



BIG DATA **PARA O BEM COMUM**

| POR SILVIA RODRIGUES FOLLADOR E JULIE RICARD

Multiplicam-se as oportunidades de uso de rastros digitais na elaboração de políticas públicas e no desenvolvimento sustentável.

Desde o início do século XX, a maioria de nossas ações e interações tem sido mediada e capturada por dispositivos eletrônicos. Os rastros de dados deixados pelo caminho resultam no que foi batizado de *big data*. Embora a exploração do *big data* tenha sido desenvolvida por gigantes da internet, que transformaram a mineração dessas *migalhas digitais* em uma de suas principais fontes de lucro, o interesse em entender como novas fontes de dados e tecnologias podem ser empregadas na formulação de políticas públicas e no desenvolvimento sustentável tem aumentado de maneira significativa.

OPORTUNIDADES PARA O BEM COMUM

Pela sua natureza, o *big data* pode fornecer informações inovadoras sobre comportamentos humanos, como padrões de mobilidade física, de comunicação e de consumo, em níveis de granularidade sem precedentes. Análises baseadas no *big data* têm sido usadas para diagnosticar uma vasta gama de situações, tais como o mapeamento de pobreza, o monitoramento de fluxos migratórios e o planejamento de transportes.

Como exemplo dessas iniciativas, vê-se a parceria entre os pesquisadores da Flowminder, uma fundação com base na Suécia, e do Programa WorldPop, da Universidade de Southampton (no Reino Unido), que criaram formas inovadoras

para a medição dos indicadores de pobreza em Bangladesh, por meio dos registros de chamadas de telefones celulares e de imagens de satélite. Essas novas maneiras de captura de informações têm o potencial de se tornarem fontes complementares valiosas para mecanismos tradicionais de censo que, sobretudo em países carentes, dependem de coletas de periodicidade irregular e de cobertura restrita.

Atualmente, pesquisadores do *think tank* Data-Pop Alliance e do MIT Media Lab buscam quantificar, pela primeira vez, em grande escala e com alto nível de detalhes, as mudanças de comportamento de consumo individuais induzidas pela criminalidade, a partir de metadados de transações bancárias, de modo a quantificar o impacto econômico de choques de violência.

Esses são apenas alguns dos exemplos de iniciativas que usam o *big data* em benefício da Agenda 2030 para o Desenvolvimento Sustentável, da Organização das Nações Unidas (ONU), como a erradicação da pobreza, o fim da fome, a promoção da saúde e a redução das desigualdades sociais. De maneira geral, análises com base em metadados derivados do uso de celulares, cartões de crédito ou débito, cartões de transporte ou registros de navegação on-line oferecem a vantagem de produzir informações mais detalhadas, potencialmente em tempo real, e menos custosas que por outros meios. Afinal, há casos em que não existem formas alternativas para o monitoramento de determinados indicadores.

DESAFIOS LIGADOS AO USO DE **BIG DATA**

Se o aparecimento do *big data* foi acompanhado de grandes expectativas em relação ao potencial para a análise dos problemas da sociedade, logo foram apontados riscos em relação à proteção de dados pessoais. As preocupações a respeito da proteção de dados pessoais cresceram à medida que foi notado pela academia que métodos de anonimização de dados podem não ser suficientes. No mesmo período, escândalos de alcance global, como as denúncias de Edward Snowden sobre a extensão das atividades de vigilância da Agência Nacional de Segurança dos Estados Unidos, ajudaram a amplificar as preocupações com a privacidade dos cidadãos.

Além disso, é necessário ressaltar o fato de que grande parte dos dados pertence ao setor privado (como operadores de

telefonia e bancos). Dessa forma, a obtenção de acesso para alavancar as fontes de *big data* coletadas por tais empresas tem sido um desafio, por motivos éticos e comerciais legítimos. De fato, até o momento, não existem sistemas nem padrões desenvolvidos para que isso seja feito de maneira escalável e de forma ética e segura.

Dada essa conjuntura, foi criado o projeto Open Algorithms (OPAL), uma inovação sociotecnológica, sem fins lucrativos, desenvolvida pelo MIT Media Lab, pelo Imperial College London, pela Orange, pelo Fórum Econômico Mundial e pela Data-Pop Alliance. Seu objetivo é elaborar um sistema resguardado, ético e replicável de acesso a dados do setor privado para impulsionar o seu uso para fins públicos. O primeiro passo consiste na criação de uma plataforma com



TIPOS DE **BIG DATA** PARA A AGENDA DO DESENVOLVIMENTO

Tipos de dados	Exemplos	Oportunidades de uso
Metadados derivados do uso de celulares	Registros de chamadas de celulares (<i>call detail records</i> , ou CDR)	Estimar a distribuição da pobreza e o <i>status</i> socioeconômico da população para informar políticas públicas.
Transporte	Cartões de transporte (bilhete único, milhas de avião etc.), passes para pagamento automático de pedágio, sistema de posicionamento global (GPS) (rastreamento de frota, rotas e horários de ônibus)	Construir a matriz de viagem dos indivíduos, de uso de redes de transporte e de estradas para melhorar o planejamento urbano.
Transações financeiras	Cartões de crédito e de débito	Quantificar comportamentos de consumo dos indivíduos para entender padrões de comportamento e/ou realizar estimativas econômicas.
Registros de navegação on-line	<i>Cookies</i> , endereços, <i>internet protocol</i> (IP), buscas on-line	Criar indicadores para dar suporte a ações dos Objetivos do Desenvolvimento Sustentável (ODS). Por exemplo, dados on-line já foram utilizados para monitorar epidemias de <i>influenza</i> e dengue (por meio do "Google Dengue Trends").
Redes sociais	<i>Tweets</i> (Twitter API), conteúdo do Facebook, vídeos do YouTube, <i>check-in</i> pelo Foursquare	Fornecer alerta antecipado sobre ameaças que vão desde surtos de doença (<i>influenza</i> , por exemplo) até insegurança alimentar.
Sensores remotos	Imagens de satélite (Administração Nacional da Aeronáutica e Espaço – NASA, Tropical Rainfall Measuring Mission – TRMM, Landsat), dados de veículos aéreos não tripulados (veículo aéreo não tripulado – VANT, <i>drones</i>)	Identificar áreas de risco de inundação, estudar mudanças na qualidade do solo, verificar disponibilidade de água e informar possíveis intervenções no setor agrícola de países em desenvolvimento.
<i>Crowd-sourced data</i>	Mapeamento (OpenStreetMap, Google Maps, Yelp), monitoramento/relatórios (U-Report)	Preencher lacunas de dados, particularmente em áreas marginalizadas, por meio da contribuição dos próprios cidadãos. Por exemplo, em Nairóbi, no Quênia, membros das comunidades Kibera, Mathare e Mukuru mapeiam zonas antes em branco no mapa, lançando mão de ferramentas como o OpenStreetMap (Map Kibera).

FONTE: DATA-POP ALLIANCE. BASEADO EM *THOUGHTS ON BIG DATA AND THE SDGS*, DE EMMANUEL LETOUZÉ E MARIA MARTINHO, 2015, E *LEVERAGING BIG DATA FOR SUSTAINABLE DEVELOPMENT TOOLKIT*, 2016.

Análises baseadas no *big data* têm sido usadas para diagnosticar uma vasta gama de situações, como o mapeamento de pobreza, o monitoramento de fluxos migratórios e o planejamento de transportes.

tecnologia de preservação de privacidade de última geração, para a qual se podem enviar algoritmos para a consulta de dados, e não o contrário, evitando assim que os dados sejam expostos a roubos e ao uso indevido. O segundo passo é a criação de um sistema de governança participativa para co-projetar, com atores locais, o desenvolvimento e o uso dos algoritmos abertos, de modo a atender às necessidades e aos padrões locais, em vez de impor perspectivas e conhecimentos externos. Atualmente, dois pilotos do OPAL estão em desenvolvimento na Colômbia e no Senegal, em parceria com seus respectivos governos e com as operadoras de telecomunicações Orange-Sonatel e Telefônica Colômbia.

Outro desafio é diminuir as desigualdades mundiais no acesso e na habilidade do uso de dados, conforme alertou a ONU em relatório recente. No início da década de 1990, o sociólogo francês Alain Desrosières já chamava atenção para o fato de que o espaço público dependeria cada vez mais de informações estatísticas acessíveis a todos. No entanto, apesar do “dilúvio de dados” em que vivemos, ainda existem lacunas de informação a respeito de certas regiões do mundo e de suas populações, em geral aquelas que já são mais vulnerabilizadas. Essa é uma das principais preocupações de nossos dias, uma vez que a “invisibilidade estatística” tem sido agravada pelo “analfabetismo digital”, configurando o que o pesquisador Emmanuel Letouzé, um dos fundadores da Data-Pop Alliance, definiu como *hiato digital*: a desigualdade de acesso e de capacidade de uso das novas tecnologias, que aprofundam as desigualdades socioeconômicas preexistentes.

Nesse sentido, uma sociedade formada exclusivamente por cientistas de dados não seria necessariamente “alfabetizada em dados”. A Data-Pop Alliance propõe uma definição holística de *data literacy* (leia mais sobre o tema no artigo *Como disseminar a linguagem dos dados*), que pode ser resumida como a vontade e a capacidade de se engajar construtivamente na sociedade por meio dos (ou sobre os) dados. Essa definição considera fundamental que os cidadãos entendam as implicações políticas, econômicas e éticas ligadas ao uso de dados para o usufruto e a garantia de seus direitos. Isso pressupõe, por exemplo, a compreensão de como Google, Apple, Facebook ou

Amazon coletam e analisam dados, o questionamento acerca da existência de consentimento prévio e a exigência de que os dados pessoais sejam protegidos.

PRÓXIMOS PASSOS NO BRASIL

A estruturação de um ecossistema emergente de *big data*, capaz de promover o uso de dados para o bem comum e o empoderamento dos cidadãos, requer uma visão colaborativa e educativa que não se resume ao desenvolvimento de capacidades técnicas e tecnológicas.

Na continuação de sua atuação institucional na América Latina, cujas metas são a promoção do uso de dados e de novas tecnologias em políticas públicas e no desenvolvimento sustentável, a Data Pop-Alliance (cofundada pelo MIT Media Lab, pela Harvard Humanitarian Initiative e pelo Overseas Development Institute) abrirá, em setembro de 2019, um escritório com sede na Cidade do México, a fim de alavancar projetos no Brasil, no Chile, na Colômbia e no México.

Em maio deste ano, a Data-Pop Alliance, a Universidade de Manchester e a FGV EAESP organizaram o evento *Big Data para o Bem Comum*, para discutir as oportunidades do *big data* no contexto brasileiro. Um dos resultados dos grupos de trabalho realizados com representantes do poder público, de empresas privadas e de organizações da sociedade civil é a consolidação de uma proposta para a implementação de um programa de alfabetização de dados no Brasil, no México e na Colômbia, em parceria com a Universidade de Manchester. Outro se refere à consolidação de uma coalizão de dados para o bem comum no Brasil, para a qual gostaríamos de convidar todas e todos os interessados. ●

PARA SABER MAIS:

- Data-Pop Alliance. Disponível em: datapopalliance.org
- OPAL Project. Disponível em: opalproject.org/
- Data-Pop Alliance. *Beyond data literacy: reinventing community engagement and empowerment in the age of data*, 2015. Disponível em: datapopalliance.org/wp-content/uploads/2015/11/Beyond-Data-Literacy-2015.pdf
- Alejandro Noriega-Campero, Alex Rutherford, Oren Lederman, Yves Montjoye, Alex Pentland. *Mapping the Privacy-Utility Tradeoff in Mobile Phone Data for Development*, 2018. Disponível em: arxiv.org/abs/1808.00160

SILVIA RODRIGUES FOLLADOR > Consultora da Data-Pop Alliance > srodriguesfollador@datapopalliance.org

JULIE RICARD > Diretora adjunta da Data-Pop Alliance > jricard@datapopalliance.org