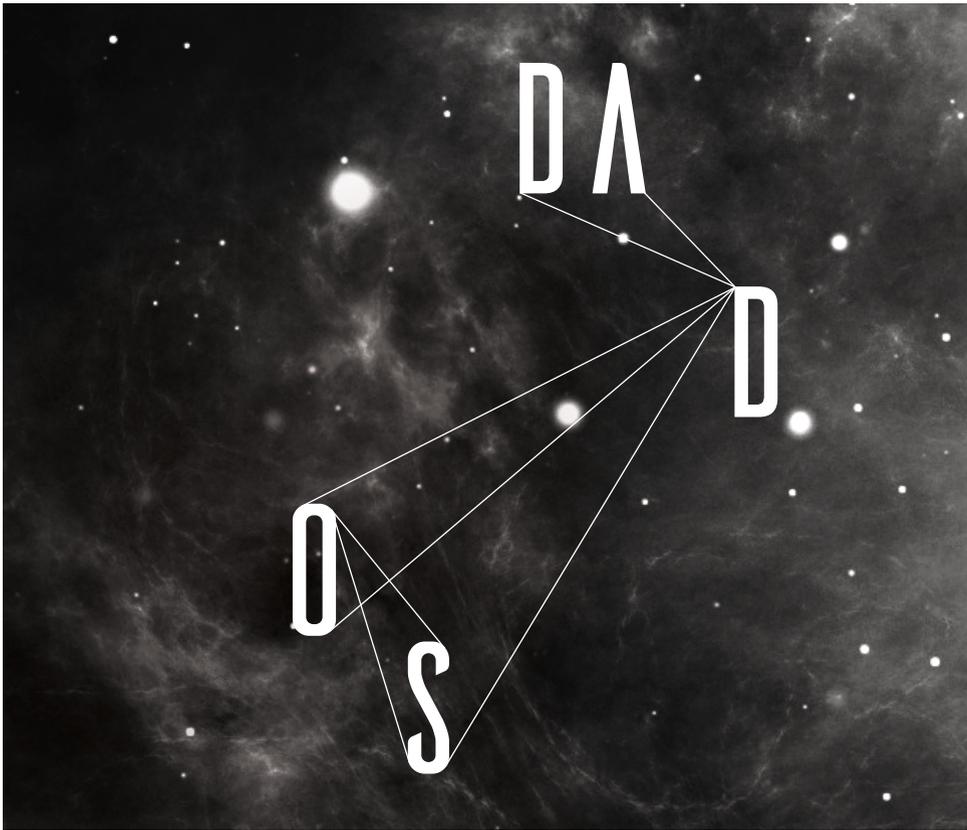


P E N S A R P O R



P E N S A R
P O R
D A D O S¹

*Um banco de dados
jamais abolirá o acaso²*



A L I N E
C O U R I

Você acorda, verifica no seu *smartphone* se possui mensagens SMS ou de WhatsApp. Checa *e-mails*, entra no Facebook e dá uma olhada na previsão do tempo. Em cinco minutos, acessou bancos de dados de servidores nos quais essas mensagens estão gravadas (Google/Gmail; Facebook e WhatsApp) e criou novos dados (mensagens, acessos, postagens). Em uma pesquisa de julho de 2017, 90% de todos os dados no mundo em 2016 foram criados somente nos dois últimos anos. (IBM MARKETING CLOUD, 2017)

Criamos dados quando nos conectamos, mandamos mensagens, escrevemos *e-mails*, fazemos pesquisas, postamos conteúdo, utilizamos aplicativos. Cada busca, clique, curtida, compartilhamento – ou mesmo a informação de quanto tempo o cursor do mouse passa sobre uma imagem – pode ser monetizado. Não fica claro, para nós, especificamente, quais dados estamos disponibilizando para essas empresas. De fato, isso depende do aplicativo e da plataforma utilizada. Ao instalar os aplicativos, clicamos em “aceitar” os contratos de licença (ou termos de uso) sem ler os contratos – que são longos, em letras miúdas e em terminologia jurídica, incompreensíveis para a grande maioria de pessoas. Além disso, muitas vezes, achamos que esses dados que geramos – e fornecemos de graça – são irrelevantes: não temos nada a esconder.

Sites como a Amazon guardam informações de cartões de crédito, endereços, produtos comprados. Usam estes últimos para nos sugerir outros tantos que julgam similares. Na internet, vivemos em bolhas, visualizando aquilo que algoritmos de busca avaliam que nos é relevante. Faça um teste: em dois computadores de pessoas diferentes, digite algo uma busca no Google e veja os resultados. Em cada computador, o resultado será diferente, já que a busca leva em consideração a identidade de quem busca (quem fez o *login*).³ O algoritmo de busca e classificação não é tornado público. Além disso, diariamente, a Google introduz mudanças em seu algoritmo de classificação. Nas palavras da empresa:

Você quer a resposta, não bilhões de páginas web, então os sistemas de classificação da Google ordenam, dentre as centenas de bilhões de páginas no nosso Index de Pesquisa, para te dar resultados *úteis* e *relevantes* em uma fração de segundos. Estes sistemas de classificação são constituídos por uma série de algoritmos que analisam o que você está procurando e quais informações retornar para você. E, à medida que evoluímos, buscamos torná-lo mais útil, refinamos nossos algoritmos para avaliar suas pesquisas e os resultados em detalhes mais finos para tornar nossos serviços funcionando melhor para você. (GOOGLE, 2017, grifo nosso)

Mas como Google escolhe o que é “útil e relevante” para cada um de nós? E como esses algoritmos “evoluem”?

MUNDOS DE DADOS

O produto da televisão. Televisão comercial. É o público. Televisão entrega pessoas a um anunciante [...] Na transmissão comercial o público paga pelo privilégio de ser vendido. É o consumidor que é consumido. [...] Você é consumido. Você é o produto da televisão. Televisão entrega pessoas.⁴ (TELEVISION..., 1973, tradução nossa)

Nos anos 1970, quando a mídia de massa mais ubíqua era a televisão, Richard Serra e Carlota Schoolman chamaram atenção para o fato de

que os comerciais exibidos nas programações das redes de TV, muito mais que proverem informação ao público, entregam sujeitos aos anunciantes, construindo subjetividades e forjando demandas e desejos.

Em 1996, foi publicada a *Declaração da Independência do Ciberespaço*, escrita por John Perry Barlow, que ilustra o sonho *nerd*-libertário de um espaço livre, onde seria possível existir de modo diverso à existência material, livres das mídias de massas, dos “governos do mundo industrial”. Acreditamos que um “vírus da liberdade” estava se espalhando.

Entretanto, atualmente, é cada vez mais difícil ser anônimo, não deixar pegadas na internet. Em cada ação *on-line*, estamos trabalhando de graça, construindo perfis que serão vendidos a anunciantes, pesquisas e empresas de tipos diversos (*marketing*, análises científicas, saúde, bancária, comercial, que falam de “prevenção” de crimes e identificação de comportamentos suspeitos). Decisões feitas com base nos perfis gerados por dados determinam, de maneira duvidosa, quem somos, quais conteúdos veremos, que tipo de produtos nos serão oferecidos, a quais serviços teremos acesso, o que podemos mostrar ou ver, quais territórios podemos cruzar.

De acordo com os dossiês publicados pelo projeto Chupadados, de autoria do coletivo Coding Rights, que se apresenta como “uma organização liderada por mulheres dedicada a promover a compreensão sobre o funcionamento de tecnologias digitais e expor as assimetrias de poder que podem ser ampliadas por seu uso”.

Sensores cada vez mais avançados possibilitam a digitalização massiva dos ambientes em que vivemos, dos nossos corpos e de nossos hábitos. Esse processo faz com que nossos movimentos, gostos de consumo, pensamentos íntimos, sentimentos e relações sejam registrados e transformados em dados, em informações de valor que são exploradas economicamente. (FELIZI; VARON, [20-])

O coletivo analisou, dentre outros, aplicativos de encontros⁵ e de acompanhamento de ciclos menstruais. Assusta saber quantos dados

são fornecidos voluntariamente, como são desprotegidos e facilmente vazados. As empresas desenvolvedoras de aplicativos acabam conseguindo o consentimento, ainda que desinformado, dos dados dos usuários, dando margem a uma série de abusos. Aplicativos de encontros coletam muito mais informações do que deveriam, capitalizando afetos e, devido a falhas, colocando em risco a privacidade e a segurança das pessoas, sem que elas fiquem sabendo. Dentre os dados coletados para exibir ou esconder o que veremos, estão: gostos e afinidades, distância e localização geográfica, amigos em comum, doenças. Usuários de aplicativos de namoro têm tendência a pagar pelo seu uso, o que aumenta as possibilidades de encontros. Isso gera mais uma forma de lucro para essas empresas. Ainda assim, não se sabe exatamente quais são as regras que ditam quais serão as pessoas que vão aparecer para cada participante.

E como Richard Serra já ressaltava, para a grande maioria que prefere usar a versão grátis desses aplicativos, visualizando anúncios dentro destes, o usuário que é o produto (como acontece no Facebook, que é “gratuito”). Os dados sobre o comportamento das pessoas em flerte são fornecidos aos anunciantes, empresas de *marketing*, pesquisa e parceiros comerciais. Isso faz com que sejam exibidas propagandas direcionadas ao perfil de quem usa, a partir da venda de suas informações a empresas dedicadas a comercializar e analisar grandes volumes de dados.

Muitas vezes, os dados recolhidos podem ficar disponíveis mesmo que você saia do aplicativo e cancele sua conta. Guardando dados, fotos e conversas, as empresas seguem lucrando com os perfis gerados pelos dados.

Em setembro de 2016, poucas semanas depois das Olimpíadas, foi revelado que um membro do exército ligado ao governo de São Paulo se infiltrou em um grupo de ativistas de oposição ao governo do presidente Michel Temer. Vinte e uma pessoas, sendo seis adolescentes, foram presas antes de um protesto contra o governo. Sob o nome Baltazar Nunes, o capitão Willian Pina Botelho criou perfis falsos em redes sociais e aplicativos como o Tinder e acabou conhecendo uma pessoa que o introduziu no grupo, levando à prisão dos ativistas.

Em relação aos aplicativos de monitoramento de saúde, a narrativa que predomina é a de que a prática do monitoramento de vários aspectos da vida através de números e ferramentas digitais (o chamado *quantified self*) levam a uma vida melhor. Alimentadas pelos nossos dados, essas ferramentas funcionam como laboratórios para a observação de padrões fisiológicos e comportamentais, que vão desde a frequência da menstruação e de sintomas associados a ela até hábitos de compra e navegação na *web* de todas as suas usuárias.



FIGURA 1: Aplicativo Maya/Lovecycles

Informações que podem ser inseridas: pressão sanguínea, mastalgia, constipação, cólicas, diarreia, tontura. Como se sente: debilitada, dispersiva, emotiva, animada, ansiosa, atrevida, preguiçosa.

FONTE: Aplicativo Maya.⁶

Sob um discurso de “empoderamento” feminino, tais aplicativos mandam diariamente “dicas” que acabam por ditar regras ou valorizar modos de vida, lembrando que a maior parte de usuárias é de jovens mulheres. O aplicativo Glow, por exemplo, exhibe mensagem de que

aos 29 anos, você é parte da esmagadora maioria de mulheres tendo filhos. Um estudo sugere que a melhor idade para ter filhos é entre os fins dos vinte e início dos trinta anos. Pesquisas indicam que esperar até essa idade está associado com mais vantagens sociais e melhor saúde para a mãe e a criança. (Glow app)

Há mais de quatro anos que aplicativos menstruais estão disponíveis nas *app stores*. A mais nova fronteira da “digitalização vaginal” é a combinação entre aplicativo e *hardware*, ou seja, entre o aplicativo e um coletor menstrual inteligente. O Looncup⁷ e o my.Flow⁸ são apresentados como os primeiros coletores menstruais⁹ “inteligentes”, que, além de coletar, avisam pelo celular, por *bluetooth*, quando o copo está cheio, como estão a cor e volume da menstruação. Ainda não estão a venda e não disponibilizaram termos de uso ou políticas de privacidade, ou seja, não é possível ainda verificar as condições de proteção dos dados coletados.

Vibradores e diafragmas conectados e que coletam informações sobre aquelas que os usam e podem colocar em risco a vida das usuárias, em caso de dados vazados, em países onde o adultério, o aborto, a homossexualidade e/ou a posse e uso de *dildos* são criminalizados.

Analisar o aplicativo Glow pode deixar mais claro o lucro que nossos dados podem gerar. Sua política de privacidade diz que a empresa pode decidir compartilhar os dados coletados pelo aplicativo com terceiros. Tais dados podem ser utilizados tanto para serviços dentro do aplicativo, quanto para indicar produtos e serviços externos que possam ser de interesse de quem forneceu os dados. Ou seja, os dados coletados incrementam os negócios das empresas parceiras. Mesmo que você pare se usar o aplicativo e cancele sua conta, seus dados continuam com a empresa, aprimorando o algoritmo, que é vendido como cada vez mais inteligente. Os dados são compartilhados com empresas de serviços de saúde, *fitness* ou pesquisa. A versão paga do Glow (50 dólares por mês) promete sucesso nas tentativas de engravidar num prazo de dez meses. Caso isso não aconteça, a empresa promete pagar o tratamento de fertilidade. O aplicativo Affirm, outro produto da mesma incubadora do Glow, funciona como agência de crédito

virtual. Isto é, depois de conseguir engravidar, tendo recebido toda a publicidade julgada adequada, pede-se um empréstimo através de um aplicativo parceiro.

Mas talvez a maior novidade em termos de produção e disponibilização de dados pessoais esteja por vir da China, que pretende implementar, até 2020, um sistema de atribuição de uma nota a cada cidadão, a partir de dados coletados durante toda sua vida. Em junho de 2014, o Conselho Estadual da China publicou um documento de nome “Planning Outline for the Construction of a Social Credit System”. Uma das ideias contidas no texto é a implementação de uma avaliação, em forma de pontuação, que indica que “tipo” de cidadão cada pessoa é. Quando a maior parte das atividades diárias são constantemente monitoradas e avaliadas, é fácil saber o que cada pessoa compra *on-line*, onde está a cada hora, quem são seus amigos e como interage com eles em redes sociais; quantas horas joga videogame, quais taxas e contas paga ou não paga. Esse Social Credit Score servirá para definir quem é confiável e quem não é, criando um *ranking* que determinará a elegibilidade para um financiamento, um emprego, a escola de seus filhos, permissão para viagens ao exterior, dentre outras vantagens.

Esses dados coletados são armazenados em servidores que existem materialmente em grandes centros de computadores. Tais dados, entretanto, estão replicados e podem ser acessados pelas “nuvens”. A nuvem, em terminologia informacional recente, é usada para se referir a qualquer tipo de informação ou de dados que não esteja armazenado no aparelho da pessoa, mas em um servidor remoto que possibilita o acesso por meio da internet. Esses dados, pelo ponto de vista do usuário, não possuem existência física, como armazenamento em disco rígido, que pode ser quebrado a qualquer momento, ou até mesmo em tradicionais arquivos em papel, que também podem ser danificados. Os dados se transformaram em impulsos elétricos que viajam rápida e constantemente entre localidades distintas e, por vezes, remotas. Essa fluidez permite que sejam acessados de qualquer lugar, a qualquer hora e por qualquer pessoa (desde que autorizada); os dados têm em sua suposta ubiquidade a garantia de serem infalíveis e duradouros. Ainda assim, essa terminologia acusa a sua condição fugaz e frágil.

No meio dessa miríade de dados (conhecida atualmente sob o termo “big data”), aquele que consegue criar informação a partir deles sai ganhando. Essa tarefa é chamada de extração de dados, mineração de dados ou *web scrapping*. Um dado fora do contexto não possui significado relevante e não conduz a nenhuma compreensão. Representa algo que não tem sentido, a princípio. Portanto, é complicado que seja usado para embasar conclusões ou respaldar decisões importantes. E se alguns dados estão presentes, muitos outros nem foram criados. Continuam imersos no ambiente, nos corpos, sem que estejamos atentos a eles.

A profissão de cientistas de dados foi eleita a profissão do século XXI devido ao volume de dados disponível para análise atualmente. Informações abertas na internet, Google Analytics, Customer Relationship Management (CRM), interações por *e-mail* e pesquisas são algumas dessas fontes possíveis. A procura por profissionais que utilizem esses dados a favor dos negócios não para de crescer, já que, dessa forma, as empresas conseguem conhecer melhor o seu cliente, prever tendências de comportamento, encontrar padrões de consumo, aumentar a conversão¹⁰ de mídias digitais e mais.

Hoje, multiplicam-se cursos de formação de “cientistas de dados”, que ensinam técnicas de coleta, tratamento e análise de dados para gerar informação e criar soluções inovadoras para negócios – em geral, utilizando linguagens de programação como Python e Structured Query Language (SQL).

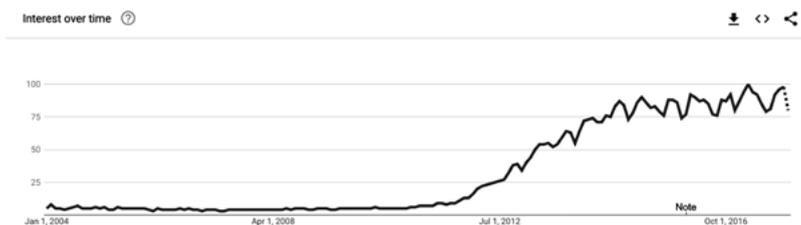


FIGURA 2: Resultado de busca do termo “big data” no Google Trends

Nota-se o crescente interesse desde, principalmente, 2012.

Esses dados extraídos servem também para aprimorar algoritmos e “ensinar” sistemas de inteligência artificial. No Facebook, por exemplo, conteúdo denunciado como impróprio é verificado por pessoas. Essas pessoas, que trabalham para o Facebook e seus usuários, têm acesso a uma linha do tempo “alternativa”, que exhibe apenas as postagens que foram denunciadas, de forma aleatória, junto a um menu sobre possíveis violações. Os moderadores só visualizam o nome do autor das publicações e não têm acesso a seus perfis completos. Sua missão é apagar, ignorar ou encaminhar a publicação para a avaliação superior – o que ocorre especialmente em casos de suicídio ou pedofilia, que, por sua vez, são encaminhados a autoridades. As decisões, pautadas por políticas internas da rede social, servem para “educar” os algoritmos, que, com o tempo, repetem as respostas automaticamente, por meio de recursos avançados de identificação de rostos ou frases ofensivas. Assim, quanto mais os funcionários trabalham, mais ensinam os algoritmos e vão se tornando cada vez menos necessários. Tornam o próprio trabalho obsoleto. Esses funcionários estão expostos a imagens de violência com animais, pessoas, mensagens de ódio. A rotatividade é alta e a maioria não completa um ano no posto.

Chama-se de banco de dados uma coleção organizada de dados, e dado, algo “extraído”, “raspado” de uma pesquisa, condição social ou *site* de internet. Dado é o menor nível de abstração da informação, sendo entendido como um aspecto de um fato em sua forma primária.

Já informação seria a ordenação e organização dos dados de forma a transmitir significado e compreensão dentro de um determinado contexto, um conjunto ou consolidação dos dados de forma a fundamentar conhecimento.

A noção de informação como algo que pode ser armazenado, transferido ou comunicado a um objeto inanimado e a noção de informação como uma quantidade definida matematicamente não surgem antes do século XX. Para Claude Shannon (1948), informação seria uma mensagem enviada por um emissor a um receptor. Ou seja, seria um sinal. Para ele, a informação não tem relação nem com a semântica nem com a pragmática. Weaver, na introdução que escreveu

ao livro de Shannon, diz que “informação não tem nada a ver com significado”, apesar de constituir um padrão. Independia do sentido. Donald MacKay (1951), três anos após Shannon propor sua definição de informação, defendeu outra abordagem para a compreensão da natureza da informação. Sugeriu que informação deve ser definida como mudança mental em um receptor, portanto, com significado, não apenas como sinal do remetente. Propôs um tipo de informação que chamou de “estrutural”: uma mensagem sobre como interpretar uma mensagem. O significado pode ser inventado. Até mesmo os dados são inventados. Não é possível prever os significados criados a partir de um dado: “todo pensamento emite um lance de dados”.¹¹ (MALLARMÉ, 1897) O que consideramos como informação depende do contexto, de onde e como está sendo produzida e usada.

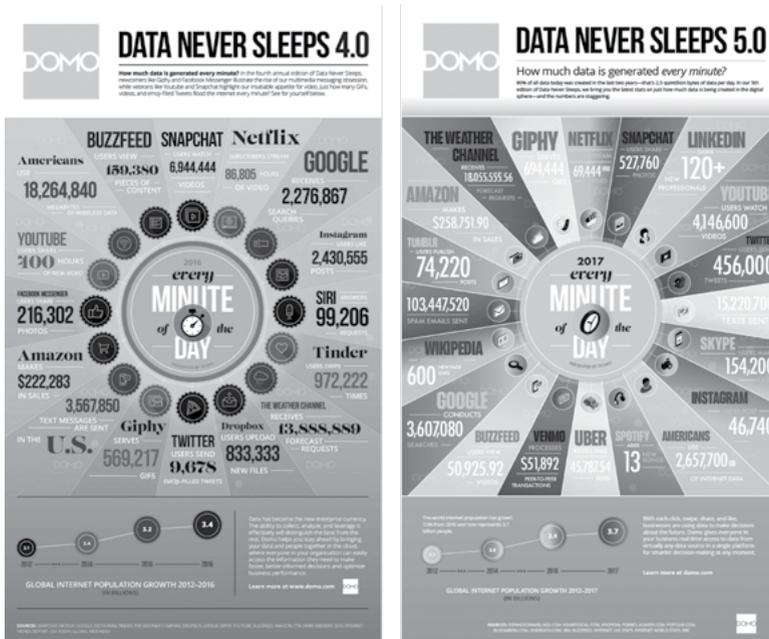


FIGURA 3: Gráficos produzidos pela empresa de *software* DOMO ilustram quantidade de dados produzidos a cada minuto, em 2016 e 2017

Na língua inglesa, a palavra “informação”, de acordo com o *Oxford English Dictionary*, aparece por escrito pela primeira vez em 1386, por Chaucer: “informacions”.¹² A palavra é derivada do latim a partir do francês, combinando a palavra “informar”, que significa “dar uma forma para a mente”, com o sufixo “ção”, que denota um substantivo de ação. Esta primeira noção refere-se a um item de treinamento ou moldagem da mente. A próxima noção de informação, “comunicação do conhecimento”, aparece brevemente em 1450.

QUÃO CONFIÁVEL É UM DADO?

Recentemente, no Brasil, uma foto compartilhada no Facebook serviu como um dado para supostamente provar aspectos repulsivos da arte contemporânea, tratada como algo inquestionável, como parte de um discurso cujo significado estaria evidente já na sua superfície. O compartilhamento dessa foto, a partir de uma interpretação pessoal do que acontecia nesta performance entre o artista e uma criança da plateia, fechou exposições, mudou leis e criou uma multidão de especialistas em arte contemporânea.

A dispersão da imagem digital reforça sua existência como sendo um aglomerado de códigos e influencia sua apresentação visual, considerando os diferentes aparelhos e programas que possibilitam tais visualizações. Hito (2009) apontou essas e outras condições como sendo parte do conceito de “imagem pobre”, declarando que ela “não é mais sobre a coisa real – o original originário. Em vez disso, é sobre suas condições reais de existência: sua circulação abundante, dispersão digital, temporalidades fraturadas e flexíveis”. Esses novos modos de circulação de imagem permitem que elas sejam entendidas não mais como objetos únicos ou cópias, mas, assim como dados em nuvens, objetos (mesmo que não físicos) que existem em múltiplos locais (mesmo que não físicos) simultaneamente, independente de dimensões fixas ou exatidão na representação das cores na tela-monitor. A imagem digital é uma cópia visível do arquivo (dado) invisível. É uma instância, uma atualização de um virtual. O digital cria a ilusão de que não existe mais diferença entre original e cópia, de que tudo que temos são cópias que se multiplicam e circulam em redes

informativas. Mas não podem existir cópias sem um original. A diferença entre original e cópia é extinta no caso da digitalização, pelo fato de que o dado original é invisível: ele existe no espaço invisível atrás da imagem, dentro do computador. Não é possível adentrar o espaço invisível atrás da imagem digital para encontrar o dado ele mesmo. Ao olhar uma imagem digital, somos confrontados a cada vez com um novo evento de visualização de um dado invisível.

Acontece que um dado, por si só, não é informação. Se um dado só existe quando é separado do seu entorno, ao realizar essa separação, já estamos agindo sobre aquilo que será visto como um dado: recortando, escolhendo, justapondo, narrando. Um “dado” pode ser fabricado, viciado, temperado. Esse episódio da foto da performance, de certa forma, se conecta à emergência, após 2016, do fenômeno das *fake news*.

Acredita-se, geralmente, que uma discussão livre e justa é base de toda sociedade que se almeja democrática. Porém, argumentações irracionais e dados falsos, como premissas, podem impossibilitar um debate justo. Do ponto de vista tecnológico e sociológico, é possível pensar que a crescente influência de falsas notícias tenha acontecido, talvez, pela maior possibilidade de criação e disseminação de conteúdos através das mídias sociais. Ou, e também, pela crescente desconfiança nas instituições sociais tradicionais: o sistema político, o sistema das mídias e a ciência – como, por exemplo, o movimento antivacinação.

Pesquisadores tentam mapear e descrever modos pelos quais a desinformação é disseminada e ampliada. Muitos desses estudos dizem respeito às mídias sociais e como as informações falsas circulam através delas.

É importante lembrar que termos como “fake” ou “falso” implicam uma ideia de verdade e algum critério para reconhecê-la. Podemos passar de uma concepção “dura” de fatos (a do realismo externo) ao seu oposto, como algumas posições niilistas segundo as quais não há fatos, mas apenas interpretações.

Nos Estados Unidos, o Facebook admitiu, em setembro, que grupos russos criaram perfis falsos ligados a movimentos sociais, como o Black Lives Matter, para divulgar memes¹³ que incitavam o ódio de grupos

racistas durante as eleições de 2016. Segundo as revelações do Senado americano, notícias falsas sobre muçulmanos também teriam sido impulsionadas por russos na rede para irritar o eleitorado xenófobo.

INTERNET DAS COISAS, *SMART CITIES*

É tão conveniente, tudo se comunica!¹⁴
(Jacques Tati, *Mon Oncle*, 1958)

A irmã do senhor Hulot mora em uma casa moderna, do meio do século XX, projetada para ser prática, estéril e funcional. Apesar das boas intenções, muitas coisas não funcionam tão bem.

O que vimos acima sobre dados e as vidas das pessoas serve para pensarmos os dados e nossos ambientes de vida, as cidades.

A internet das coisas supõe um mundo de objetos conectados, acessíveis pelas redes, que geram dados que servem de *input* para que outras máquinas funcionem. *Smart cities*, as cidades inteligentes, são exaltadas como um benefício para todos, mas às custas de maior vigilância, centralização de poder e controle. Podem existir a partir de elementos domésticos e urbanos conectados. Na Inglaterra, por exemplo, quando um ônibus, ao enviar sua geolocalização à central de controle, tem um atraso diagnosticado, isso serve para que os sinais de trânsito de seu caminho a percorrer tenham um tempo de luz verde mais longo que o normal, permitindo que ele chegue a seu destino no horário determinado.

Segundo o estudo conduzido por Natusch, Felizi e Varon ([201-]), desde a implementação do bilhete único no estado do Rio de Janeiro, os dados dos passageiros de ônibus não estão seguros. Aprovado em 2009 e implementado no ano seguinte, a emissão, comercialização e distribuição do bilhete único está a cargo da Federação das Empresas de Transportes de Passageiros do Rio de Janeiro (Fetranspor). Esta entidade congrega 10 sindicatos, reunindo cerca de 200 empresas de ônibus em todo o estado. É também a proprietária da RioPar Participações S.A., uma *holding* que controla diferentes aspectos

relacionados ao transporte coletivo – e que tem em um dos seus braços a RioCard TI, responsável pelo gerenciamento dos dados coletados pelos cartões de bilhete único.

Ou seja, todos os dados coletados no sistema ficam nas mãos da iniciativa privada, principalmente do empresário Jacob Barata, dono direto ou indireto de cerca de 25% da frota de ônibus da cidade do Rio. Além disso, seu filho é um dos principais sócios da M2M Solutions, que atende a Secretaria Municipal de Transportes do Rio de Janeiro (SMTR) no monitoramento em tempo real, via GPS, das viagens de ônibus pela cidade. A Linktrans, companhia de tecnologia de informação que gera relatórios a partir dos dados de GPS para as empresas de ônibus, também tem na família Barata seus sócios majoritários.¹⁵

Quanto à Riocard, mesmo que pertença a três sócios principais, tem como presidente do Conselho de Administração Jacob Barata Filho, herdeiro do império do pai e sócio ou diretor de várias empresas que operam ônibus no Rio de Janeiro. Cabe lembrar que em julho de 2017 Barata Filho foi preso por envolvimento no esquema de corrupção na área de transporte público no governo de Sérgio Cabral. Desde então, vai e volta da prisão, já que conta com a proteção do ministro Gilmar Mendes, que o liberou três vezes. Este último foi padrinho de casamento da filha do acusado. Só para termos um panorama geral de quem lida com nossos dados.

Seria lógico esperar que o poder público, ainda mais por conceder a exploração do serviço à iniciativa privada, tivesse diretrizes e leis claras sobre a gestão dos dados pessoais (e privados) coletados, mas isso não acontece. Não há termo legal que deixe claro como os dados do transporte dos cidadãos são utilizados e a quem são transferidos, seja no setor público ou privado. Para fiscalizar o cumprimento dos contratos de concessão, as Secretarias de Transportes da cidade do Rio e do Estado do Rio de Janeiro (Setrans) recebem relatórios com os dados coletados no sistema do cartão de transporte.

O governo recebe apenas relatórios usados para a prestação de contas das empresas de ônibus, sendo que a Coordenação e Controladoria Geral do Bilhete Único, órgão da Setrans, têm acesso à base de dados para auditar o sistema contra fraudes, mas não controla o

que a federação faz com os dados, como, por exemplo, se os vende para *databrokers*.¹⁶ Conclusão: o enorme banco de dados sobre os passageiros fica nas mãos de uma entidade privada, controlada por um pequeno número de pessoas nada confiáveis.

Mais grave ainda, está sendo implementado um sistema de biometria. Segundo a Fetranspor, esse sistema garantirá mais segurança e eficiência no controle das gratuidades. Dados ainda mais sensíveis serão coletados através do uso dos bilhetes únicos, sem nenhuma garantia de que funcionarão adequadamente. O público desses transportes – em geral, as camadas menos favorecidas – vira produto, o que pode gerar mais fontes de lucro para esse pequeno número de empresários.

Você sabia que o seu cartão RioCard não pode ser emprestado?

Este ônibus possui câmera que identifica
por foto o titular do cartão cadastrado
(Bilhete Único Intermunicipal ou gratuidade).



Faça sua parte, utilize corretamente seu
cartão e evite o cancelamento do seu benefício.
Não empreste seu RioCard.



www.cartaoiocard.com.br
Para você: 21.2527-4500 Gratuidade: 21.4393-3737
f /riocard.tudoandamehor t @riocardL

RIOCARD
tudo andamehor

FIGURA 4: Cartaz em ônibus do Rio de Janeiro, incriminando o empréstimo do bilhete único e justificando a necessidade da biometria para a segurança do sistema

FONTE: Bazani (2017).

Isso não é alarmismo. Em 2015, o Departamento de Proteção e Defesa do Consumidor (DPDC) notificou a Telefônica/Vivo por vender dados de GPS de seus clientes. Informações sobre a movimentação de pessoas serviriam para o desenvolvimento de estratégias comerciais, de *marketing* e de serviços. Atualmente, não sabemos se os dados de transporte no Rio de Janeiro são utilizados para esse fim. E nem mesmo existe legislação que impeça isso.

Outras estratégias tecnológicas para “ordenamento”, monitorização e administração de cidades são redes de câmeras, balões de vigilância (também com câmeras), *drones* e bloqueadores de rede de telefonia celular.

O *big data* pode gerar formas de discriminação mais intensas e mais precisas, baseadas em dados. Como argumenta Steyerl (2016), atualmente a visão humana vem perdendo importância e está sendo substituída pela filtragem, decifração e reconhecimento de padrões em meio a esse mar de dados.

É pensando sobre esses dados digitais que podemos mais globalmente pensar no significado da palavra “dado”. Cada pessoa é muito mais do que esses dados extraídos, essas partes que, em geral, dão uma ideia muito restrita da complexidade do todo. Não somos a soma dos dados extraídos de nós. Ou você acha que é?

NOTAS

- 1 “Dado”, do latim *datum*, plural *data*, participio passado de *dare*, “dar”, tinha também o significado de “conceder”, “entregar”. Na língua brasileira: adjetivo, algo que se deu, fornecido. Determinante indefinido: “num dado momento”. Substantivo masculino: cada um dos elementos conhecidos de um problema. Informação que se constitui como elemento necessário para uma avaliação. Resultado de pesquisa ou cálculo. Informação relativa a um indivíduo. Informação capaz de ser processada por um sistema informático. Aquilo que foi combinado. O que é habitual. Elemento inicial de qualquer ato de conhecimento antes de ser elaborado no processo cognitivo. E o objeto “dado”, uma peça de múltiplos lados numerados, tinha, em árabe, o nome de *az-zahr*, dando origem ao “azar”, que tem sentido mais próximo de eventualidade, acaso, do que de algo preciso.
- 2 O poema de Mallarmé, “Un coup de dès jamais n’abolira le hasard”, foi publicado em 1897 na revista *Cosmopolis*.
- 3 Para saber mais sobre algoritmos usados pelo Google, ver em: <<https://search-engineland.com/8-major-google-algorithm-updates-explained-282627>>.
- 4 “The Product of Television. Commercial Television. Is the Audience. Television delivers people to an advertiser. ... In commercial broadcasting the viewer pays for the privilege of having himself sold. It is the consumer who is consumed. [...] You are consumed. You are the product of television. Television delivers people.”
- 5 Alguns aplicativos de encontro: Happn, Grindr, Tinder, Ok Cupid e Match.com (os três últimos do grupo Match Group). O lucro desse grupo é de mais de US\$ 2 bilhões ao ano.
- 6 O antigo Love Cycles.
- 7 Ver mais informação em: <<https://www.kickstarter.com/projects/700989404/looncup-the-worlds-first-smart-menstrual-cup>>.
- 8 Para saber mais, ver: <<http://www.trackmyflow.com/mission.html>>.
- 9 Um coletor menstrual é um copinho de silicone hipoalérgênico e antibacteriano que é colocado na entrada da vagina para coletar o sangue da menstruação. Ou seja, é um dispositivo interno, de barreira, que coleta o fluxo ao invés de absorvê-lo.

- 10 No vocabulário do *marketing* digital, conversão é a execução de uma ação que leva uma pessoa a comprar um produto ou serviço. Por exemplo, quando um visitante executa uma ação através de uma chamada em um *site*, quando o visitante “paga” o produto com um compartilhamento em alguma de suas redes sociais, ou um *e-mail* com oferta que faz alguém realizar uma compra no *site*.
- 11 “Toute pensée émet un coup de dés”.
- 12 “Whanne Melibee hadde herd the grete skiles and reasons of Dame Prudence, and hire wise informacions and techynges”. (OXFORD ENGLISH DICTIONARY, 2010)
- 13 A expressão “meme de internet” é usada para descrever um conceito ou uma ideia através de imagem (em movimento ou não) que se espalha via internet. O termo foi derivado a partir do conceito de meme, criado por Richard Dawkins em *The Selfish Gene* (1976) para descrever pequenas unidades de cultura que se espalham de pessoa para pessoa através da cópia ou imitação.
- 14 “C’est si pratique, tout communique!”.
- 15 A influência e poder dos Barata extrapola os limites territoriais do Rio de Janeiro, chegando ao Ceará e Portugal.
- 16 Empresas que compilam e vendem dados.

REFERÊNCIAS

- BARLOW, J. P. *A Declaration of the Independence of Cyberspace*. Davos, 1996. Disponível em: <<https://www.eff.org/pt-br/cyberspace-independence>>. Acesso em: 15 out. 2017.
- BAZANI, A. Ônibus do Rio bloqueiam 34 mil cartões de gratuidades por uso indevido e RioCard lança campanha. *Diário do transporte*, São Paulo, 22 fev. 2017. Disponível em: <<https://diariodotransporte.com.br/2017/02/22/onibus-do-rio-bloqueiam-34-mil-cartoes-de-gratuidades-por-uso-indevido-e-riocard-lanca-campanha/>>. Acesso em: 13 jan. 2018.
- BOTSMAN, R. Big data meets Big Brother as China moves to rate its citizens. *Wired*, [S.l.], 21 Oct. 2017. Disponível em: <<http://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion>>. Acesso em: 15 nov. 2017.
- DAWKINS, R. *The Selfish Gene*. Oxford: Oxford University Press, 1976.
- FELIZI, N.; VARON, J. *O que é o Chupadados e por onde anda essa entidade?* [20-]. Disponível em: <<https://chupadados.codingrights.org/introducao/>>. Acesso em: 1 jan. 2018.
- GOOGLE. *Como funcionam os algoritmos da pesquisa*. 2017. Disponível em: <<https://www.google.com/search/howsearchworks/algorithms/>>. Acesso em: 15 nov. 2017.
- GOOGLE TRENDS. Disponível em: <<https://trends.google.com/trends/explore?date=all&q=big%20data>>. Acesso em: 1 jan. 2018.
- GROYS, B. *From Image to Image File - and Back: Art in the Age of Digitalization*. Cambridge: MIT Press, 2008.
- IBM MARKETING CLOUD. *10 Key Marketing Trends for 2017 and Ideas for Exceeding Customer Expectations*. 2017. Disponível em: <<https://public.dhe.ibm.com/common/ssi/ecm/wr/en/wrl12345usen/watson-customer-engagement-watson-marketing-wr-other-papers-and-reports-wrl12345usen-20170719.pdf>>. Acesso em: 15 nov. 2017.

JAMES, J. *Data Never Sleeps 4.0*. 2017a. Disponível em: <<https://www.domo.com/blog/data-never-sleeps-4-0/>>. Acesso em: 1 jan. 2018.

JAMES, J. *Data Never Sleeps 5.0*. 2017b. Disponível em: <<https://www.domo.com/blog/data-never-sleeps-5-0/>>. Acesso em: 1 jan. 2018.

LOGAN, R. K. *Que é informação?* a propagação da organização na biosfera, na simbolosfera, na tecnosfera e na econosfera. Rio de Janeiro: Contraponto: PUC-Rio, 2012.

MANOVICH, L. Database as a symbolic form. *Millennium Film Journal*, [S.l.], n. 34, 1999.

MACKAY, D. M. The nomenclature of information Theory with Postscript on Structural Information-Content and Optical Resolution. In: MACKAY, D. M. *Information, Mechanism and Meaning*. Cambridge: MIT Press, 1951. p. 165-189.

MALLARMÉ, S. Un coup de dés jamais n'abolira le hasard. *Cosmópolis*, Paris, t. 6, n. 17, p. 417-427, 1897.

MON ONCLE. Direção de Jacques Tati. Intérpretes: Yvonne Arnaud; Jean-François Martial; Adelaide Danieli. Paris: Alter Films, 1958. 1 bobina cinematográfica (117 min), son., color., 35 mm.

NATUSCH, I.; FELIZI, N.; VARON, J. *Bilhete único: concentração de dados e dinheiro no transporte público do Rio*. [201-]. Disponível em: <<https://chupadados.codingrights.org/com-o-riocard-seus-dados-passeiam-pelo-rj-e-ninguem-sabe-onde-vao-descer/>>. Acesso em: 1 jan. 2018.

OXFORD ENGLISH DICTIONARY. Oxford: Oxford University Press, 2010.

TELEVISION Delivers People. Produção de Richard Serra e Carlota Fay Schoolman,. [S.l.: s.n.], 1973.1 videocassette (6 min.), son., color.

SHANNON, C. A Mathematical Theory of Communication. *Bell System Technical Journal*, New York, v. 27, p. 379-423, 623-656, July/Oct. 1948.

STEYERL, H. A Sea of Data: Apophenia and Pattern (Mis-)Recognition. *E-flux*, [S.l.], n. 72, Apr. 2016. Disponível em: <<http://www.e-flux.com/journal/72/60480/a-sea-of-data-apophenia-and-pattern-mis-recognition/>>. Acesso em: 15 nov. 2017.

STEYERL, H. In defense of the poor image. *E-flux*, [S.l.], n. 10, Nov. 2009. Disponível em: <<http://www.e-flux.com/journal/in-defense-of-the-poor-image/>>. Acesso em: 15 nov. 2017.