. O MasterCard, analisando 65 bilhões de transações de 1,5 bilhão de clientes em 210 países, descobriu o seguinte: o consumidor que enche o tanque do carro por volta das 16 horas tende a gastar entre 35 e 50 dólares na hora seguinte num mercado ou restaurante. De novo: ninguém sabe por quê.

Em outro impacto colossal, a própria democracia representativa poderá ser substituída pela democracia direta. No mundo do Big Data, um cidadão comum pode ter tanta informação quanto um senador. Hoje, o iPhone de um camelô na Praia de Ipanema acessa, numa semana, mais informação do que Napoleão Bonaparte teve em toda a vida. É uma nova civilização. Na utopia de Malone: "Pela primeira vez na história, todas as relações e transações poderão se dar entre iguais". Na economia, a produção em massa está sendo substituída pela produção customizada, que agora será sucedida pela produção personalizada. No século XX, o valor transitou de bens físicos (terras, fábricas) para intangíveis (marcas, propriedade intelectual). Agora, nova transição — desta vez para os dados. Na economia industrial, o petróleo é o ouro negro. Na economia do futuro, o ouro é o dado.

A Axiom, localizada no Arkansas, é



ROB HANCOURT/MACEJARIE UNIVERSITY 2012/FROM THE HUMAN FACE OF BIG DATA

INFORMAÇÕES NO FUNDÔ DO OCEANO

Desde 2007, a organização de cientistas marinhos Integrated Marine Observing System (em inglês, Sistema Integrado de Observação Marinha) rastreia os oceanos que banham a Austrália com uma rede de submarinos autônomos, estações de monitoramento e uma imensa variedade de sensores flutuando na água, colocados no fundo dos oceanos ou acoplados a animais como elefantes-marinhos. O objetivo é coletar terabytes de informações para servir de base a pesquisas científicas sobre a fauna, a flora, a qualidade da água e o movimento de correntes. O projeto já reuniu mais de 300 estudiosos, que estão publicando uma média de 1 000 artigos acadêmicos por ano. É um número sem precedente para esse campo da ciência, que antes dependia de coletas de dados que demoravam meses, por vezes anos, para estar disponíveis.

considerada o gigante das "corretoras de dados", como são chamadas as empresas que coletam, analisam e vendem dados. Tem 23 000 servidores que processam 50 trilhões de informações por ano. Tem 1 500 informações sobre 500 milhões de consumidores no mundo. Seu cadastro inclui a maioria dos adultos dos EUA. As corretoras de dados já foram apelidadas de "ciberazzi", mistura de cibernetico com paparazzi. O grosso das informações vem do universo digital. Diz o empresário David Gorodysky, dono da AnchorFree, empresa de software do Vale do Silício: "Há mais dados on-line do que dentro da casa das pessoas". Ou seja: o cuidado de fechar a porta de casa ao sair deve ser redobrado na vida digital.

A abundância e a variedade de dados seriam inúteis se fosse impossível analisá-los. O gênio humano, inspirado em séculos de matemática e filosofia, criou o instrumento adequado para o trabalho: o algoritmo. A pesquisa de Cambridge, que relacionou QI à voz de Freeman, empregou um algoritmo capaz de descobrir características não reveladas no Facebook. O resultado é espantoso. Em 95% dos casos, os algoritmos descobriram a raça do usuário. Em 88%, a orientação sexual. Em 80%, religião e posição política. São algoritmos complexos? Michal Kosinski, da equipe de Cambridge, responde: "São muito simples. Nossa meta não era impressionar, mas mostrar que se pode deduzir muita coisa com base no nosso rastro digital". Portanto, para que o Big Data cumpra a promessa de refazer a civilização, um instrumento essencial são os algoritmos, o assunto da reportagem seguinte.

COM REPORTAGEM DE FILIPE VILICIC, DE BOSTON

UMA REPORTAGEM SOBRE ALGORITMO

DETERMINAR SE
VALE A PENA LER

NÃO

SIM

JÁ SEI TUDO
SOBRE O ASSUNTO

SIM

NÃO

VAI
AJUDAR A
ENTENDER
O BIG DATA

Criado na Antiguidade, ele triunfou na era digital. É a ferramenta para extrair riqueza de montanhas de dados sem utilidade

ANDRÉ PETRY, DE NOVA YORK

VOU LER E
DECIFRAR O MUNDO



FEDEX 2012/FROM THE HUMAN FACE OF BIG DATA

Na tarde de 6 de maio de 2010, o mercado financeiro em Nova York começou a despencar. O índice Dow Jones, que mede o desempenho das ações das trinta maiores empresas americanas, entrou em queda livre sem que ninguém entendesse a razão. Eram 14h42. Raramente o Dow Jones oscila mais de 300 pontos durante um dia inteiro. Naquele tarde, às 14h47, apenas cinco minutos depois do começo da queda, o índice já havia desabado 700 pontos, deflagrando uma onda de pânico entre investidores. Era então, e ainda hoje o é, a maior queda da história em tão pouco tempo. Quase 1 trilhão de dólares viraram pó em questão de segundos. De repente, sem razão aparente, o índice voltou a subir — furiosamente. Em um minuto, o Dow Jones estava no patamar normal. O caso ficou conhecido como o crash-relâmpago da Bolsa de Nova York.

Tocando fogo nos bastidores da crise, não havia centenas de nervosos investidores, nem multidões de operadores do mercado disputando lances freneticamente. Havia algoritmos, apenas algoritmos — esses diretores invisíveis que silenciosamente e incessantemente executam tarefas no universo digital. Na definição clássica, algoritmos são uma sequência de instruções, como a que está na página ao lado como título desta reportagem, que permite que se chegue a uma conclusão sobre que tipo de

ação tomar — no caso, ler ou não ler esta reportagem. A origem do termo remonta ao matemático persa do século IX cujo nome entrou na corrente do pensamento ocidental pelo árabe clássico como Al-Khwarizmi — que virou *algorismos* em latim. No título da página ao lado, a sequência está escrita em português. No mundo da computação, ela vem escrita numa linguagem que os programas entendem. São camadas e mais camadas de notações binárias — quanto mais camadas houver, mais complexo será o algoritmo.

A ideia de uma sequência algorítmica foi inspirada na vida cotidiana. Para escovar os dentes, por exemplo, usamos um algoritmo. Primeiro, é preciso pegar a escova e o creme dental, abrir a tampa do creme, apertar o tubo sobre as cerdas da escova, pôr a escova na boca e esfregá-la contra os dentes. Quem não seguir essa ordem — querendo, por exemplo, aplicar o creme dental antes de abrir o tubo — vai fracassar. De onde se pode concluir que há algoritmos bons e algoritmos ruins. Mas não é uma questão simples de resolver. Depois de meses de investigação sobre o crash-relâmpago em Nova York, descobriu-se que o sistema computadorizado de compra e venda de ações dos EUA entrara em colapso, produzindo a queda e a subsequente subida das ações. Os algoritmos eram ruins? Até hoje não se sabe se saíram do controle por acidente ou por malfilia

O BIG DATA EM PACOTES

A americana FedEx, a mais famosa transportadora de cargas do mundo, é pioneira no uso do Big Data. Desde a fundação, em 1973, quando operava em 25 cidades, a empresa monitora os pacotes que transporta, o que resulta em um enorme banco de dados. Algoritmos permitem aos computadores da FedEx processar em tempo real e sem interrupção cerca de 2,5 milhões de transações diárias realizadas em 220 países por 90 000 veículos e 690 aviões. Nos anos 80, a empresa tornou-se uma das primeiras companhias privadas a escanear cada encomenda. Hoje, a FedEx usa sensores que medem a temperatura, a localização e a exposição à luz de pacotes — dados que são transmitidos diretamente ao cliente.

de programadores espertalhões, dispositos a faturar milhões de dólares em segundos. O então senador Ted Kaufman, de Delaware, resumiu a perplexidade geral: “Os algoritmos estão fechando os negócios. É tudo automático. No fundo ninguém sabe o que está acontecendo nessas transações”.

Os algoritmos podem ser muito simples. Por exemplo: um algoritmo para descobrir o número mais alto de uma lista qualquer, aleatória. Em português, a sequência ficaria assim:

- 1) Considere que o primeiro item é o maior número;
- 2) Examine cada um dos itens restantes

na lista e, se algum item for maior que o número até então mais alto, tome nota; 3) Terminada a lista, o último item anotado será o maior.

O mesmo resultado poderia ser encontrado por um algoritmo menos elegante. Digamos, por exemplo, que o algoritmo anotasse todos os itens da lista com dois algarismos, depois examinasse um por um anotando apenas o maior. Em seguida, examinaria todos os itens com três algarismos, desprezando sempre o menor — e assim por diante. No fim, o algoritmo encontraria a resposta certa, mas depois de percorrer um caminho longo e tortuoso. É um algoritmo ruim. O professor Thomas Cormen, chefe do departamento de ciência da computação do Dartmouth College, em New Hampshire, explica que o hardware de um computador pode ser excepcional, o sistema operacional pode ser fantástico, mas, se os algoritmos forem ruins, nada funcionará direito. O algoritmo é o cérebro. Ou melhor: a rede neuronal. Como os neurônios do nosso cérebro, eles se ativam e se desativam conforme a necessidade, são dinâmicos, velozes e — o mais incrível — capazes, com a prática, de aprender e melhorar seu desempenho.

Algoritmos possantes, no entanto, são mais eficientes que nossos neurônios: mais rápidos, mais precisos (e muito mais baratos). Por isso, realizam tarefas que nos são impossíveis, como decodificar uma montanha caótica, quase infinita, de dados — o Big Data. Eis a anatomia da revolução que está acontecendo: a tectônica farta de dados pode ser entendida em segundos desde que analisada por algoritmos bem concebidos. Com isso, os algoritmos têm maiores possibilidades do que nossos neurônios de calcular todas as graduações de luz que Caravaggio pintou em suas obras-primas, decifrar o acorde de abertura que George Harrison usou em *A Hard Day's Night* e, quem sabe um dia, descobrir, afinal, se Capitu traiu Bentinho com seu amigo Escobar.

“Os algoritmos estão, literalmente, por toda parte”, diz Christopher Steiner, autor de um livro sobre o assunto. Eles regulam o preço de mercadorias conforme a oferta e a procura. Escre-



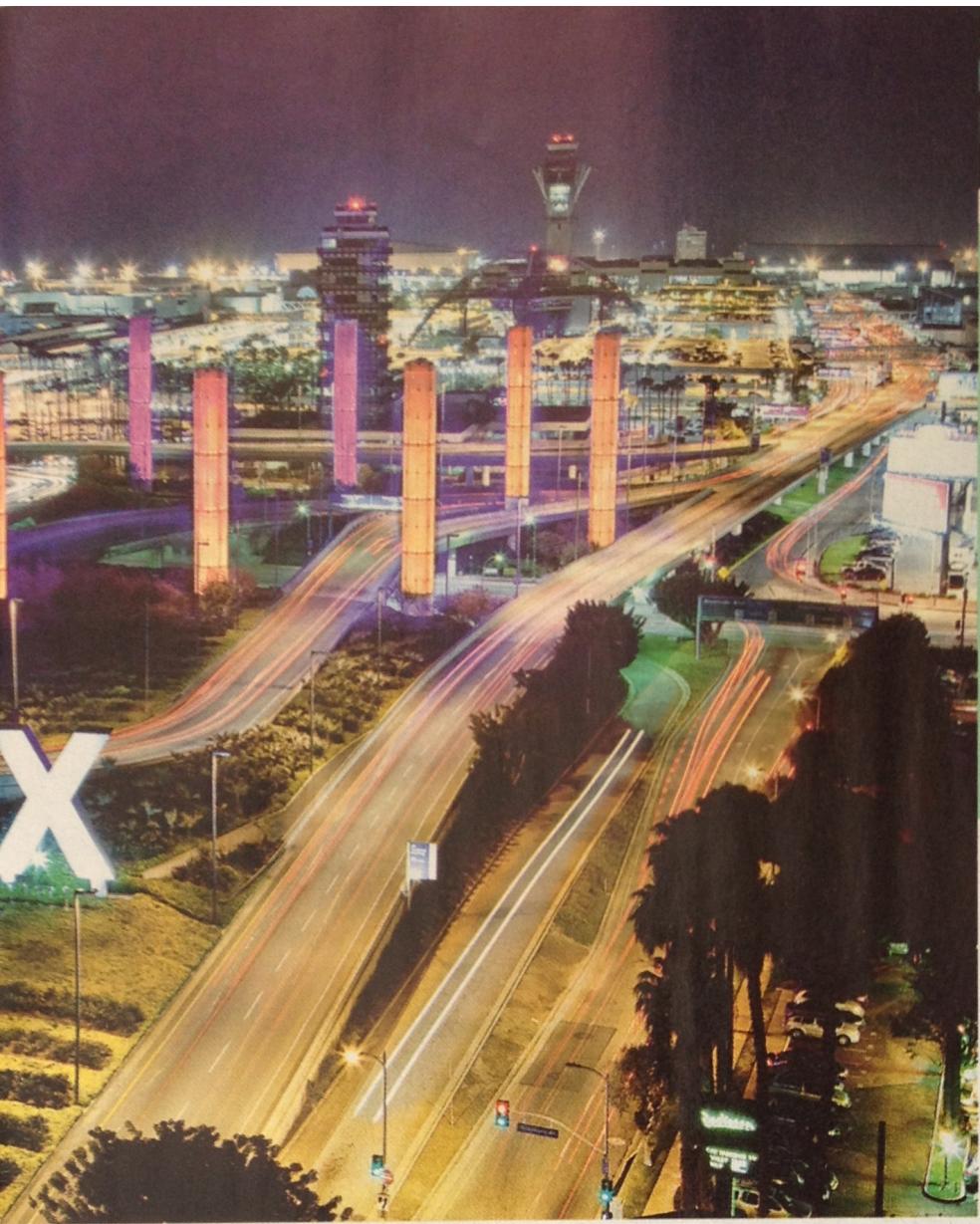
SCOTT STULBERG/CORBIS/LATINSTOCK

vem sinfonias emocionantes como se fossem um compositor polonês. Fazem diagnósticos com mais competência que médicos. Vencem os campeões mundiais de xadrez. A Amazon usa algoritmos para recomendar livros ao gosto dos clientes. A Netflix faz a mesma coisa para indicar filmes. (Em 2009, a Netflix promoveu um concurso de 1 milhão de dólares para melhorar seu sistema de recomendação. O grupo vencedor fundiu 107 abordagens algorítmicas diferentes. Apropriadamente, chamava-se Caos Pragmático.) Os algoritmos já dirigem carros, como mostram testes do Google na Califórnia, com a vantagem de que jamais excedem o limite de velocidade. Nevada e Flórida já têm lei autorizando o uso de carros sem motorista nas estradas.

Um professor bem preparado corrige trinta redações por hora. Um programa comercial utiliza algoritmos que corrigem 16 000 redações em vinte se-

gundos. A edX, empresa da Universidade Harvard e do Instituto de Tecnologia de Massachusetts (MIT), introduziu um sistema que corrige textos automaticamente e é oferecido de graça a qualquer escola que queira usá-lo. Na saúde, os algoritmos estão ficando soberanos. No mês que vem, a Heritage Provider Network, empresa de saúde, vai premiar com 3 milhões de dólares o criador de um algoritmo capaz de identificar pacientes com grande possibilidade de precisar ser hospitalizados no prazo de doze meses. A farmácia da Universidade da Califórnia, em São Francisco, é atendida por um farmacêutico-robô. Ele memoriza as receitas médicas, confronta com a ficha médica do paciente, checa o risco de interação medicamentosa e, claro, sabe na íntegra a bula de todos os remédios.

Em Wall Street, onde chegaram ainda na década de 70, os algoritmos criaram uma realidade à parte. Eles coman-



dam mais de 60% das operações no pregão, e fazem uma disputa feroz por inteligência e velocidade. Desde os Rothschild, a rapidez no mercado financeiro é essencial. Com seu eficiente sistema de pombos-correio, os banqueiros foram os primeiros a saber em Londres da derrota de Napoleão em Waterloo, ficando em vantagem na negociação dos títulos ingleses. Em Wall Street, hoje, além de incrivelmente velozes, os algoritmos precisam simular um comportamento aleatório. O algoritmo de um grande fundo de pensão, por exemplo, não pode revelar que está programado para comprar milhões de ações de uma determinada empresa. Se o fizer, algoritmos espertos perceberão, saltarão em cima das ações com uma fração de milésimos de segundo de vantagem e depois vão revendê-las para o algoritmo pateta. Alguns algoritmos ficam meses a fio escondidos, à espreita, só observando os outros.

Quando identificam um padrão de compra ou venda, dão o bote. Há algoritmos concebidos para dissimular suas ações. E outros feitos para identificar a dissimulação.

Mas será que, em outros campos da atividade humana, os algoritmos são mesmo eficientes? Será que são bons para formar um casal nos sites de relacionamento? Os psicólogos duvidam. Será que suas canções têm a vitalidade da obra de um Bach? Os especialistas dizem que não, mas o público leigo não nota a diferença. Será que deixam fluir a criatividade inovadora nas redações? Algoritmos gostam de poesia? Aprovariam um texto de James Joyce? De Guimarães Rosa? Por serem uma expressão matemática, os algoritmos carregam uma aura de verdade que pode ser enganosa. Eles são escritos por humanos e podem trazer parcialidades e preconceitos — voluntários ou não — que comprometem sua suposta objetivida-

CONTRA O TERRORISMO

Em 2007, o aeroporto de Los Angeles começou a utilizar algoritmos baseados na Teoria dos Jogos para orientar seu esquema de segurança contra possíveis ataques terroristas. Inspirado em algoritmos usados para o pôquer — nos quais se desconhece o próximo movimento do adversário —, o professor Milind Tambe, da Universidade do Sul da Califórnia, criou o programa que examina os planos que os terroristas podem considerar as suas melhores opções em termos de destruição e morte. E, com base nisso, orienta decisões. A ideia está se espalhando por outros aeroportos e também começou a ser aplicada na área de segurança de transporte.

de. Um algoritmo que seleciona as frases mais “importantes” do Twitter precisa ser programado para definir o que é “importante” — e, nisso, elimina-se a neutralidade. Explica Nick Diakopoulos, especialista em aplicativos de notícias: “Pode ser fácil sucumbir à falácia segundo a qual, por serem sistemáticos, os algoritmos são, de algum modo, mais objetivos. Mesmo os robôs têm preconceitos”. Até algoritmos aparentemente neutros podem ser falseados. A manipulação do algoritmo de pesquisas do Google já tem até nome: chama-se Google bombing. É o ato de criar associações negativas — o nome de um ator com a palavra “viciado”, por exemplo. Toda pesquisa sobre o ator virá junto com pesquisa de drogas.

Sendo mais velozes e eficientes do que nós, os algoritmos acabarão nos substituindo? Em alguns campos, isso já é realidade. Mas é improvável que o cenário seja aterrador. Por mais capazes que sejam de simular emoções humanas, os robôs ainda serão robôs, algoritmos. Ao contrário deles, a humanidade tem uma história cujos milênios foram imprimindo em nosso DNA características que nos fazem únicos. Diz o biólogo Edward Wilson, do alto de seus 83 anos e brilhante produção acadêmica: “Cada grande passo na nossa trajetória pelo labirinto evolucionário deixou um carimbo no nosso DNA. A humanidade é realmente única, talvez mais do que jamais tenhamos sonhado”.