

BIG DATA:
Como extrair volume, variedade, velocidade e valor
da avalanche de informação cotidiana
(Viktor Mayer- Schönberger; Kenneth Cukier, Editora Elsevier, 2013)

Giovanna Oliveira Lima de ABREU¹

No esclarecedor e muito oportuno livro “Big Data”, os autores Viktor Mayer-Schönberger, um professor de governança e regulamentação da internet, integrante do Oxford Internet Institute, da Universidade de Oxford, e Kenneth Cukier, editor de dados da revista *The Economist*, argumentam que a natureza da vigilância mudou.

Para os autores, as pessoas são o somatório de suas relações sociais, de suas interações e conexões *on line* com os conteúdos. Atualmente, quando o objetivo é conhecer alguém verdadeiramente, os analistas observam os dados em torno do indivíduo, os contatos e ligações estabelecidas, relacionando praticamente todos os aspectos da vida para, a partir das ferramentas e sistemas do Big Data, identificar padrões de comportamento, fazer correlações e análises preditivas.

Pela primeira vez na história, temos a capacidade de coletar e analisar quantidades massivas de informações de forma mais fácil e barata. O volume de dados disponível mais do que dobra a cada dois anos e os algoritmos aperfeiçoam-se rapidamente, ao passo que, em razão quase inversamente proporcional, os custos de armazenamento decrescem. Técnicas de análise de dados, antes acessíveis apenas às agências de espionagem, laboratórios de pesquisa e grandes conglomerados comerciais são, paulatinamente, democratizadas. De acordo com Mayer-Schönberger e Cukier, não tardará e o Big Data será uma parte imprescindível da solução para os problemas globais mais prementes como o enfrentamento das mudanças climáticas, a erradicação de doenças, além da promoção da boa governança e do desenvolvimento econômico.

^{1 1} Mestranda do Programa de Pós-Graduação em Comunicação (PPGC/ UFPB). Integrante do Grupo de Pesquisa em Processos e Linguagens Midiáticas (Gmid/PPGC). abreu.giovanna@gmail.com.

Na verdade, o advento do Big Data é o espelho da evolução tecnológica social. Trata-se da nossa sobejante capacidade de captar miríades de informações, analisá-las de imediato e tirar conclusões, por vezes, profundamente surpreendentes. Um fenômeno emergente, de amplitude crescente e ações tão distintas que atendem desde uma busca por melhores tarifas de bilhetes de avião até a dataficação de textos contidos em milhões de livros. Nosso crescente poder na computação entra em ação para descobrir epifanias nunca antes imaginadas. Estamos diante de uma revolução emparelhada à Internet.

Os aquedutos permitiram o crescimento das cidades; a imprensa facilitou o Iluminismo, e os jornais permitiram a ascensão do Estado. Mas essas infraestruturas estavam voltadas para o fluxo – de água e de conhecimento, assim como o telefone e a internet. Em contrapartida, a dataficação representa um essencial enriquecimento da compreensão humana. Com a ajuda do Big Data, não mais veremos o mundo como uma sequência de acontecimentos explicados como fenômenos naturais ou sociais, e sim como um Universo composto essencialmente por informações (p.66).

Na ausência de um bom termo que sintetize os eventos contemporâneos, os autores cunharam *dataficação* como o vocábulo mais apropriado. A ideia é que há um componente quantitativo em tudo que fazemos e que os dados são indispensáveis ao aprendizado da sociedade. As várias dimensões da realidade são transformadas em dados. As informações são usadas das maneiras mais inusitadas, revelando um valor latente e implícito. Trata-se de uma mudança que subverte séculos de práticas consagradas e desafia nossa compreensão mais básica de como tomamos decisões e compreendemos a realidade.

Buscando imprimir força ao discurso, o livro lança mão de sete capítulos para tentar definir o conceito de Big Data, para falar das transformações sofridas pelos sujeitos e pela sociedade como um todo. Exemplos práticos pululam nas páginas da obra, mostrando-nos o valor científico, econômico e social do Big Data: a Amazon recomenda os livros ideais, aumentando efetivamente volume de vendas. O Facebook conhece as preferências dos usuários, por isso a publicidade dirigida é tão certa. Há, ainda, casos no varejo, na saúde, na política e na segurança, tanto pública quanto privada.

Todavia, se os primeiros capítulos parecem exaltar os pontos positivos do Big Data, como não poderia deixar de ser, os aspectos mais preocupantes também ganham

destaque. Em “Riscos” e “Controle”, capítulos oito e nove, respectivamente, Mayer-Schönberger e Cukier fazem críticas à análise previsiva dos fatos. Há milênios a sociedade acumula experiências na compreensão e no exame do comportamento humano, mas como um algoritmo pode ser regulado? A indagação dos autores está firmada na mudança contemporânea de paradigma. Se antes as decisões eram tomadas com base em amostras exatas e causais, agora os dados são grandes, rapidamente processados e toleram imprecisão. As máquinas apontam os caminhos. Vivemos, então, sob a ditadura dos dados e a reformulação iminente do conceito de privacidade.

Essa transformação não só dificulta a proteção da privacidade como também apresenta uma ameaça totalmente nova: castigo com base em propensões, isto é, a possibilidade de usar previsões de Big Data sobre as pessoas para julgá-las e puni-las, antes mesmo que elas ajam, o que renega a ideia de justiça e livre-arbítrio... Corremos o risco de sermos vítimas da ditadura dos dados, na qual adoramos as informações e os resultados de nossas análises e acabamos usando-os de forma equivocada. Com responsabilidade, o Big Data é um instrumento útil de tomada de decisão. Se usado sem sabedoria, ele se torna um instrumento de poderosos que podem transformá-lo em uma fonte de repressão, seja ao frustrar clientes e funcionários ou, pior, ao atacar cidadãos (p.105-106).

A verdadeira revolução, percebe-se, não está nas máquinas propriamente, mas nos dados e na maneira como eles são usados. Ao alterarmos a quantidade, mudamos a essência da análise. À medida que a escala aumenta, as imprecisões quantitativas também crescem. Noções de causalidade dão lugar à correlação e aos padrões. Trata-se de um modelo que não revela a razão, a causa dos acontecimentos, mas faz previsões sobre os fatos. Sendo assim, ao trabalharmos com probabilidade e punição, como o Big Data possibilita, nos tornamos reféns dos dados, culpados prévios por ações ainda não materializadas.

O fato é que, à primeira vista, os dados podem não parecer informações pessoais explícitas, mas com os processos de análise, facilmente indicam a quem se referem ou facilitam a dedução de detalhes íntimos da vida de qualquer pessoa. No contexto do Big Data, portanto, a anonimização perfeita é impossível. A ideia que fazemos de privacidade foi corroída, e a liberdade está ameaçada, entre outras razões, porque a vigilância foi potencializada.

O valor das informações coletadas revela-se nos usos secundários, não planejados que, agora, são atribuídos aos dados. Entramos em um mundo de constantes

previsões, no qual nem sempre seremos capazes de explicar os motivos de nossas ações. A base de uma previsão de algoritmo pode ser assaz complexa visto que o Big Data opera em uma escala que transcende nossa compreensão comum. É impossível para o ser humano investigar os motivos precisos para as escolhas apontadas pelos programas, afinal são resultados fundamentados por enormes bancos de dados e cálculos estatísticos. De forma bastante perspicaz, os autores nos fazem, então, refletir: “O que significaria um médico não poder justificar uma intervenção sem o consentimento do paciente, uma vez que ele pode contar com um diagnóstico com base no Big Data? O padrão jurídico precisará mudar de “causa provável” para “causa probabilística”? Em caso afirmativo, quais as implicações para a liberdade e a dignidade humanas?” (p.11).

De certo modo, é possível deduzir que a maneira como controlamos e lidamos com a informação terá que mudar. É no último capítulo, intitulado “Depois”, que os autores estabelecem as bases dessas mudanças e, conseqüentemente, do nosso futuro, explicitando, para tanto, algumas maneiras de controlar o Big Data. Mayer-Schönberger e Cukier afirmam que simples alterações das regras vigentes não serão o bastante (p.120). As mudanças devem ser práticas e eficientes. É possível, por exemplo, que o conceito de privacidade implique em uma maior responsabilização dos usuários dos dados por seus atos. Afinal, parece lógico que os usuários precisam ter tanto o ônus da análise, quanto a responsabilidade pelo uso das informações.

Concomitantemente, a noção de justiça também precisará ser redefinida para que se possa garantir a liberdade de agir: cada pessoa deve ser avaliada por seu comportamento, não pela propensão. Embora seja assim que o sistema jurídico funcione atualmente, com o fortalecimento do Big Data será possível prever as ações humanas cada vez com mais exatidão. Ao garantir proteção à ação das pessoas, assevera-se que as medidas tomadas pelo governo, por exemplo, não serão firmadas em análises objetivas dos dados, mas em fatos reais, na responsabilidade e no comportamento. Outro ponto interessante apontado pelos autores é a necessidade de profissionais especializados para trabalhar como revisores das análises e previsões do Big Data, evitando, assim, formas mais invasivas de regulação.

Embora estejamos tratando de um fenômeno em franco desenvolvimento, talvez a lição mais profunda relacionada ao Big Data seja a de que devemos tirar proveito de

suas análises para agir mais acertadamente, experimentando com velocidade e explorando mais pistas sem que, para isso, seja necessária uma compreensão mais detalhada dos fatos.

Não podemos esquecer que o Big Data é um recurso e um instrumento, criado para informar, não para explicar. Ele aponta para a compreensão do fenômeno, mas também pode apontar para o equívoco, dependendo da utilização.