

Big Data e autonomia do usuário: os rumos de um futuro datafocado¹

Giovanna, ABREU²
Marcos NICOLAU³

Resumo

Há tempos vigilância deixou de ser uma simples questão de escrutínios individuais e dos inerentes receios à perda da privacidade. Vivemos, atualmente, uma experiência ambígua, dirigida por agências especializadas, com fins múltiplos, que não isentam ninguém. Com o fenômeno do Big Data capta-se uma torrente de dados, gerando possibilidades de conexões, até então, impensadas. Nos primórdios da computação, os legisladores perceberam que alguns usos da tecnologia ameaçariam a privacidade, por isso criaram leis direcionadas à proteção das informações. Ocorre que na era da computação pervasiva e da dataficação, qualquer tentativa de salvaguardar esses dados parece insuficiente. Assim, este artigo traz uma reflexão sobre o oposicionismo entre as ideias de liberdade como atributo essencial para o compartilhamento de informações *on-line* e a ditadura imposta pelo controle dos dados, além de analisar as implicações dessa vigilância sobre o poder de autonomia do usuário.

Palavras-chave: Big Data; autonomia do usuário; privacidade

Introdução

O big Data é produto da evolução contínua da computação e da comunicação, uma conseqüência da comunicação pervasiva e da ubiquidade da informação, da midiatização: pessoas, máquinas e ambientes comunicam-se através do mesmo código, compartilhando informações, diuturnamente.

A ambiência proporcionada pelas redes digitais metamorfoseou as práticas e relações sociais. A comunicação ganhou novos matizes que, por meio dos fluxos, se depreende infinitos significados e sentidos para a informação apreendida. O valor de um dado já não se mantém apenas no uso primário, mas surgem fins secundários não planejados.

¹ Artigo apresentado no Eixo 2 – Vigilância, Criptografia, Ativismo e Redes Sociais Federada do VIII Simpósio Nacional da Associação Brasileira de Pesquisadores em Cibercultura realizado de 03 a 05 de dezembro de 2014.

² Mestranda do Programa de Pós-Graduação em Comunicação (PPGC/ UFPB). Integrante do Grupo de Pesquisa em Processos e Linguagens Midiáticas (Gmid/PPGC).

³ Professor pós-doutor do Programa de Pós-Graduação em Comunicação (PPGC/UFPB). Coordenador do Grupo de Pesquisa em Processos e Linguagens Midiáticas (Gmid/PPGC).

O desenvolvimento de cidades inteligentes, de programas de saúde mais eficientes ou de sistemas de segurança mais efetivos, inevitavelmente, implica no tratamento dos mais diversos tipos de informação. A dataficação de todos esses dados, ainda que seja com o intuito de criar produtos e serviços para aumentar o grau de satisfação das pessoas, terá um impacto determinante sobre a privacidade.

Ainda que não seja mais possível pensar em privacidade no âmbito do indivíduo, é possível pensar em regras que estabeleçam limites para a coleta e o uso de todos esses dados, agora, disponíveis. As maneiras como produzimos, divulgamos e interagimos com a informação implicam em mudanças irrevogáveis dos valores sociais. O aumento da quantidade de dados, da capacidade de processá-los e dos usos dos dados proporcionados pelo Big Data tem gerado novos riscos para as nossas liberdades individuais.

Desse modo, este artigo tem como objetivo apresentar uma reflexão sobre o oposicionismo entre as ideias de liberdade como atributo essencial para o compartilhamento de informações on-line e a ditadura imposta pelo controle dos dados, além de analisar as implicações dessa vigilância sobre o poder de autonomia do usuário.

2 Entendendo o Big Data

Os seres humanos, há tempos, usam dados para aprender mais sobre o mundo, seja no sentido informal das observações diárias ou, mais intensamente, nas últimas décadas, no sentido formal de unidades quantificadas que podem ser manipuladas por poderosos algoritmos.

Os sumérios, por exemplo, ainda em 8.000 a.C., usavam contas de barro para controlar a quantidade de bens comercializados. Na Mesopotâmia, a escrita foi a forma desenvolvida pelos burocratas para manter o controle das informações. Todavia, em diversos momentos da história, a tarefa de coleta e análise de grandes quantidades de dados coube às instituições mais poderosas, como a igreja e o Estado que sempre tentaram manter o controle da população através da detenção de informações (MAYER-SCHÖNBERGER; CUKIER, 2013).

Entretanto, se fizermos uma comparação com os dias atuais, não foram muitos os dados coletados ao longo da história, visto que os instrumentos e métodos, em

muitos casos, eram imprecisos, caros e lentos – produto dos limites naturais da interação informacional em uma era analógica. Como alternativa, desenvolvemos práticas elaboradas para, a partir de pequenas amostras, obtermos os melhores resultados. Contudo, quando a necessidade de amostras maiores surgia, os instrumentos disponíveis não davam conta do processamento, e novas técnicas precisavam ser desenvolvidas.

A evolução tecnológica, então, imprime força e notoriedade ao Big Data, ou “mega dados”, em uma tradução literal, que trata da *dataficação*⁴ de uma massiva quantidade de informações (MAYER-SCHÖNBERGER; CUKIER, 2013). Na verdade, o conceito de Big Data não é novo, mas está recebendo uma grande atenção por razões como o aumento da capacidade de armazenamento, do poder de processamento e da enorme disponibilidade de dados.

Big Data vem chamando atenção pela acelerada escala em que volumes, cada vez, maiores de dados são criados pela sociedade. Já falamos comumente em exabytes de dados gerados a cada dia. Zetabytes começa a ser uma escala real... O que era futuro há uma década, os terabytes, hoje, temos em nossas próprias casas... Big Data é a simples constatação prática que o imenso volume de dados gerados diariamente excede a capacidade das tecnologias atuais de os tratarem adequadamente. (TAURION, 2011)⁵.

Ainda segundo Taurion (2011), diretor de novas tecnologias aplicadas da IBM Brasil, o Big Data pode ser entendido a partir de cinco características principais, também conhecidas como “os cinco Vs do Big Data”: volume, variedade, velocidade, veracidade e valor.

O volume, cuja estimativa é que dobre de tamanho a cada 18 meses, está ligado aos exabytes de dados gerados diariamente. A variedade refere-se aos dados que vem de sistemas estruturados e não estruturados. Antes do Big Data os dados não estruturados, como *tweets*, *posts* no Facebook, geolocalização, vídeos e traços do

⁴ Dataficação um fenômeno é colocá-lo em um formato quantificado de modo que possa ser tabulado e analisado. “O termo refere-se à coleta de informações de tudo que existe – inclusive informações que nunca foram pensadas como tal... – e à transformação disso em dados que possam ser quantificados. Esse conceito nos permite usar as informações de novas maneiras, como na análise previsiva... Como resultado, revelamos o valor latente e implícito das informações” (MAYER-SCHÖNBERGER; CUKIER, p. 10).

⁵ Texto retirado do artigo O caos conceitual e os 5 vs do big data. Disponível em: <http://cio.com.br/opiniaio/2012/05/11/o-caos-conceitual-e-os-5-vs-do-big-data/>. Acesso em: 20 set.2014.

comportamento inseridos em um contexto só podiam ser compreendidos por pessoas. Outro conceito determinante é o de velocidade, variável que permite o tratamento dos dados em “tempo real”. A veracidade, por sua vez, traz a certeza de que os dados fazem sentido e são autênticos. Quanto ao valor, é inerente ao processo que as organizações obtenham benefícios tangíveis, isto é, extraiam lucro das informações captadas da massa de dados.

Em essência, o avanço rumo ao Big Data é uma evolução da antiga busca da humanidade por medir, registrar e analisar o mundo, com a diferença de que, agora, qualquer tipo de informação digital pode ser analisado em tempo real. Todavia, é preciso saber como extrair e tabular as informações dos lugares mais improváveis. Muitas vezes, elementos que nunca foram tratados como dados ou pensados como detentores de informações de qualidade têm, hoje, se devidamente analisados, um enorme valor.

Portanto, não seria exagero afirmar que, embora sejamos capazes de coletar e analisar informações em escala massiva, ainda há um limite no volume de dados gerenciáveis. De fato, os problemas de contagem e tabulação não existem mais da mesma forma: sensores, GPSs e internet, por exemplo, coletam dados passivamente, enquanto os computadores lidam com os números de forma cada vez mais eficiente. Governos e empresas não têm apenas uma amostra, não raro se servem de todos os dados disponíveis acerca do problema. Examina-se todo o universo do fenômeno, acurando a análise, compreendendo melhor os padrões, tornando-se capaz de fazer previsões acerca dos fatos.

O Big Data, todavia, não dará respostas a perguntas mal formuladas ou processos equivocados. Mais do que o desejo de coletar, quantificar e registrar, a dataficação exige um conjunto específico de instrumentos e profissionais. Isso significa que apenas dispor dos dados não é suficiente, mas é essencial utilizar meios e conhecimentos apropriados para interpretá-los.

Muitas empresas não entendem o que é ou como extrair valor das ferramentas de Big Data. A manipulação dos dados, comumente, ocorre de maneira equivocada. Em numerosas ocasiões, os dados não provêm do sistema adequado, contêm erros ou não se encontram no formato devido para a sua integração com outros dados. Há, ainda, a carência de recursos semânticos eficazes, que compreendam o contexto e a

mensagem dos dados na medida das necessidades. Um dos exemplos que melhor retratam as dificuldades em lidar com as enormes quantidades de associações advindas do Big Data é a Google Flu Trends: ferramenta construída pela Google para antecipar as estatísticas dos Centros de Controle e Prevenção de Doenças (CDCs), que baseiam os resultados de suas pesquisas nas informações repassadas pelos médicos das clínicas e hospitais.

Ocorre que a rápida disseminação da doença seguia em caminho contrário ao, relativamente longo, período necessário para transferência das informações às organizações centrais. A alternativa criada pela Google soou, obviamente, como a panacéia contemporânea. Afinal, a empresa que recebe e salva mais de três bilhões de pesquisas diariamente, tinha muitos dados com os quais trabalhar.

A ideia era encontrar uma relação entre as pessoas infectadas e os termos pesquisados na internet relacionados à doença. Embora nem todas as pesquisas fossem feitas por doentes, um padrão surgiu quando as buscas relacionadas ao H1N1 foram comparadas. O que os pesquisadores da Google não contavam é que os resultados obtidos somassem mais que o dobro da quantidade de casos de H1N1 nos Estados Unidos. Por um período maior que dois anos, os números divulgados foram superiores em 100 das 108 semanas do surto de inverno.

Assim como os editores do jornal Chicago Tribune acreditaram que poderiam prever o vencedor das eleições de 1948 – e eles estavam errados – a Google acreditou que apenas os seus métodos de Big Data eram capazes de produzir, em tempo real, uma imagem mais precisa das tendências da gripe do que os antigos métodos de coleta e análise de informações. Isso é uma forma de "arrogância automatizada" ou talvez causada pelo Big Data, e pode ser vista em todo esse falatório em torno do Big Data, hoje. Só porque empresas como a Google podem acumular uma quantidade impressionante de informações sobre o mundo não significa que eles serão sempre capazes de processar essas informações para produzir um quadro preciso do que está acontecendo, especialmente se as informações recolhidas forem erradas (WALSH, 2014) ⁶.

⁶ Trecho da reportagem de Brian Walsh, Google's flu project shows the failings of big data, publicada na revista TIME. Disponível em: <http://time.com/23782/google-flu-trends-big-data-problems/>. Acesso em: 23 set.2014.

O cenário, à primeira vista, parecia perfeito, mas os engenheiros da Google equivocaram-se grandemente. A verdade é que o Big Data pode nos deixar tão fascinados com a promessa e o poder suscitado que as limitações são esquecidas: as correlações encontradas não implicam relações de causa e efeito, isto é, as relações entre os termos de busca estabelecem-se por diversas circunstâncias, com os mais variados fins. No caso do Google Flu Trends, em específico, as pessoas selecionadas não necessariamente estavam infectadas, e poderiam estar em busca de informações para ajudar alguém doente, para aprender os métodos de prevenção ou por simples curiosidade.

Os resultados das pesquisas que se valem do Big Data, portanto, podem estar acorrentados à previsões que não condizem com o conteúdo real da informação. Exacerba-se, então, a antiga crença de que os números são infalíveis. Embora o Big Data seja um caminho sem volta, o fato de se tratar de um fenômeno relativamente novo e em curso abre precedente para que a qualidade dos dados ainda seja baixa, tendenciosa, mal analisada ou usada deturpadamente.

É preciso compreender que nem toda situação humana complexa pode ser reduzida a linhas em um gráfico, porcentagens ou números em uma tabela.

3. Quando tudo são dados

Apesar dos equívocos aos quais ainda estamos sujeitos e das já consagradas técnicas de coleta de dados, é inegável que o Big Data representa mudanças significativas na forma como analisamos informações e, por sua vez, no modo como entendemos e organizamos a sociedade.

Pensar em Big Data é lembrar que as amostragens saíram de cena para que grandes quantidades de dados ou todas as informações acerca de um determinado problema sejam consideradas. Essa mudança subverte séculos de práticas consagradas e desafia nossa compreensão mais básica de como tomamos decisões e compreendemos a realidade.

Há uma caça ao tesouro em andamento, motivada pelas ideias a serem extraídas dos dados e pelo valor adormecido que pode ser despertado por uma mudança de causalidade para correlação, mas não há apenas um tesouro. Cada banco de dados tem um valor intrínseco oculto, e há uma corrida para descobri-lo e captá-lo. O

Big Data altera a natureza dos negócios, dos mercados e da sociedade (MAYER-SCHÖNBERGER; CUKIER, 2013, p.10).

Perceber que há um valor especial em um amontoado de dados que falta em quantidades menores foi o âmago da questão. Não se trata mais de informações exatas, pequenas e causais, mas de padrões e probabilidades, de previsões. Os dados estão, paulatinamente, tornando-se indispensáveis ao aprendizado e desenvolvimento da sociedade. O Big Data atinge todas as práticas humanas, desde as comunicacionais, econômicas, sociais, políticas até as jurídicas. Dois dos grandes incentivadores desse alcance prodigioso são os valores decrescentes do armazenamento e das diversas ferramentas para tratar grandes volumes de dados em tempo real.

Não seria precipitado afirmar que, atualmente, tudo são dados. O mundo é, em si mesmo, composto por dados. Geolocalização da natureza, objetos e pessoas já configuram, claramente, informação. As palavras tornaram-se dados, são “estatisticamente significantes” e têm suas conexões estabelecidas através dos algoritmos. As interações revestiram-se de atributos que as caracterizam como dados. Nossas relações, experiências, crenças, os elementos intangíveis do cotidiano foram transformados em dados e podem ser usados, independente da nossa vontade, para os mais distintos fins.

A quantidade de dados digitais no mundo deve crescer do atual 1,8 zettabyte para 7,9 zettabytes em 2015, como prevê a IDC. Isso significa que, daqui a três anos, toda a informação do mundo poderia ser armazenada em 493 bilhões de iPads. A Gartner, outra empresa de consultoria, sustenta que esse volume de informações expandir-se-á nos próximos anos a um ritmo de, no mínimo, 59%⁷, e a centelha que aciona essa explosão é a proliferação de aparelhos e plataformas que geram dados diuturnamente. São celulares, GPS, redes sociais, câmeras e sensores diversos.

Uma prova significativa que endossa esse alcance foi o ocorrido durante o Fórum Econômico Mundial, em 2012: o austero encontro de Davos abriu espaço em sua agenda de debates sobre mercados financeiros e conjuntura macroeconômica para discutir a informação sob diversos aspectos. Após o evento, o fórum publicou o estudo "Big Data, grande impacto: novas possibilidades para o desenvolvimento

⁷ Informações retiradas da reportagem De ‘moneyball’ a davos: o big data se abre para o mundo, publicada no site do Jornal O Globo. Disponível em: <http://oglobo.globo.com/sociedade/tecnologia/de-moneyball-davos-big-data-se-abre-para-mundo-4460918#ixzz3ERvctOuG>. Acesso em: 25 set.2014.

internacional", apontando o Big Data como um poderoso aliado para a solução de problemas sócio-econômicos.

Em princípio, pode parecer que juntar todos esses dados em uma aplicação prática seja algo para um futuro longínquo. Contudo, muitas ideias já são realidades bem-sucedidas. Soluções tecnológicas para o problema das super populações nas cidades, por exemplo, estão sendo criadas em vários lugares do mundo, tanto por pequenas empresas quanto por indivíduos, multinacionais e governos.

O conceito de *smart cities*, ou cidades inteligentes, definidas como ambientes inteligentes, que embutem tecnologias da informação e da comunicação, criando ambientes interativos, que trazem a comunicação para o mundo físico. A partir desta perspectiva, em uma cidade inteligente as tecnologias de comunicação e de informação, bem como os sistemas de sensores, desaparecem à medida que se tornam embutidos nos objetos físicos e nos ambientes (STEVENTON; WRIGHT, 2006).

Municípios como Songdo, na Coreia do Sul, e Masdar, em Dubai estão sendo construídos, literalmente, desde a fundação, com quase toda a infraestrutura conectada à internet⁸. Para atingir a meta de ser uma cidade sem trânsito, as ruas possuem sensores no asfalto que analisam o tempo de deslocamento dos veículos em engarrafamentos. Os postes de iluminação pública também possuem sensores que diminuem a intensidade das luzes quando não há ninguém nas ruas. Até as garrafas de refrigerante seguirão o conceito “inteligente”, computando descontos nos impostos dos moradores que jogarem o recipiente no cesto de reciclagem correto.

Empresas de tecnologia como a IBM e a Siemens, criaram departamentos de pesquisa na área. Instituições de ensino, como o MIT, possuem centros de investigações e protótipos para cidades inteligentes, e muitos governos apostam no conceito. No setor público, a União Européia foi uma das pioneiras. Lançou, em 2007, um programa de incentivo para que setenta cidades médias pré-selecionadas invistam em inovações. Capitais como Barcelona, que não participaram do programa, fizeram seus próprios investimentos. A cidade espanhola reconstruiu um bairro em uma antiga área industrial do século 19, há muito esquecida pelos habitantes de Barcelona e pelos

⁸ Informações retiradas da reportagem Asia constrói cidades inteligentes “do zero” ao custo de US\$ 102bi. Disponível em: <http://www1.folha.uol.com.br/fsp/tec/116712-asia-constroi-cidades-inteligentes-do-zero-ao-custo-de-us-102-bi.shtml>. Acesso em: 26 set.2014.

turistas. O Distrito22@⁹, como ficou conhecido, a exemplo das cidades inteligentes orientais, funciona como um laboratório de testes de soluções urbanas.

O Big Data tem usos marcantes na área do varejo também. A empresa norte-americana Walmart¹⁰ é reconhecida como referência por transformar os dados captados *on line* em soluções para impulsionar as vendas de suas lojas físicas. Os *softwares* desenvolvidos pela empresa conseguem, entre outras coisas, monitorar momentos de discussões intensas sobre esportes na internet em diferentes cidades dos Estados Unidos para, em questão de horas, as lojas físicas dessas regiões exibirem nas vitrines produtos dos times em questão.

Com mais pessoas comprando pela *web* e usando aplicativos de *smartphone* para fazer listas, encontrar ofertas e comparar preços, as diferenças entre as empresas de *e-commerce* e lojas físicas começam a se confundir. Assim, em 2011, a varejista criou a Walmart Labs, uma subsidiária de tecnologia considerada como parte crucial do plano da empresa para expandir seus negócios na internet. A meta é ultrapassar, em diversos sentidos, gigantes como a Amazon.

Outro feito interessante ocorreu em agosto de 2012, quando a Walmart se tornou o primeiro grande varejista a lançar um motor de busca caseiro, ao invés de usar a tecnologia de um fornecedor. O *software*, cunhado Polaris, utiliza algoritmos de busca semântica capazes de compreender alguns significados implícitos presentes em pesquisas *on line*. O Polaris pode reconhecer, por exemplo, a possibilidade de que uma mulher ao digitar "*flat*"¹¹ esteja em busca de sapatos, não de uma televisão de tela plana.

Um pequeno ajuste no algoritmo de *e-mail* que trabalha com um banco de dados de milhões de contas de clientes da empresa foi capaz de estabelecer ligações entre a previsão meteorológica local de um destinatário e as sugestões de compra que seriam enviadas como *e-mail marketing*. Para o vice-presidente do Walmart Labs, Chris Bolte, líder da equipe responsável pela captação do usuário, a vantagem do Walmart sobre muitas empresas *on line* encontra-se na pesada escala de operações

⁹ Informações retiradas do *site* <http://www.22barcelona.com/>. Acesso em: 26 set.2014.

¹⁰ Informações retiradas da reportagem Walmart's new high-tech labs: you're not in arkansas anymore. Disponível em: <http://www.technologyreview.com/news/429589/walmarts-new-high-tech-labs-youre-not-in-arkansas-anymore/>. Acesso em: 27 set.2014

¹¹ Flat, termo homônimo, em inglês, que pode significar “plano”, como no caso de “tela plana”, ou “baixo”, no caso de sapatos.

físicas que combinadas ao Big Data revelam "um amplo panorama da economia vigente".¹²

Especificamente, no Brasil essas aplicações ainda são limitadas. A tecnologia é usada de forma pouco madura, muito restrita aos dados estruturados, tradicionais. Setores como o varejo, as telecomunicações, o petrolífero, o mercado financeiro e as áreas de mídia e entretenimento começam a perceber que vale a pena tentar compreender a enxurrada de dados que os cercam.

A Renner, gigante do varejo nacional, por exemplo, usa Big Data para identificar as necessidades imediatas dos consumidores. A empresa monitora, em tempo real, o fluxo de mercadorias das lojas ao cruzar os dados de localização emitidos pelos GPS dos caminhões dos seus fornecedores com os níveis dos seus estoques. A rede também acompanha a aceitação dos seus produtos de forma instantânea nas redes sociais. Assim, é possível saber se a roupa do comercial recentemente veiculado agradou. Em uma reportagem do Jornal O Globo¹³, a empresa informou que brevemente será possível deduzir as estratégias dos concorrentes analisando os preços dos fornecedores.

A tecnologia que cruza coordenadas de GPS, *posts* de redes sociais e todo tipo de dado em tempo real, modificando a realidade já existe e está invadindo empresas e governos tanto do Brasil, quanto do mundo. A promessa é de uma revolução, tanto na sociedade, enquanto entidade coletiva, quanto no âmbito do indivíduo. Mudarão aspectos muito particulares da nossa vida e o modo como pensamos, obrigando-nos a transformar, pragmaticamente, princípios, há muito instaurados.

Outro exemplo brasileiro que merece destaque é o caso do pré-sal, nome dado às reservas de hidrocarbonetos em rochas calcárias que se localizam abaixo de camadas de sal. O petróleo é extraído de camadas entre cinco e sete mil metros de profundidade abaixo do nível do mar. A executiva de tecnologia para a região das Américas da EMC¹⁴, Patrícia Florissi, afirma que a viabilização do pré-sal como

¹² Informações retiradas da reportagem Walmart's new high-tech labs: you're not in arkansas anymore, publicada na revista MIT Technology Review.

¹³ Informações retiradas da reportagem De 'moneyball' a davos: o big data se abre para o mundo, publicada no site do Jornal O Globo

¹⁴ EMC Corporation é uma empresa multinacional norte-americana reconhecida como uma das maiores fornecedoras de software do mundo. A empresa negocia também sistemas para infraestrutura de informação e serviços relacionados.

reservatório petrolífero passível de exploração só foi possível por causa do Big Data e da economia em nuvem.¹⁵

A discussão sobre a existência de uma reserva petrolífera na camada pré-sal ocorre desde a década de 1970. Os geólogos da Petrobrás acreditavam no fato, mas não possuíam tecnologia suficiente para a realização de pesquisas mais avançadas. Ocorre que a tecnologia torna ágil o processamento de dados sísmicos captados pelas sondas que procuram petróleo no fundo do mar. Milhões de variáveis exigem intermináveis simulações de imagens, e apenas os sistemas que lançam mão do Big Data são capazes de realizar o trabalho em tempo hábil.

A medicina, a política, todas as áreas do conhecimento humano podem ser beneficiadas pelo Big Data. A verdade é que sempre que a dataficação é bem-sucedida, obtemos muito valor agregado pelas informações subjacentes, e grandes ideias são reveladas.

4. Autonomia do usuário e privacidade em foco

Em sentido literal, autonomia pode ser definida como a faculdade de uma entidade se governar por leis próprias, de estabelecer as próprias normas. O vocábulo, que deriva do grego antigo, *autós* (de si mesmo) e *nomos* (lei), tem o seu conceito aplicado em áreas diversas como a moral, a filosofia, a política e a educação.

Embora a ideia de autonomia não seja absoluta, a definição que, por hora, parece mais apropriada é a do *Vocabulário Técnico e Crítico da Filosofia*: "a condição de uma pessoa ou de uma coletividade cultural, que determina, ela mesma, a lei à qual se submete" (LALANDE, p. 115). Se autonomia pode ser uma "condição", então é possível afirmar que ela se dá no mundo, e não apenas na consciência dos sujeitos. A construção desse "direito" envolve o poder, a liberdade de conceber as próprias leis e a capacidade de realizar. O pensar autônomo precisa ser também fazer autônomo (ZATTI, 2007). Para que haja autonomia, então, os dois aspectos devem estar presentes. Contudo, é no mundo que a capacidade de realizar é posta em prática.

¹⁵ Informação retirada da reportagem A revolução do Big Data, publicada no site do Jornal Extra. Disponível em: <http://extra.globo.com/noticias/celular-e-tecnologia/a-revolucao-do-big-data-4471312.html>. Acesso em: 26 set.2014.

As leis naturais e civis, as convenções sociais, além de outros condicionantes atuam como limitadores da autonomia.

Ao longo da história, a ideia de autonomia vai sendo elaborada, ganhando força e adquirindo significados diferentes. O pensamento grego desenvolveu as primeiras noções, mas o conceito surgiu, de fato, na modernidade (ZATTI, 2007). A concepção de dignidade humana fundada por Descartes foi, em certo sentido, recuperada por Kant. Para o filósofo, a autonomia da vontade é considerada o princípio supremo da moralidade. A autonomia verdadeira é dependente de uma lei promulgada pela vontade que, por sua vez, deve ser uma lei universal válida para todo ser racional. Caso contrário, a lei estaria condicionada a algum interesse subjetivo, e a vontade seria, então, dependente do objeto de interesse. Seria uma vontade heterônoma. Quando a vontade é autônoma, promulga leis universais isentas de todo interesse, que reclamam a obediência por puro dever (KANT *apud* ZATTI, 2007).

A essa definição de autonomia se prende a ideia de dignidade da pessoa e de toda a natureza racional. O ser racional ao participar da legislação universal, ao submeter-se à lei que ele próprio se confere, é fim em si, não possui valor relativo, mas uma dignidade, um valor intrínseco (LALANDE, 1999).

Trazendo esses conceitos para o contexto contemporâneo, percebemos que o Big Data não apenas corrói a nossa autonomia, aumentando os riscos para a privacidade, mas altera as características desses riscos. Os usos secundários que surgem durante a coleta e o processamento das grandes quantidades de dados deslocam o foco para probabilidade: algoritmos, cada vez mais potentes, calcularão a probabilidade de pessoas comprarem um produto, terem um ataque cardíaco, não conseguirem pagar um empréstimo ou cometerem um crime, por exemplo.

É certo que o Big Data marca um passo importante na busca da humanidade por quantificar e perceber o mundo. Novas formas de compreensão estão surgindo. A sociedade, de várias maneiras, tem, paulatinamente, deixado a causalidade de lado, para se aproveitar da correlação (MAYER-SCHÖNBERGER; CUKIER, 2013).

A partir das correlações, informações aparentemente díspares estabelecem uma forte relação de interdependência, criando associações. Trata-se de uma mudança de abordagem, na qual o motivo do acontecimento não é importante, apenas o evento em si. A Amazon, por exemplo, com base na coleta de dados e nas correlações

estabelecidas, recomenda produtos específicos para seus consumidores com base em suas preferências de compra. Quando uma correlação parece firme, a probabilidade de uma conexão torna-se alta também. Tanto é verdade, que mais de um terço das vendas da Amazon resultam dos sistemas de recomendação e personalização (MAYER-SCHÖNBERGER; CUKIER, 2013).

Inúmeras cidades nos Estados Unidos já utilizam o “policimento preventivo”: usando análises de Big Data, selecionam as ruas, grupos e pessoas que devem ter vigilância redobrada, pois um algoritmo identificou maior probabilidade de um crime ocorrer. As correlações, portanto, nos ajudam a captar o presente e a prever o futuro.

A prevenção de comportamentos perigosos ou ilegais é uma das bases da sociedade moderna. Medidas preventivas diminuem a liberdade das pessoas. Entretanto, muitos as vêem como um preço pequeno quando comparadas ao mal evitado. O que não pode passar despercebido é que, em tempos de Big Data, esse caminho pode terminar em uma encruzilhada.

Apesar de os sistemas ainda serem embrionários, a promessa é aperfeiçoar os mecanismos. Algoritmos farão previsões baseadas em análises de dados tão precisas que as pessoas não se tornarão culpadas pelo que fizeram, mas pelo que pretendiam fazer. Vislumbra-se, assim, um futuro de previsões individualizadas do comportamento (MAYER-SCHÖNBERGER; CUKIER, 2013).

Enquanto as previsões comportamentais usuais buscam algumas variáveis para formar um modelo mental acerca de um determinado tema, na análise não causal do Big Data, é possível, a partir de uma grande massa de informações, identificar os dados mais adequados para formular uma conclusão. Almeja-se, com previsões mais granuladas, distinguir pessoas em vez de grupos. Uma operadora de seguros, por exemplo, que atualmente leva em conta características como histórico familiar e problemas anteriores de saúde, poderia escapar das identidades grupais e fazer acordos mais individualizados.

Em muitos casos, as previsões podem parecer animadoras e vantajosas. Todavia, é preciso perceber que se os algoritmos pudessem fazer previsões com clareza e sem falhas, não teríamos mais escolhas sobre o futuro. Os dados podem não parecer informações pessoais explícitas, mas com os processos de análise facilmente dizem a quem se referem ou viabilizam deduções de detalhes íntimos da

vida de qualquer pessoa. Dessa forma, a autonomia e a possibilidade de viver com liberdade seriam negadas.

É bem verdade que previsões perfeitas são impossíveis. A análise do Big Data preverá, então, a probabilidade de alguém ter um determinado comportamento no futuro (MAYER-SCHÖNBERGER; CUKIER, 2013). Seja como for, sérias questões éticas emergem. As luzes estão acesas sobre temas como privacidade, confidencialidade, transparência e identidade. Quem, de fato tem direitos sobre todos esses dados gerados? Há limites para os tipos de inferências que podem ser feitas? Que decisões podem ser tomadas sobre as pessoas com base nas informações extraídas da dataficação?

Muitas perguntas, por enquanto, permanecerão sem respostas. O que se pode afirmar é que, de certo modo, a maneira como controlamos e lidamos com dados terá que mudar, assim como vem mudando o nosso conceito de privacidade. A era do Big Data exigirá novas regras para salvaguardar a liberdade e a privacidade das pessoas.

Considerações finais

O avanço dos processos que envolvem o Big Data é inequívoco. A capacidade de empresas e governos para coletar e analisar informações cresce exponencialmente, transformando tudo, desde negócios à política, esportes, culinária e os indivíduos, especificamente.

Na verdade, Big Data é muito mais do que apenas correlacionar tabelas de banco de dados e criar algoritmos de reconhecimento de padrões. Trata-se de dinheiro e poder. Big Data, em sentido lato, é o aumento massivo da vigilância, do lucro e do poder institucional, em detrimento da autonomia das pessoas. Com a ascensão da disponibilidade da Internet e a popularização dos sistemas de *web*, incluindo as redes sociais, alcançamos um número inimaginável de informação, que aumenta a cada dia. A aquisição do WhatsApp pelo Facebook¹⁶ e todo o assunto em torno da NSA¹⁷ demonstram o enorme valor dos dados, hoje.

¹⁶ Os usuários mais jovens estão gastando cada vez mais tempo em serviços de mensagens móveis que não possuem anúncios e oferecem mais privacidade. Com o WhatsApp, o Facebook imediatamente ganhou acesso aos dados móveis de centenas de milhões de usuários.

¹⁷ *National Security Agency* - NSA é a agência de segurança dos Estados Unidos, criada em 4 de novembro de 1952 com funções relacionadas a Inteligência de sinais, incluindo interceptação e

Certamente, não se trata, apenas, de segurança nacional e publicidade dirigida. O problema é que a nossa capacidade de revelar padrões e novos conhecimentos advindos de grandes amontoados de dados anteriormente não examinados está evoluindo mais rápido do que nossas atuais diretrizes legais e éticas.

Se não formos capazes de estabelecer limites para essa nova sociedade digital, então correremos um sério risco de vermos naufragar direitos vitais para o bom funcionamento da sociedade em prol da inovação e da conveniência.

Referências bibliográficas

ASIA CONSTROI CIDADES INTELIGENTES “DO ZERO” AO CUSTO DE US\$ 102BI. Disponível em: <http://www1.folha.uol.com.br/fsp/tec/116712-asia-constroi-cidades-inteligentes-do-zero-ao-custo-de-us-102-bi.shtml>. Acesso em: 26 set.2014

DE ‘MONEYBALL’ A DAVOS: O BIG DATA SE ABRE PARA O MUNDO. Disponível em: <http://oglobo.globo.com/sociedade/tecnologia/de-moneyball-davos-big-data-se-abre-para-mundo-4460918#ixzz3ERvctOuG>. Acesso em: 25 set.2014.

GOOGLE’S FLU PROJECT SHOWS THE FAILINGS OF BIG DATA. Disponível em: <http://time.com/23782/google-flu-trends-big-data-problems/>. Acesso em: 23 set.2014.

LALANDE, André. *Vocabulário Técnico e Crítico da Filosofia*. São Paulo: Martins Fontes, 1999.

MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. *Big Data: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana*. Rio de Janeiro: Elsevier, 2013.

O CAOS CONCEITUAL E OS 5 VS DO BIG DATA. Disponível em: <http://cio.com.br/opiniaio/2012/05/11/o-caos-conceitual-e-os-5-vs-do-big-data/>. Acesso em: 20 set.2014.

STEVENTON, Alan; WRIGHT, Steve. *Intelligent spaces: the application of pervasive ICT*. London: Ed. Springer, 2006.

WALMART'S NEW HIGH-TECH LABS: YOU'RE NOT IN ARKANSAS ANYMORE. Disponível em: <http://www.technologyreview.com/news/429589/walmarts-new-high-tech-labs-youre-not-in-arkansas-anymore/>. Acesso em: 27 set.2014.

ZATTI, Vicente. AUTONOMIA E EDUCAÇÃO EM IMMANUEL KANT E PAULO FREIRE. Porto Alegre: EdiPUCRS, 2007. Disponível em: <http://www.pucrs.br/edipucrs/online/autonomia/autonomia/autonomia.html>. Acesso em: 28 set.2014.

criptoanálise. Dedicar-se também à proteção das comunicações americanas. A NSA, que é parte do Departamento de Defesa dos Estados Unidos, está envolvida em escândalos relacionados à vigilância global, com inúmeros programas de vigilância eletrônica ao redor do mundo.