

# Economic Analysis of Law Review

## Big Data e Concorrência: Quando Big Data é Uma Variável Competitiva em Mercados Digitais e Deve Ser Considerada na Análise Concorrencial?

*Big Data and Competition: When Big Data is a Competitive Variable in Digital Markets and Should Be Considered in Competitive Analysis?*

Carlos Ragazzo<sup>1</sup>  
FGV Direito RIO

Gabriela Monteiro<sup>2</sup>  
FGV Direito RIO

### RESUMO

As etapas da análise de atos de concentração giram em torno da noção de poder mercado e da probabilidade de seu exercício. Nesse aspecto, são analisadas diversas variáveis, como a existência de vantagem exclusiva detida por um agente incumbente. Neste ponto, surge a discussão sobre se a exploração de big data pode representar uma vantagem competitiva para o agente econômico que a detém. Não há uma resposta única para o debate, sendo necessário avaliar uma série de fatores, no caso concreto, sob a ótica da cadeia de valor do big data, como: (i) a presença de efeitos de rede (user and monetisation feedback loops), (ii) a existência de economias de escala e escopo; (iii) a disponibilidade e a acessibilidade a dados e às tecnologias para a sua exploração; e (iv) a variação do valor dos dados coletados ao longo do tempo.

**Palavras-chave:** *Big data*. Dados. Concorrência. Antitruste. Vantagem competitiva. Variável competitiva.

**JEL:** K1, K15

### ABSTRACT

The stages of the analysis of concentration acts revolve around the notion of market power and the probability of its exercise. In this aspect, several variables are analyzed, such as the existence of exclusive advantage held by an incumbent agent. At this point, there is an emerging debate on whether the exploitation of big data can represent a competitive advantage for the economic agent that holds this capacity. There is no single answer to this discussion, and it is necessary to evaluate several factors, in the concrete case, from the perspective of the big data value chain, such as: (i) the presence of user and monetization feedback loops, (ii) the existence of economies of scale and scope; (iii) the availability and accessibility to data and technologies used for its exploitation; and (iv) the variation in the value of the data collected over time.

**Keywords:** Big data. Data. Competition. Antitrust. Competitive advantage. Competitive variable.

**R:** 08/05/18 **A:** 01/10/18 **P:** 31/12/18

<sup>1</sup> E-mail: carlos.ragazzo@fgv.br

<sup>2</sup> E-mail: gabriela.monteiro@veirano.com.br

## 1. Introdução

Na análise de atos de concentração econômica, as autoridades de defesa da concorrência analisam os efeitos líquidos das operações que lhes são submetidas sobre o bem-estar econômico, para ao final decidirem se elas devem ser aprovadas, reprovadas ou aprovadas com determinadas restrições. Nessa análise, é tradicionalmente aplicado um método a partir dos fundamentos jurídicos e econômicos e dos objetivos que embasam a política de defesa da concorrência, para os quais as mais variadas teorias econômicas já contribuíram.

Esse método tradicional de controle de atos de concentração é composto, em geral, pelas seguintes etapas: (i) definição do mercado relevante; (ii) análise dos *market shares* (*proxy* utilizada para a identificação de presunção de poder de mercado) e do nível de concentração dos mercados relevantes afetados; (iii) avaliação, a partir das condições do mercado, da probabilidade do exercício do poder de mercado decorrente da concentração gerada pela operação; (iv) avaliação do poder de compra existente no mercado afetado pela operação ou criado, no caso de mercados de insumos; (v) avaliação de eficiências econômicas geradas pela operação; (vi) análise dos efeitos líquidos da operação; e (vii) verificação da possibilidade de imposição de remédios<sup>3</sup>.

Como se pode notar, as análises feitas ao longo das etapas desse método giram em torno da noção de poder mercado (*i.e.*, habilidade para definir preços acima do nível competitivo), pressupondo-se que danos ou restrições à concorrência em um mercado delimitado somente podem ser causados pelos agentes econômicos que detêm esse poder. Todavia, a simples demonstração da existência de poder econômico não é suficiente para se presumir que a operação gerará efeitos anticompetitivos (unilaterais ou coordenados) e implicará perda de bem-estar. Desse modo, a análise dos efeitos anticompetitivos potencialmente decorrentes da concentração econômica prossegue com uma avaliação da probabilidade de exercício do poder de mercado adquirido ou reforçado (*i.e.*, se as empresas concentradas serão capazes de elevar o preço acima do nível competitivo lucrativamente).

Particularmente na análise de poder de mercado, após a mensuração do nível de concentração dos mercados relevantes e verificação do atingimento dos patamares a partir dos quais a intervenção antitruste se torna necessária, também é feita uma avaliação da probabilidade do exercício desse poder. Isso porque, embora possa funcionar como uma *proxy*, a avaliação de *market shares* pode não refletir o poder de mercado dos agentes econômicos envolvidos na operação (SULLIVAN; HARRISON; 2014, p. 28), de modo que é necessário, ainda, que as autoridades de defesa da concorrência analisem a estrutura do mercado relevante para avaliar se existem fatores que podem impedir ou limitar o seu exercício. Dessa forma, também são avaliados fatores como, por exemplo, a existência de barreiras à entrada de novos concorrentes no mercado relevante, a significância dos rivais que já atuam nesse mercado, a pressão realizada por importações, a existência de poder de portfólio e, ainda, o poder de barganha dos compradores.

Nessa etapa, são analisadas diversas variáveis que podem indicar se o exercício de poder de mercado será provável, ou não, no cenário após a concentração. Particularmente para a análise das condições de entrada, por exemplo, o Gui- H do CADE elenca uma série de fatores que podem consistir em barreiras, como por exemplo: (i) a existência custos irrecuperáveis (*sunk costs*), que não

---

<sup>3</sup> Em 2016, o Conselho Administrativo de Defesa Econômica (CADE) lançou um novo Guia de Análise de Atos de Concentração Horizontais (Guia-H) para orientar a análise realizada pelo órgão nas etapas acima, de acordo com as melhores práticas adotadas em importantes jurisdições concorrenciais, como a norte-americana e a europeia. Conquanto a metodologia prevista no Guia-H não possua um caráter obrigatório ou vinculante para o CADE, nem seja exauriente no que diz respeito a todos os possíveis métodos de análise de casos concretos, a publicação do referido documento tem os objetivos de guiar os seus membros a empregar as melhores práticas e auxiliar os agentes econômicos a compreender cada uma de suas etapas, técnicas e critérios.

podem ser reembolsados pelo agente econômico produtor quando este decide sair do mercado relevante; (ii) a existência de economias de escala e/ou economias de escopo; (iii) a presença de exigências legais ou regulatórias, que estabelecem requisitos mínimos para a instalação e funcionamento da uma empresa, condicionando os investimentos dos potenciais entrantes em termos de tempo, espaço e tecnologia; (iv) o grau de integração da cadeia produtiva; (v) a fidelidade de consumidores a marcas já estabelecidas; e (vi) a ameaça de reação de agentes econômicos já instalados (CADE, 2016, p. 27).

A existência de vantagem exclusiva detida por um *player* que já esteja instalado no mercado também é um fator que pode consistir em barreira à entrada e, por isso, deve analisado pela autoridade de concorrencial. O Guia-H do CADE define vantagem exclusiva como:

a propriedade, o acesso exclusivo ou o acesso condicionado a qualquer insumo de produção com oferta restrita ou inelástica a preços pré-concentração (por exemplo, acesso a fornecedores e distribuidores, contratos com ofertantes de serviços e bens de capital especializados, localização, controle ou propriedade de patentes, além de conhecimentos tácitos e intangíveis de tecnologias) (CADE, 2016, p. 28).

É exatamente esse aspecto que surge a discussão entre *big data* e concorrência. Mais recentemente, com o aumento no número de operações em setores que exploram dados de forma maciça, particularmente o digital<sup>4</sup>, e uma série de casos emblemáticos, a doutrina especializada vem debatendo se a exploração da capacidade de *big data* pode efetivamente representar uma verdadeira vantagem competitiva para o agente econômico que a detém.

De um lado, o principal argumento posto por parte da doutrina é o de que a capacidade de *big data* pode representar uma significativa barreira à entrada que levaria à manutenção no mercado de poucas grandes empresas incumbentes com poder de mercado, particularmente nos setores *online*. De outro lado, contudo, se argumenta que a exploração de *big data* tem o condão de aperfeiçoar os produtos ou serviços ofertados e reforçar o processo competitivo, sendo que, no caso específico dos serviços *online*, haveria baixas barreiras à entrada e, inclusive, a desnecessidade de explorar essa capacidade para adentrar no mercado (TUCKER; WELLFORD, 2014, p. 1).

Dito isto, com base na literatura especializada, o presente artigo tem o propósito de avaliar se e em quais circunstâncias a capacidade de *big data* pode ser considerada uma vantagem competitiva para as empresas que a exploram em mercados digitais, exigindo a sua inserção nas análises concorrenciais<sup>5</sup>, particularmente no que concerne à avaliação de barreiras à entrada e rivalidade. Para tanto, o presente trabalho foi dividido em 6 seções, incluindo esta introdução. Na seção 2 é apresentado, em linhas gerais, o conceito de *big data*, destacando-se os atributos dessa capacidade e a cadeia de valor de sua exploração. Na seção 3, são apresentadas as características econômicas que são frequentemente atribuídas a dados e ao ecossistema de plataformas que

---

<sup>4</sup> Como o valor que pode ser extraído da exploração dessa capacidade de *big data* depende, em larga escala, do volume, da qualidade (variedade) e da velocidade com que os dados são coletados e processados, a aquisição de agentes econômicos que já detenham amplos conjuntos de dados à sua disposição tem se tornado uma opção crescente para as empresas que atuam no setor digital (KAROLCZYK; LE CROY, 2018, p. 2). Essa tendência é confirmada pelo aumento expressivo no número de aquisições e fusões impulsionadas por dados envolvendo empresas que exploram *big data* (as chamadas *data-driven mergers*), muitas delas com um claro e evidente propósito de aquisição de dados (OCDE, 2016a, p. 5).

<sup>5</sup> Essa escolha não foi trivial, mas em razão de esses agentes econômicos serem aqueles que mais exploram e utilizam em seus modelos de negócios tecnologias de processamento de dados, sobretudo pessoais (MAYER-SCHÖNBERGER; CUKIER, 2014, p. 6). Além disso, a análise pretendida será feita sob o prisma do controle de estruturas. Essa limitação de escopo foi feita em função do aumento do número de operações realizadas em mercados envolvendo a exploração de *big data*, muitas delas com um claro e evidente propósito de aquisição de dados (OCDE, 2016a, p. 5), e pelo fato de que, no âmbito do controle de estruturas, a matéria ainda está em estágio bastante inicial.

exploram *big data*, em torno das quais giram os debates acerca da significância competitiva dessa capacidade. Já na seção 4, são apresentadas considerações sobre a inserção da variável competitiva *big data* na análise concorrencial. Por fim, as seções 5 e 6, respectivamente, contêm as conclusões e referências bibliográficas deste trabalho.

## **2. Big Data: Sobre O Que Estamos Falando?**

Antes de analisar a questão da vantagem competitiva proporcionada pela sua exploração, é necessário compreender, em linhas gerais, o significado de *big data*.<sup>6</sup>

Segundo Hu et al. (2014, p. 654/655), dada a sua popularidade, a definição do termo *big data* é bastante diversa, de modo que atingir um consenso é uma atividade difícil. Fundamentalmente, a expressão *big data* capturaria um volume de dados que não se qualificaria apenas por ser amplo, mas também por conter outras características relevantes. Apesar da enorme variedade de definições existentes na literatura para *big data*, os autores reconhecem que 03 (três) formas distintas de definição desempenham um importante papel na modelagem desse conceito: (i) a definição atributiva (*i.e.*, baseada nos seus principais atributos), (ii) a definição comparativa (*i.e.*, que define *big data* em termos métricos, ainda que não estanques), e (iii) a definição arquitetural (*i.e.*, com enfoque na estrutura do *big data*, que demanda novas tecnologias de processamento). Definições atributivas são frequentemente encontradas na literatura para o termo *big data*, especialmente a partir do trabalho de Doug Laney, que cunhou os seus famosos “3 Vs” dimensionais (volume, variedade e velocidade). Muitas vezes, contudo, o que se verifica não é uma proposta de definição exclusivamente atributiva, comparativa ou arquitetural, mas um misto de todas elas.

Assim, em linhas gerais, o termo se refere a um grande volume de dados digitais variados, que são coletados, armazenados e processados analiticamente em alta velocidade, particularmente com o objetivo de identificar padrões e comportamentos e fazer correlações, extraindo-se, dessa forma, informações relevantes que lhes conferem valor e podem ser utilizadas em processos de decisão informada nas mais variadas áreas. Desse modo, mais do que um conjunto volumoso de dados variados, *big data* consiste na capacidade de processar analiticamente esses dados em alta velocidade (ou, muitas vezes, quase em tempo real), de modo que a tecnologia analítica importa tanto quanto o volume de dados variados para a extração de informação valiosa.<sup>7</sup>

---

<sup>6</sup> A partir da análise da literatura, é possível perceber que muitos trabalhos não chegam a estudar profundamente o que é *big data*, nem seus atributos e cadeia de valor, de modo que acabam tratando da questão simplesmente com a aplicação de testes de disponibilidade e acesso a dados, desconsiderando diversas outras barreiras que podem surgir na exploração dessa capacidade. No entanto, como será exposto, dados por si só têm um valor muito reduzido ou nenhum, se não forem devidamente processados e analisados. Eles são um insumo para o *big data*, mas este não se resume somente à coleta de dados. O valor da exploração da capacidade de *big data* é extraído das informações geradas pelo processamento e análise em alta velocidade de um grande volume de dados variados, o que também demanda tecnologia e acontece ao longo de toda uma cadeia de valor, que, por esse motivo, deve ser considerada nas análises concorrenciais (RUBINFELD; GAL, 2017, p. 33).

<sup>7</sup> Mais recentemente, vem se notando esforços doutrinários para conceituar o termo *big data*, especialmente do ponto de vista do direito da concorrência. Nesse sentido, por exemplo, Boutin e Clemens (2017, p. 4) propuseram uma definição específica para esse termo que pudesse efetivamente expressar a sua relevância no âmbito do direito da concorrência, considerando os intensos e ambíguos debates que vêm sendo travados nessa seara sobre o tema. Para tanto, os autores analisaram o funcionamento de mercados baseados em dados e os casos já analisados por autoridades de defesa da concorrência nesses setores, bem como estudaram a literatura existente sobre *big data*, particularmente na área de TI. Tomando em consideração a natureza dinâmica dos mercados baseados em dados e os atributos usualmente conferidos ao *big data* (volume, variedade e velocidade), Boutin e Clemens (2017, p. 4, EALR, V.9, n° 3, p. 150-177, Set-Dez, 2018

### a. Os Atributos do *Big Data*

Como visto, o volume de dados não é a única característica importante do *big data* (OCDE, 2013a, p. 11). Em fevereiro de 2001, Doug Laney publicou pesquisa intitulada “*3D Data Management: Controlling Data Volume, Velocity, and Variety*”, a partir do qual começou-se a falar nos chamados “3Vs” do *big data*. Conforme explica Gill Press, uma década depois, esses “3Vs” tornaram-se qualidades bastante aceitas na literatura para a delimitação do *big data*, conquanto este termo não tenha propriamente aparecido na pesquisa de Doug Laney (PRESS, 2013).

O elevado volume de dados coletados, armazenados e processados analiticamente é uma das principais características do *big data*, que não é expresso em nenhuma medida específica pelo fato de ser continuamente crescente e, por natureza, uma *moving-definition*. A velocidade do *big data*, por sua vez, está relacionada à elevada rapidez com que os dados são gerados, acessados, captados, inseridos nos sistemas de análise, processados e atualizados. Por fim, a variedade do *big data* é uma característica relacionada à sua heterogeneidade e complexidade. Isto é, os dados coletados são originados a partir de variados processos de digitalização, possuindo diversas fontes, formas de armazenamento e tipos de informação, mas podem ser combinados e sintetizados (compreendidos) por ferramentas de processamento analítico.

Cumpram ressaltar que, para além dessas propriedades técnicas, outras qualidades foram posteriormente acrescentadas pela literatura na sua delimitação<sup>8</sup>, destacando-se entre elas o Valor. Esse atributo “está relacionado ao crescente valor socioeconômico a ser obtido do uso de *big data*” (OCDE, 2013a, p. 12, tradução nossa). Ou seja, essa propriedade técnica do *big data* se refere aos benefícios sociais e econômicos extraídos a partir da sua exploração, sendo ao mesmo tempo “uma causa e uma consequência do aumento no volume, variedade e velocidade” dos dados (OCDE, 2016a, p. 6, tradução nossa).

### b. A Cadeia de Valor do *Big Data*

O processo de transformação de dados em *big data* é representado por uma cadeia de valor. Assim, a cadeia de valor do *big data* é o caminho percorrido pelos dados “brutos” digitais até que se possa extrair valor econômico deles. O resultado final é um ativo informacional, que pode ser utilizado em variados processos de tomada de decisão informada. A partir do advento da “Internet das Coisas”, cada vez mais, os dados de indivíduos poderão ser extraídos e inferidos a partir de suas atividades e hábitos cotidianos, sendo baixa a possibilidade de que haja erros nas análises desses dados para a retirada de valor comercial dos mesmos. A fusão e a exploração desses dados pessoais com outros tipos de dados (condições do tempo, dados industriais, eventos públicos, etc.) permitem novas correlações e a extração de novas informações que possibilitam os agentes econômicos, por exemplo, a entenderem melhor ou explorarem um determinado mercado. Dessa forma, enquanto dados “crus” podem ser gratuitos ou de baixo valor (dependendo da forma como

---

tradução nossa) sugerem que o termo não seja simplesmente visto como um ativo, mas como uma capacidade tecnológica que pode transformar de forma significativa qualquer mercado, consistente na “habilidade de coletar e analisar um grande volume de dados que contêm uma variedade de informações de forma tempestiva”.

<sup>8</sup> A literatura tem apontado ainda os atributos da veracidade, validação e variabilidade, conquanto – cabe observar – ainda não haja um consenso com relação a eles. A veracidade se refere à confiabilidade e qualidade dos dados, de que depende a precisão das atividades de processamento e de análise (JAPKOWICZ; STEFANOWSKI, 2016, p. 3). Trata-se de atributo relevante para garantir a veracidade da análise e a revelação de informações precisas. Alguns autores se referem ainda ao atributo da validação, que corresponderia à capacidade de garantir que as múltiplas fontes de dados façam sentido quando sintetizadas e de analisar informações sensíveis sob padrões de segurança e governança (BAGNOLLI, 2016, p. 9). Finalmente, a variabilidade do *big data* se refere às mudanças constantes nos seus significados em função de inconsistências ao longo do tempo e de mudanças conforme as interpretações que lhes são atribuídas (JAPKOWICZ; STEFANOWSKI, 2016, p. 3), o que é particularmente relevante quando a exploração dos dados se baseia em processamento de linguagens.

são coletados), o processo pelo qual a informação é extraída da fusão desses dados gera valor no mercado (OCDE, 2016a, p. 6).

Essa cadeia de valor do *big data* possui 05 (cinco) elos principais interligados e complexos, conforme ilustrado na Figura 1 abaixo. Em cada elo dessa cadeia é adicionado valor aos dados brutos digitais que são capturados.



**FIGURA 1:** Cadeia de valor do *big data*  
Elaboração própria

O crescimento do volume de dados gerados é, em especial, uma consequência da migração de muitas atividades econômicas, sociais e públicas para a *internet*, bem como da implantação e interconexão de sensores fixos e móveis, da ubiquidade das atividades *online* e do desenvolvimento de dispositivos eletrônicos cada vez mais poderosos<sup>9</sup>, menores, inteligentes e baratos e que estão mais facilmente acessíveis, assim como a própria *internet*. Com os avanços tecnológicos, todos esses fatores levaram a uma queda nos custos de coleta, armazenamento, processamento e análise de dados (OCDE, 2016a, p. 6/8), impulsionando o surgimento de uma economia digital e permitindo o advento de modelos de negócios baseados na exploração de *big data* (OCDE, 2016a, p. 5).

Nesse cenário, o segundo elo da cadeia de valor do *big data* se refere à aquisição de dados em alta velocidade, que são capturados principalmente de forma direta de indivíduos, entidades, sensores, etc. (*online* ou *off-line*). No entanto, os dados também podem ser indiretamente adquiridos no mercado por meio de terceiros intermediários, como os chamados *data brokers*, que são empresas especializadas na coleta de dados e no desenvolvimento de perfis de consumidores (GERADIN; KUSCHESKY, 2013, p. 1).<sup>10</sup> Com essa coleta de dados “brutos” digitalizados já há um ganho de valor, uma vez que, além de coletados, os dados também são filtrados e “limpos” antes de serem armazenados e processados.

No terceiro elo da cadeia, os dados coletados são armazenados em servidores internos ou, externamente, em serviços de computação em nuvem (*cloud computing*). Quanto maior a quantidade de dados a serem armazenados, maiores precisam ser os *data centres*, o que exige, por partes das empresas, maior infraestrutura e investimento em recursos de computação, armazenamento e rede (RIFFE, 2016).

Apesar disso, o uso de dados tem sido facilitado, em grande medida, pelo declínio nos custos de seu armazenamento (OCDE, 2013a, p. 9).<sup>11</sup> e pelo desenvolvimento de diversas tecnologias novas que possibilitam que as empresas mudem suas estratégias. Conforme observado

<sup>9</sup> A OCDE (2013a, p. 10) reporta que a Lei de Moore, que previa que o poder (capacidade) de processamento de computadores dobraria a cada 18 (dezoito) meses, se realizou em grande medida.

<sup>10</sup> Nos Estados Unidos, a *Federal Trade Commission* (FTC) define *data brokers* como empresas que têm como principal negócio “coletar informações pessoais sobre consumidores de uma variedade de fontes e agregar, analisar, e compartilhar essas informações, ou informações derivadas daquela, para fins como comercializar produtos, verificar a identidade de um indivíduo e detectar fraudes” (EUA, 2014a, p. 3, tradução nossa), por exemplo.

<sup>11</sup> Com relação a esse aspecto, a OCDE (2013a, p. 9) também reporta que o custo médio por *gigabyte* de discos rígido era de USD 56,00 (cinquenta e seis dólares) em 1998, passando para USD 0,05 (cinco centavos de dólar) em 2012.

pela OCDE (2016a, p. 14), as infraestruturas de TI que suportam as operações realizadas por empresas na exploração de *big data* têm sido, muitas vezes, fornecidas por terceiros prestadores desses serviços, que não só fornecem *softwares* para processamento de dados, mas também serviços de computação em nuvem e de armazenamento, funcionando como verdadeiros *data centres* externos onde empresas podem armazenar e processar seus dados por demanda, convertendo seus custos fixos em custo variáveis (CRE, 2017, p. 13), o que é particularmente relevante já que essas infraestruturas são robustas e usualmente exibem economias de escala (OCDE, 2016a, p. 14; CRE, 2017, p. 13).

No quarto elo da cadeia de valor do *big data*, os dados coletados e armazenados são processados e analisados em alta velocidade com o propósito de se extrair informações relevantes, especialmente padrões e correlações. Esse processamento significa “a exploração, transformação, e modelagem dos dados com o objetivo de destacar dados relevantes, sintetizando e extraindo informações úteis escondidas com alto potencial do ponto de vista dos negócios” (CURRY, 2016, p. 30, tradução nossa). Conforme observam Rubinfeld e Gal (2017, p. 2), os dados, por si só, têm pouco valor, sendo o seu processamento que lhes confere este atributo ao transformar um enorme volume de dados não-estruturados em informações e informações derivadas (*i.e.*, aquelas que não são geradas diretamente a partir dos dados, mas que deles podem ser inferidas), sejam elas descritivas ou preditivas.

Essa análise é feita por meio de diversos *softwares* e algoritmos, de modo que o fenômeno do *big data* é intrinsecamente associado ao de *data analytics*, que é, “a capacidade de processar e analisar dados não-estruturados diversos em (quase) tempo real” (OCDE, 2014, p. 11, tradução nossa). Assim – convém reiterar –, dados não são o único insumo para que um agente econômico consiga alcançar a capacidade tecnológica de *big data*, tendo as tecnologias de processamento um papel igualmente relevante.

O desenvolvimento *in-house* dessas tecnologias depende de mão de obra bastante qualificada, investimento e, ainda, da própria capacidade “autodidata” desses programas de aprender com os próprios dados que os alimentam (*self-learning*). No entanto, conforme reportado pela OCDE (2013a, p. 10), as ferramentas de processamento de dados têm se tornado cada vez mais poderosas, sofisticadas, onipresentes e baratas, o que facilita a pesquisa, o rastreamento e a correlação de dados também por diversos agentes econômicos dos mais variados setores da econômica. Esses programas também podem ser obtidos com terceiros, o que pode ajudar a converter custos fixos em custos variáveis (CRE, 2017, p. 14). Até mesmo as grandes plataformas, como a Amazon, o Google e a Microsoft fornecem serviços de *machine learning algorithms* como parte de seus serviços de computação em nuvem, de modo que muitas empresas acham, cada vez mais, conveniente ter os seus dados processados e explorados por meio de infraestruturas de TI externas (OCDE, 2016a, p. 14). O fenômeno da computação em nuvem tem tido um papel importante no aumento da capacidade de armazenamento e processamento de dados (OCDE, 2013a, p. 11) ao permitir que organizações acessem a nuvem de forma flexível, elástica e por demanda (modelo “*pay-as-you-go*”), sem grandes investimentos em ativos de informática. Outro fator que tem contribuído para o barateamento do processamento de dados é o desenvolvimento de *softwares* abertos (*open source*) que cobrem uma ampla variedade de funções necessárias para o *big data*. Além disso, como já exposto, cada vez mais *data brokers* e empresas especializadas em *data analytics* também têm oferecido, no mercado, dados processados para uso em atividades econômicas.

Finalmente, no quinto elo da cadeia de valor do *big data*, as informações relevantes extraídas do processamento e da análise de dados são utilizadas para basear decisões informadas. Essas informações podem ter diversas aplicações, como a melhoria de processos produtivos e da organização industrial, possibilitando a geração de eficiências produtivas e dinâmicas, o aperfeiçoamento de produtos em função de uma melhor compreensão acerca das necessidades dos

consumidores, a personalização de estratégias de *marketing*, o desenvolvimento de produtos ou serviços baseados em dados (CRE, 2017, p. 14), entre outras variadas aplicações. Conforme reportado pela OCDE (2013a, p. 12), as organizações fazem um uso, cada vez maior, de fluxos de dados em suas operações, sendo o aumento no volume de investimentos em atividade de gerenciamento e análise de dados um indicativo dessa relevância dos dados para os negócios. Nesse contexto, conforme será apresentado na próxima seção, à exploração de *big data* tem sido atribuída crescente relevância competitiva.

### **3. Relevância Competitiva do *Big Data* em Mercados Digitais**

É um consenso na literatura antitruste e econômica a existência de diversos benefícios oriundos da exploração de *big data* em plataformas *online* para as empresas e consumidores (preocupações relacionadas à perda de privacidade destes últimos à parte). Por um lado, a exploração de *big data* pode proporcionar diversos benefícios econômicos. Isso porque essa capacidade pode contribuir para melhorar produtos, serviços e a própria eficiência de seus processos produtivos, bem como prever tendências de mercado e reforçar a segmentação de consumidores por meio do desenvolvimento de modelo de negócios mais dirigidos, com publicidade direcionada e de ofertas personalizadas, por exemplo (OCDE, 2016a, p. 8; ALEMANHA; FRANÇA, 2016, p. 9). Além disso, a exploração dessa capacidade também pode propiciar novas oportunidades de negócio, uma vez que os dados coletados durante a prestação de um determinado serviço ainda podem fornecer informações que sejam aproveitáveis para um serviço distinto. De fato, estudos apontam que empresas alicerçadas em inovação a base de dados (*data-driven innovation*) se beneficiam, em média, de um crescimento em produtividade de 5% (cinco por cento) a 10% (dez por cento) maior do que o crescimento de companhias similares que não adotam essa estratégia (OCDE, 2016a, p. 8).

Debate-se na doutrina, entretanto, se a exploração dessa capacidade pode efetivamente conferir a um agente econômico uma vantagem competitiva duradoura e sustentável. Lambrecht e Tucker (2017, p. 11) apontam que alguns autores argumentam que *big data* pode conduzir para um novo tipo de vantagem, enquanto outros questionam se esse é efetivamente o caso. Apesar disso, conforme observam as autoras, a questão sobre se *big data* pode conferir uma vantagem competitiva sustentável tem recebido pouca atenção sistemática.

Dessa forma, uma vez apresentado o conceito de *big data*, o objetivo da presente seção é o de apresentar as características econômicas que frequentemente são atribuídas aos dados e ao ecossistema de plataformas que exploram *big data*, em torno das quais giram os debates acerca da significância competitiva dessa capacidade.

Como será exposto, além de fatores inerentes ao próprio funcionamento das plataformas *online*, estudos têm destacado certas características econômicas dos dados e de fenômenos do ecossistema de sua utilização nas plataformas *online* que devem ser levadas em consideração na análise concorrencial para se avaliar se a exploração de *big data* pode, ou não, representar uma vantagem competitiva em um determinado caso. O *Center on Regulation in Europe* (CRE), por exemplo, observa que nos modelos de negócio geralmente adotados pelas empresas da *internet* há um grande interesse e investimento na coleta de dados (particularmente pessoais) por meio do oferecimento de produtos gratuitos, que, em geral, exibem mais efeitos de rede e de experiência que os modelos tradicionais. Todavia, conforme observado pelas autoridades de defesa da concorrência alemã e francesa, as discussões sobre esses fatores ainda estão acontecendo, e o que a prática tem demonstrado é que essas características e fenômenos têm impactos ambíguos sobre a concorrência, de modo que não há conclusões universais válidas sobre o tema, sendo necessária



uma análise refinada no caso a caso (ALEMANHA; FRANÇA, 2016, p. 27).

#### a. Potencialização de Efeitos de Rede: *Feedback Loops*

Efeitos de rede têm tido um papel central nas discussões sobre a exploração de *big data* por plataformas *online*. Essa centralidade no debate se encaixa no argumento de que a exploração dessa capacidade potencializaria os efeitos de rede identificados nas plataformas de múltiplos lados (THE ECONOMIST, 2017). A OCDE se refere a efeitos de rede derivados de dados (*data-driven network effects*) que são resultantes de 02 (dois) ciclos de retroalimentação (*feedback loops*), os quais se assemelham a um fenômeno de *leaning-by-doing*: (i) o “*user feedback loop*” e (ii) o “*monetisation feedback loop*” (OCDE, 2016a, p. 10).<sup>12</sup> De um lado, uma plataforma que possua maior número de usuários de seus serviços consegue coletar mais dados para melhorar a qualidade desses serviços e, dessa forma, conquistar novos usuários. Do outro lado, plataformas que exploram dados de usuários também podem melhorar o direcionamento de anúncios e monetizar os serviços, obtendo mais recursos para investir em sua qualidade e atrair novamente mais potenciais usuários e anunciantes. São esses ciclos intermináveis que podem fazer com que fique muito difícil para quaisquer potenciais entrantes e rivais competir contra uma empresa incumbente com um maior número de clientes (OCDE, 2016a, p. 10).

O estudo produzido pelos órgãos de defesa da concorrência alemão e francês se refere aos *feedback loops* como “efeitos de bola de neve” (“*snowball effects*”) que, em razão do acesso diferenciado a dados, poderiam reforçar a marginalização de concorrentes menores com relação às poucas e grandes empresas que operam em determinados setores econômicos onde a coleta e o uso de dados são geralmente vistos como atividades especialmente importantes, como no caso das plataformas *online* de buscas e de rede social, por exemplo (ALEMANHA; FRANÇA, 2016, p. 13). É que, na ocorrência de *user feedback loop* em plataformas incumbentes, esses efeitos podem aumentar as dificuldades e os custos para novas plataformas obterem dados e se estabelecerem no mercado *online* (CRE, 2017, p. 35). Já na existência de *monetisation feedback loops*, uma nova plataforma com menos usuários e menos dados terá maiores dificuldades de direcionar seus anúncios e atrair anunciantes do que as plataformas estabelecidas, conseqüentemente dispondo de menos recursos para melhorar seus serviços.

Esses *feedback loops*, portanto, podem suscitar preocupações competitivas, na medida em que, ao cabo, podem conduzir à monopolização de mercados, em que as estratégias dos seus *players* se baseiam em dados. Sobretudo no contexto das plataformas *online* de múltiplos lados, os efeitos desses *feedback loops* têm sido considerados perigosos, já que a coleta e a exploração de dados pode reforçar os efeitos de rede tradicionalmente verificados nessas plataformas, quando um aumento no número de usuários de uma plataforma a possibilita de coletar mais dados que seus

---

<sup>12</sup> O *user feedback loop* está relacionado ao número de usuários que utilizam um produto ou serviço oferecido pela plataforma online, funcionando de seguinte forma: (1) quanto mais usuários, mais dados são coletados; (2) mais dados coletados ajudam a aprimorar os serviços de forma geral e, ainda, de forma personalizada; (3) melhores serviços atraem mais usuários, dos quais são coletados mais dados, reiniciando-se automaticamente o ciclo. O *monetisation feedback loop* é o da monetização e funciona da seguinte forma: (1) quanto mais usuários, mais dados são coletados; (2) mais dados significam melhores possibilidade de direcionamento de publicidades quando dados adicionais são necessários para identificar perfis e aperfeiçoar os algoritmos de direcionamento; (3) o melhor direcionamento da publicidade aumenta as chances de os usuários clicarem nos anúncios que lhes são disponibilizados e, conseqüentemente, aumenta as receitas da plataforma no modelo comumente usado de pagamento por clique (“*pay-per-click*”); (4) o aumento nas chances de o usuário clicar no anúncio e comprar o produto anunciado atrai mais anunciantes; (5) a atração de mais anunciantes também aumenta as receitas da plataforma; e (6) essa possibilidade de monetização aumenta a possibilidade de investimentos para melhorar os serviços da plataforma, o que atrai mais usuários, recomeçando todo o ciclo novamente.

concorrentes, levando a produtos ou serviços de maior qualidade e, conseqüentemente, a um aumento na sua participação de mercado. Portanto, ainda que a plataforma não pratique qualquer conduta anticompetitiva, esses *feedback loops* podem reforçar a sua dominância e impedir concorrentes ou potenciais entrantes de ganhar clientes (OCDE, 2016a, p. 11).

O CRE, entretanto, observa que a existência do *user feedback loop* depende da relação entre os dados e a qualidade do serviço que, por sua vez, depende do tipo de dado e da sua aplicação. Em muitos casos, a qualidade do serviço depende apenas marginalmente dos dados do usuário, de modo que o *user feedback loop* pode ser raramente existente. Além disso, mesmo quando esse *feedback loop* existe, se o custo de coleta de dados for bem pequeno, os efeitos de rede também serão (CRE, 2017, p. 36).

Com relação ao *monetisation feedback loop*, o CRE chama a atenção para a necessidade de se verificar, caso a caso, a existência desses efeitos de rede. Isso porque eles dependerão do volume, dos tipos de dados e dos algoritmos *self-learning* usados. Neste aspecto, a presença de *feedback loops* se subordina à avaliação das relações entre (i) a quantidade de dados de usuários e o aperfeiçoamento dos algoritmos (para o direcionamento de anúncios, por exemplo), e entre (ii) o aperfeiçoamento destes algoritmos e a atração de novos clientes no mesmo lado ou em outros lados da plataforma *online* (anunciantes, por exemplo). Outro fator a ser levado em consideração diz respeito ao nível de investimento das receitas auferidas pela plataforma com publicidade no financiamento do aperfeiçoamento de seus serviços para usuários (CRE, 2017, p. 37).

Além disso, é relevante observar que pode haver situações em que esses mesmos efeitos de rede podem ser benéficos para a própria entrada de novos *players* em um dado mercado, resultando no decréscimo dos *market shares* das empresas incumbentes. Quando empresas entrantes conseguem atrair usuários por outros motivos, como a introdução de novas tecnologias ou a prestação de serviços com funcionalidades diferenciadas, isso pode aumentar a atratividade dos serviços ou produtos atraindo futuros usuários devido aos mesmos efeitos de rede. Nesse aspecto, tais efeitos também podem estimular a concorrência ao conferir ao entrante o potencial de que sua base de usuários cresça de forma rápida (ALEMANHA; FRANÇA, 2016, p. 28). Os efeitos de rede derivados de dados nas plataformas *online*, portanto, podem amplificar o processo de ganhar ou de perder usuários (OCDE, 2016a, p. 10).

## **b. Economias de Escala e de Escopo Associadas à Exploração de *Big Data***

Conforme pontuado pelas autoridades de defesa da concorrência alemã e francesa, “a significância das vantagens competitivas associadas com a coleta e exploração de dados pode depender crucialmente de se aqueles dados precisam ser coletados em grande escala e/ou escopo”, ou seja, se essa significância só puder ser conquistada por uma incumbente grande e/ou diversificada (ALEMANHA; FRANÇA, 2016, p. 47, tradução nossa).

Economias de escala surgem quando o custo incremental de criar uma unidade adicional declina conforme a escala da produção aumenta. É o que ocorre na exploração de *big data*, cuja estrutura de custos é caracterizada por elevadas economias de escala e de escopo, o que pode facilitar a concentração de mercado nas mãos de alguns poucos *players* (OCDE, 2016a, p. 11). Isso ocorre porque a infraestrutura para o tratamento e o uso de informação é bastante incomum, envolvendo elevados custos afundados iniciais e custos marginais próximos a zero. No caso de *big data*, as tecnologias da informação necessárias para coletar, armazenar e processar dados podem ser muito custosas, por envolver grandes *data centres*, servidores, *softwares* analíticos de dados, conexão à *internet* com *firewalls* avançados e significativos investimentos em recursos humanos

(cientistas da computação e programadores, por exemplo). Contudo, quando o sistema está totalmente operacional, o dado incremental pode treinar e aperfeiçoar os algoritmos a custos baixos e, por consequência, incrementar a qualidade dos próprios produtos ou serviços oferecidos (OCDE, 2016a, p. 11). Ou seja, os próprios *feedback loops* podem levar a economias de escala (SOKOL; COMERFORD, 2016, p. 12).<sup>13</sup> Portanto, a capacidade de *big data* necessita de uma determinada escala de dados para processamento para ser mais lucrativa, sendo, portanto geralmente explorada por *players* maiores que conseguem ter mais acesso a recursos financeiros e dados (OCDE, 2016a, p. 11). O mesmo pode ser dito com relação à variedade de dados, pois o valor do *big data* depende de sua capacidade de combinar diferentes tipos de dados, o que indica a exibição de economias de escopo.<sup>14</sup>

A exibição de economias de escala e de escopo, no entanto, é uma questão controversa na literatura. Para alguns autores, a escala é uma preocupação menor. Graef (2015, p. 12) explica que os benefícios relacionados à disponibilidade de dados para a prestação de serviços em ambos os lados da plataforma (usuários e anunciantes) podem estar sujeitos a retornos de escala decrescentes, de modo que o valor de ter um dado adicional declina conforme o volume de dados aumenta. Por isso, os limites dos benefícios nos 02 (dois) lados da plataforma dependeriam do volume em que os retornos de dados extras começam a reduzir. Esta característica por si só, contudo, não teria o condão de afastar eventuais preocupações relacionadas à economia de escala na exploração de dados. Isso porque, como também observa Graef (2015, p. 12), se os benefícios de um dado adicional começam a declinar somente quando há um montante bastante elevado de dados que integram o *big data*, largos conjuntos de dados podem dar origem a, respectivamente, uma vantagem competitiva para a plataforma incumbente e barreiras à entrada de potenciais concorrentes. Além disso, o estudo conjunto produzido pelos órgãos concorrenciais alemão e francês destaca que o decréscimo no valor marginal de dados ocorre geralmente quando esses são usados para inferências, como no caso das plataformas de busca. Contudo, nem todos os dados são usados exclusivamente com esses propósitos. Mesmo neste caso, para que as inferências sejam úteis e permitam a extração de observações valiosas, o nível de dados a partir do qual se possa constatar um decréscimo de retorno de escala pode ser alto para algumas atividades, sendo de todo modo uma vantagem competitiva para incumbentes. Ainda que o volume necessário para fazer inferências em alguns casos possa ser relativamente baixo, o número de inferências que precisam ser feitas diariamente é alto, o que aumenta o número de dados necessários (ALEMANHA; FRANÇA, 2016, p. 50/51).

O CRE (2017, p. 33) pondera que o retorno marginal de ter mais dados depende, em grande medida, do tipo de dado e de análise que deles é feita, não se referindo, desse modo, exclusivamente à quantidade de dados coletados. Por exemplo, no caso de ferramentas de buscas, se argumenta que a economia de escala é menor para solicitações de busca que são feitas com maior frequência (“*head queries*”) do que para as solicitações de buscas que são feitas com menor frequência (“*tail queries*”) pelos usuários do serviço na plataforma *online*. Isso porque o valor adicionado por uma palavra-chave buscada mais raramente<sup>15</sup> seria maior e contribuiria melhor para o aperfeiçoamento dos resultados da busca. Todavia, como também apontado pelo CRE (2017 p. 33), o nível dessas economias de escala não é claro. Alguns autores consideram que a economia de escala é baixa até

<sup>13</sup> No mesmo sentido, Graef (2015, p. 9) explica que com relação à coleta, armazenamento e análise de dados, os custos envolvidos em organizar os instrumentos necessários para essas atividades são tipicamente custos fixos, enquanto os custos marginais da crescente produção de *big data* são baixos. Como resultado, economias de escala são criadas, as quais podem dar origem a barreiras à entrada de novos concorrentes.

<sup>14</sup> Rubinfeld e Gal (2017, p. 15) se referem, ainda, a “economias de velocidade”, que consistiriam na rapidez com que os dados são coletados, permitindo previsões e respostas rápidas para usuários e concorrentes, como no caso do fenômeno de *now-casting*.

<sup>15</sup> Essa raridade da palavra-chave, entretanto, pode ser bem ampla, incluindo eventos bastante recente com relação aos quais ainda não há informação.

para *tail queries* e que há um retorno marginal decrescente de dados para ambos os *head* e *tail queries*, enquanto outros autores consideram que efetivamente há economias de escala para *tail queries*. Diante dessas discussões, portanto, o CRE (2017, p. 33/34) conclui que a extensão das economias de escala e de escopo é uma questão empírica a ser testada em cada caso concreto com base no tipo de dado e na sua utilização. Em especial, deve ser averiguada a necessidade de ter dados em maior quantidade e variedade para aperfeiçoar a qualidade do uso e do algoritmo.

### **c. Disponibilidade e acesso a dados**

Há entendimento, por parte da doutrina especializada, de que dados são ubíquos e amplamente disponíveis a baixo custo (TUCKER; WELLFORD, 2014, p. 2), de modo que concorrentes seriam capazes de obter os mesmos dados ou dados semelhantes que lhes possibilitariam a extração das mesmas informações relevantes obtidas pelas empresas líderes do mercado.

Com base na doutrina, Sokol e Comerford elencam as seguintes razões para que isso aconteça: (i) usuários/consumidores estão constantemente gerando dados e deixando “rastros” de suas necessidades e preferências a partir do acesso à *internet* e do uso de dispositivos eletrônicos; (ii) dados podem ser fácil e rapidamente coletados de consumidores durante a utilização dos serviços disponibilizados por uma plataforma, sendo que ambos os dados e as ferramentas necessárias para sua coleta e análise estão prontamente disponíveis de várias fontes; (iii) os custos marginais de produção e distribuição de *big data* seriam próximo a zero; (iv) há várias fontes alternativas de dados disponíveis para as empresas, o que reflete a extensão com que consumidores deixam seus rastros digitais na *internet*; e (v) dados podem ser adquiridos de terceiras fontes, de modo que mesmo antes do lançamento de um serviço por uma plataforma, esta já pode ter se beneficiado com informações sobre as preferências de consumidores (SOKOL; COMERFORD, 2017, p. 6).

Os órgãos concorrenciais alemão e francês apontam 03 (três) características que contribuem para a elevada disponibilidade de dados em mercados digitais: (i) o caráter não rival de dados; (ii) a existência de *data brokers*; e (iii) a particularidade de mercados digitais estarem mais propensos a coletar dados.

Nessa linha, os dados são considerados bens não-rivais, o que significa que, em tese, o fato de uma dada empresa ter coletado um dado não impede outras empresas de obterem dado idêntico, ainda que por meio de outras atividades.<sup>16</sup> Além disso, dados também são considerados bens não excludentes, de modo que o mesmo dado de um consumidor pode, em tese, ser usado por diferentes empresas ao mesmo tempo. Essas características econômicas dos dados teoricamente os tornam bens públicos e implicam a redução nos custos de coleta e um barateamento no preço desses insumos. Afinal, a coleta de um dado por uma empresa não impede que o mesmo dado seja simultaneamente capturado por outras, incluindo *data brokers* que podem vendê-los no mercado, nem que os dados possam ser duplicados.

Todavia, como bem ponderam Stucke e Grunes (2016, p. 42), se dados realmente fossem amplamente disponíveis, então as empresas não incorreriam em custos, muito menos ofereceriam serviços gratuitos para obtê-los. Elas simplesmente buscariam dados publicamente disponíveis. É que a natureza não-rival somente contribui para a ampla disponibilidade de dados quando estes

---

<sup>16</sup> Essa possibilidade é particularmente relevante no contexto de *single-homing*, onde concorrentes ainda podem ter acesso ao mesmo tipo de dado a partir da prestação de serviços distintos.

podem ser acessados. Ou seja, a natureza não-rival de dados não garante que eles estejam acessíveis a todos os concorrentes (ALEMANHA; FRANÇA, 2016, p. 36/37). Nesse sentido, Graef (2015, p. 6/7) observa que argumentos sobre a disponibilidade de dados devem ser flexibilizados, considerando que podem ser identificadas situações em que plataformas *online* poderão ser capazes de excluir concorrentes, impedindo ou restringindo o acesso a dados para os quais há poucos ou nenhum substituto. Ainda segundo o autor, empresas cujos modelos de negócio se baseiem na aquisição e monetização de dados têm incentivos para limitar os seus acervos para uso próprio.

O CRE (2017, p. 30) faz referência a 03 (três) tipos de barreiras que enfraquecem ou até mesmo retiram a natureza não-rival de dados: (i) técnicas, como a criptografia de dados, que podem tornar mais difícil ou mesmo impossível a sua coleta; (ii) contratuais; e (iii) legais. Além destas, no entanto, há diversos outros tipos de barreiras que podem bloquear ou dificultar o acesso a dados, como, por exemplo: (iv) barreiras regulatórias e autorregulatórias; (v) barreiras financeiras; (vi) barreiras tecnológicas e, ainda, (vi) barreiras culturais (OCDE, 2015, p. 191), entre outras.

As barreiras contratuais são verificadas quando um agente econômico estipula termos e condições que proíbem ou limitam a coleta, determinados usos, o compartilhamento ou a transferência dos dados capturados para terceiros. Além disso, condições de exclusividade também podem ser impostas no fornecimento e aquisição de dados, impedindo rivais e potenciais concorrentes de os acessarem.

No caso das barreiras legais, regulatórias e autorregulatórias<sup>17</sup>, é o escopo das normas que pode ter impactos sobre a atividade de coleta realizada no segundo elo da cadeia de valor do *big data*, conforme sejam (i) proibidos ou limitados a coleta, a transferência e o compartilhamento de dados, ou (ii) conferidos a titularidade e o direito de propriedade sobre os dados coletados. No primeiro caso, pode-se citar como exemplo as regras que restringem, em algum grau, a coleta de dados pessoais. Já para a segunda situação, pode-se citar como exemplo a proteção autoral que recai sobre determinados bancos de dados.<sup>18</sup> Além disso, nessa categoria, também podem ser encaixadas as restrições decorrentes da existência de monopólios legais, quando os dados coletados pela firma monopolista não podem ser acessados nem copiados por concorrentes interessados em utilizá-los na exploração de outros serviços que estejam sujeitos a um regime de concorrência.<sup>19</sup> Essas restrições, no entanto, também podem recair sobre as próprias tecnologias

---

<sup>17</sup> Por exemplo, com o objetivo de promover a autorregulação da indústria de publicidade *online* e promover a transparência e o controle de dados do consumidor à luz de questões relacionadas à sua privacidade, após o julgamento do caso Google/DoubleClick, a FTC, nos Estados Unidos, anunciou um conjunto de princípios para as atividades de *behavioral advertising* (RAMIREZ, 2016, p. 9). Entre esses princípios, por exemplo, consta a orientação de que os sites que coletam dados de consumidores os informem sobre essa atividade e possibilitem que eles possam optar por ter, ou não, os seus dados capturados para fins publicitários (esses princípios podem ser encontrados na seguinte página eletrônica: <<https://www.ftc.gov/news-events/press-releases/2007/12/ftc-staff-proposes-online-behavioral-advertising-privacy>>. Acesso em: 24 de novembro de 2017). Após a divulgação desses princípios, a *Digital Advertising Alliance*, uma entidade privada que congrega publicitários, redes de publicidade e editores, desenvolveu e divulgou os seus próprios princípios e, ainda, os codificou em um programa de autorregulação (CASTRO, 2011, p. 10).

<sup>18</sup> Conforme Graef (2015, p. 7), “*in addition, providers of online platforms may rely on intellectual property and trade secret law to protect the data they have collected. Copyright and sui generis database protection are of particular relevance with regard to intellectual property rights. While users hold copyright over original posts, photos and videos uploaded over the databases they have created on the basis of the data collected about users*”.

<sup>19</sup> Em 2014, por exemplo, a *Autorité de la Concurrence*, órgão concorrencial francês, determinou que a empresa de energia elétrica GDF Suez fornecesse a concorrentes acesso à sua base de dados de clientes corporativos e domésticos. A decisão foi tomada após as acusações feitas por outro fornecedor, segundo as quais a GDF Suez usava o banco de dados históricos dos clientes em tarifas reguladas para comercializar seus contratos de gás não regulados, setor em que havia competição (FRANÇA, 2014).

utilizadas na captura de dados. Esse é o caso, por exemplo, da proibição que incide sobre o uso de determinados programas na coleta de dados.<sup>20</sup>

Com relação ao aspecto financeiro, é importante considerar que, como visto, a coleta direta de dados pode requerer investimentos significantes e envolver diversos tipos de custos fixos para colocar em funcionamento um *data center* capaz de capturar e explorar um grande volume variado de dados. Custos proibitivos envolvidos na aquisição de determinados dados também podem consistir em uma barreira à sua coleta por plataformas menores ou potenciais entrantes (ALEMANHA; FRANÇA, 2016, p. 37/38), de modo que, também nesses casos, a natureza não-rival será elidida e o acesso exclusivo a um volume e uma variedade de dados superiores pela plataforma incumbente pode significar uma vantagem competitiva.<sup>21</sup>

Outra dificuldade a ser considerada ainda na coleta direta de dados pela plataforma *online* é a de que os dados são geralmente capturados como um subproduto no momento em que o usuário utiliza o serviço, como visto na abordagem dos *feedback loops*. Disso decorre que a coleta dos dados exigirá de um rival ou potencial concorrente esforços para a entrada em 02 (dois) níveis de atividades (*two-level entry*), de modo que será mais difícil para um agente econômico acessar tais dados, particularmente se eles forem únicos e difíceis de serem replicados, ou se houver limitações para a sua portabilidade (RUBINFELD E GAL, 2017, p. 19). Além disso, pela mesma razão, para um potencial entrante obter acesso aos mesmos dados diretamente, ele também precisará fornecer para um número suficientemente amplo de usuários os mesmos tipos de serviços (ou serviços que permitam acesso aos mesmos dados ou a dados similares dos quais se possa extrair informações igualmente relevantes para o negócio). Essa estratégia, entretanto, pode demandar investimentos elevados em pesquisa e desenvolvimento, e a conquista de uma massa suficiente de usuários pode ser particularmente difícil, já que a própria qualidade do novo serviço ou produto pode depender do próprio número de usuários (e dos dados que deles são coletados) em função dos efeitos de rede e de experiência e das economias de escala analisadas nas seções precedentes. Além disso, em alguns casos, pode ser particularmente difícil convencer usuários a fornecer dados pessoais para plataformas entrantes, especialmente quando eles ainda não estão suficientemente seguros quanto à utilização de seus dados pela plataforma ou somente estão dispostos a fornecê-los quando recebem serviços de alto valor agregado (ALEMANHA; FRANÇA, 2016, p. 38).

O fenômeno do *multi-homing*, que ocorre quando os usuários utilizam diversos provedores para obter o mesmo tipo de serviço *online*, é frequentemente mencionado no contexto das discussões sobre a disponibilidade e o acesso a dados (TUCKER; WELLFORD, 2014, p. 3; SOKOL; COMERFORD, 2016, p. 6).<sup>22</sup> Todavia, uma situação de *multi-homing* perfeita tem sido considerada rara, sendo mais comum a verificação de situações intermediárias (entre o *multi-homing* e o *single-homing*) devido a vários custos de transferência (*switching costs*), que podem ser gerados por

---

<sup>20</sup> Por exemplo, na União Europeia, a partir da entrada em vigor da Diretiva 2009/136/EG, o uso *cookies* foi restringido, impondo-se a implementação de mecanismos *opt-in*, por meio dos quais se exige a permissão dos usuários para a utilização dessa ferramenta em cada um dos *sites* por ele visitados (MENDES, 2014, p. 103). Conforme observam Rubinfeld e Gal (2017, p. 24), ao cabo, essa restrição produziu impactos competitivos indiretos ao criar uma vantagem comparativa para o Google sobre seus rivais e potenciais concorrentes, uma vez que tal agente econômico dispõe de outras tecnologias para coleta de dados, o que é feito principalmente por meio sua ferramenta de buscas *online*.

<sup>21</sup> Além disso, na outra ponta, limitações nas tecnologias a que têm acesso determinados usuários também podem beneficiar a coleta de dados pelo fabricante de um determinado *hardware* ou dispositivo eletrônico que seja mais utilizado. Com relação a esse aspecto, Rubinfeld e Gal (2017, p. 13) citam como exemplo o acesso à *internet* por meio de celulares em países em desenvolvimento, o que cria uma vantagem competitiva tecnológica para os fabricantes desses dispositivos e se traduz em barreira à coleta de dados por fabricantes de computadores, por exemplo.

<sup>22</sup> É o caso, por exemplo, quando um indivíduo possui distintos aplicativos de comunicação instantânea em seu *smartphone* (Whatsapp, Facebook Messenger, Skype, Snapchat, Kik, etc.)

razões exógenas (*e.g.* custos de aprendizagem do usuário, por exemplo) ou endógenas (*i.e.*, relacionadas ao comportamento da plataforma já estabelecida, como nos casos de efeitos de rede e de limitações que podem ser impostas à portabilidade de dados, por exemplo) e resultar em efeitos *lock-in*. Com relação a esses últimos aspectos, observa-se que a própria coleta de dados pela incumbente pode aumentar os custos de transferência, já que o provedor mais utilizado pelo usuário tem mais informações sobre ele e, portanto, maior capacidade para personalizar os serviços ou produtos que oferta para esse mesmo usuário. Mesmo nos mercados onde os serviços ou produtos são oferecidos a “preço zero” para usuários finais de um lado da plataforma – e que, por esse mesmo fator, estariam mais propensos a *multi-homing* –, o fenômeno não é perfeito. Isso porque, nesses casos, os usuários prestam mais atenção na qualidade do serviço/produto oferecido (dimensão não relacionada a preço), sendo que novos concorrentes, por não terem acesso ao mesmo volume e variedade de dados, podem não estar capacitados para prestar serviços com a mesma qualidade das empresas incumbentes, considerados os efeitos de rede e de experiência já analisados. Desse modo, o fenômeno do *multi-homing* não necessariamente será relevante para constranger a vantagem competitiva de concorrentes em mercados relacionados a dados (ALEMANHA; FRANÇA, 2016, p. 28/29).

Não apenas custos de transferência elevados podem manter usuários em uma mesma plataforma, mas também os “atalhos” (vieses cognitivos) que os levam a escolher por um determinado serviço. Como observado por Rubinfeld e Gal (2017, p. 20), usuários podem ter um papel relevante na criação ou robustecimento da barreira à entrada na coleta de dados ao escolher por um determinado serviço, o que pode ser afetado por uma série de fatores como a “(des)informação relativa à qualidade e ao preço de bens ou serviços concorrentes dos quais dados são coletados, os custos de transferência, e a (des)informação ou (in)diferença quanto ao preço indireto que eles pagam em termos de perda de privacidade”. Nesse contexto, uma estratégia utilizada em mercados digitais para manter usuários presos a uma plataforma e, dessa forma, limitar o acesso a dados é a instalação de *softwares default* (*i.e.*, que já vêm pré-instalados) nos dispositivos eletrônicos, como navegadores de *internet* e ferramentas de busca padrões, os quais dificilmente são substituídos pelo usuário (KESAN; SHAH, 2006, p. 601).

A possibilidade de aquisição de dados por meio de terceiros também é um fator frequentemente alçado aos debates sobre a disponibilidade e o acesso a dados, sendo considerado um serviço alternativo ou complementar à coleta direta desse insumo (TUCKER; WELLFORD, 2014, p. 3). *Data brokers* funcionam como intermediários que podem coletar dados de uma enorme variedade de fontes<sup>23</sup> e revende-los em um mercado de dados. O acesso intermediado a dados pode ser menos custoso, já que a plataforma deixa de incorrer em custos fixos associados à implantação de um *data center*, transformando-os em custos variáveis incorridos sob demanda (ou seja, a plataforma pode comprar apenas o volume e a variedade de dados de que precisa). Esses *data brokers* podem ainda oferecer outros tipos de serviços, como *cloud computing* e *data analytics* para o armazenamento, processamento e análise de dados para terceiros, reduzindo ainda mais os custos fixos incorridos na exploração de *big data*.

No entanto, o acesso a dados via *data brokers* também pode apresentar algumas desvantagens. Afinal, além de custos de transação<sup>24</sup>, esses intermediários também podem ser constrangidos pelas mesmas barreiras erguidas à coleta direta de dados anteriormente apresentadas,

<sup>23</sup> Além da coleção própria, dados podem ser públicos, detidos por autoridades públicas, adquiridos de outras empresas (plataformas, bancos, seguradoras, administradoras de cartões, etc) ou por meio de arranjos contratuais com proprietários de páginas eletrônicas que permitem a utilização de tecnologias de rastreamento, como *cookies* e *pixels*, etc.

<sup>24</sup> Custos proibitivos para a localização dos dados demandados, identificação de seus titulares, negociação e contratação do intermediário, e monitoramento do contrato podem tornar inviável a coleta indireta de dados. O mesmo ocorrerá se o valor cobrado pelo acesso a um conjunto de dados também for elevado.

de modo que o volume e a variedade de dados muito dificilmente serão comparáveis àqueles diretamente coletados pelas maiores plataformas *online*, que, em alguns casos, se beneficiam da vasta quantidade de dados que pode ser coletada por meio de outros serviços que também sejam prestados pela plataforma (ALEMANHA; FRANÇA, 2016, p. 39).<sup>25</sup> Conquanto algumas plataformas também funcionem como *data brokers* e compartilhem ou comercializem seus dados, os incentivos para tanto podem ser insuficientes em razão dos elevados custos que podem ser incorridos ao longo da cadeia de valor do *big data* até que se possa retirar informações relevantes dos dados explorados e usá-las no negócio e da redução significativa da vantagem competitiva, que possivelmente não compensa as receitas auferidas com o compartilhamento (OCDE, 2015, p. 192).

26

Além de possibilidades mais limitadas de coleta, *data brokers* também podem ter limitações legais, contratuais ou tecnológicas para compartilhar determinados dados com terceiros, especialmente no caso de dados mais sensíveis, como os pessoais, os confidenciais e os sigilosos (ALEMANHA; FRANÇA, 2016, p. 41).

Ainda é relevante considerar que, conquanto alguns tipos de dados possam ser adquiridos por meio de *data brokers* e mesmo outros terceiros, as informações que muitas plataformas *online* precisam para operarem seus algoritmos e *softwares* são específicas e podem não estar disponíveis de pronto no mercado (GRAEF, 2015, p. 9). Essa questão é bastante relevante no caso de dados que perdem o valor rapidamente ao longo do tempo, exigindo que as plataformas recorram constantemente a compras de dados atualizados.

A possibilidade de substituição de dados também é um fator relevante na discussão sobre a disponibilidade e o acesso a dados. Estes, por si só, têm baixo ou nenhum valor. Como visto na cadeia do *big data*, o seu valor decorre principalmente da extração de informações que sejam consideradas relevantes para o processo de tomada de decisão, aperfeiçoamento de serviços ou produtos e algoritmos e, ainda, para a criação de outros tantos novos<sup>27</sup>. Assim, se as mesmas informações podem ser retiradas da análise de outros tipos de dados, a exclusividade sobre determinados dados não consistirá em vantagem competitiva. Esse argumento é particularmente reforçado no contexto das plataformas digitais, onde se considera que a coleta de dados pode ser mais fácil. Entretanto, é importante considerar nessa análise que dados que hoje são tidos como substitutos podem não ser mais no futuro, sendo importante considerar a natureza veloz desses mercados.

Por fim, no que tange à questão da disponibilidade e do acesso, Grunes e Stucke (2016, p. 40) advertem ainda para o fato de que é possível que agentes econômicos detentores de dados

---

<sup>25</sup> Nessa perspectiva, como ponderado pela autoridade concorrencial canadense em seu estudo preliminar da matéria, a despeito da existência de dados oferecidos por *data brokers*, algumas empresas dedicam recursos significantes para coletar, desenvolver e manter seus próprios dados, o que sugere que esses intermediários não ofertam todos os dados que são valorizados (CANADÁ, 2017, p. 8).

<sup>26</sup> Nesse sentido, conforme observado pela OCDE (2015, p. 192), um dos motivos para a ausência de incentivos é que os custos do compartilhamento de dados são percebidos como superiores aos benefícios privados esperados desse compartilhamento.

<sup>27</sup> É importante esclarecer que, epistemologicamente, uma coleção de dados gera informação e diferentes linhas de informação formam conhecimento (THE ECONOMIST, 2010). Ou seja, no sentido técnico, dados são matéria-prima da informação e informações são matéria-prima do conhecimento (FRAZÃO, 2017). Essa distinção, conforme explica Frazão (2017), não é inútil para o debate no campo do Direito da Concorrência, porque “os dados importam, do ponto de vista econômico, na medida em podem ser convertidos em informações necessárias ou úteis para a atividade econômica. Os dados precisam, portanto, ser processados e trabalhados para que possam gerar valor. Se tal constatação não afasta a importância em si dos dados isolados ou ‘crus’, tem o importante papel de realçar o fato de que o mero acesso a dados sem a possibilidade efetiva e eficiente de transformá-los em informação, pode ser insuficiente para resolver diversos problemas competitivos”.



também se engajem em estratégias anticompetitivas para impedir que rivais e potenciais concorrentes tenham acesso aos mesmos dados.<sup>28</sup> Além de buscar concentrações econômicas, esses agentes têm incentivos para a prática de condutas exclusionárias e comportamentos predatórios. Entre essas práticas, a literatura cita como exemplo a implementação de mecanismos que dificultam ou impedem a portabilidade de dados, recusas de contratar e comportamentos discriminatórios, entre outras (NING; WU; ZHAO, 2016).

#### d. Valor Fugaz dos Dados

Conforme o estudo conjunto das autoridades de defesa da concorrência alemã e francesa, o fato de o valor de dados poder decrescer rapidamente no tempo é um outro fator que pode limitar a vantagem competitiva baseada em dados de que desfrutam grandes incumbentes. O valor de muitos dados é temporário, o que significa que eles apenas têm relevância por um curto período de tempo e que sua taxa de depreciação é alta. Esse é o caso, por exemplo, de dados como os de localização utilizados em aplicativos de navegação. Em razão disso, alguns autores argumentam que o controle de dados não pode, por si só, conferir uma vantagem competitiva sustentável (TUCKER; WELLFORD, 2014, p. 4).

Todavia, é importante observar que a característica de transitoriedade não é comum a todos os dados.<sup>29</sup> Alguns deles, como nome, gênero e data de nascimento, por exemplo, podem ter um valor mais permanente, sendo a sua taxa de depreciação menor, o que pode conferir um benefício mais perene que o de dados temporários (CRE, 2017, p. 34).

Além disso, como visto na análise das características técnicas do *big data*, este não é somente volume. Em um modelo de negócios que explora essa capacidade, a rapidez com que os dados devem ser coletados, processados, analisados e atualizados importa, sendo uma tendência que essas atividades ocorram em (quase) tempo real. A atualidade do *big data*, enquanto uma capacidade que permite a tomada de decisões informadas, é importante em diversas situações. Nelas, conquanto os dados possam perder valor muito rápido, esse valor já foi explorado pela empresa com capacidade de coletá-los e analisa-los tempestivamente. Nesses casos, a velocidade com que precisam ser capturados os dados em um curto espaço de tempo gera restrições relevantes sobre os concorrentes. Conforme Graef (2015, p. 10), os tipos de dados que perdem valor em um curto período de tempo fazem com que as empresas tenham que continuar coletando informações atualizadas sobre as preferências e os interesses de seus consumidores para que possam ser capazes de retornar respostas relevantes para usuários e oferecer serviços de publicidade direcionada para anunciantes. O controle sobre esses dados pode até não conferir, por si só, uma vantagem competitiva durável. Contudo, onde o acesso a dados é exclusivo e a coleta está sujeita a economias de escala e efeitos de rede uma vantagem competitiva pode emergir.

---

<sup>28</sup> Segundo Stucke e Grunes (2016, p. 40), “*businesses – to obtain or maintain their competitive advantage – will have strong incentives to limit their competitors’ access to these datasets, prevent others from sharing datasets, and will likely be adverse to data-portability policies that threaten their data-related competitive advantage. Companies will battle over who gets the valuable consumers data*”. No mesmo sentido, Kuilwijk (2016, p. 5) explica que “*business models that rely on securing Big Data a competitive edge may also devise strategies in order to prevent their rivals from gaining access to that same data. After all, a company whose fundamental business model is based on collecting data will need to ensure an advantage over their competitors. This must be approached with caution, as the temptation to prevent rivals from accessing the data may be deemed exclusionary under antitrust law*”.

<sup>29</sup> Conforme o estudo conjunto das autoridades concorrenciais alemã e francesa: “*First, as regards the impact of data obsolescence, it should be kept in mind that although the value of some data may diminish relatively quickly over time, not all data are transient in value. Depending on the market, some data, such as gender, names address, data of birth, job, etc., may not lose their value over time. Thus, a company having such data at its disposal may have a lasting advantage over its competitors*” (ALEMANHA; FRANÇA, 2016, p. 50)

Além disso, conforme observado pelo CRE (2017, p. 34), mesmo dados históricos podem ser usados para aperfeiçoar ou desenvolver novas funcionalidades e algoritmos, sendo que essas inovações terão valor maior e mais permanente, capitalizando-se o valor fugaz desses dados e transformando-o em valor mais permanente por meio de utilidades e desenvolvimentos.

#### **e. Disponibilidade e Acesso a Tecnologias de Processamento e Análise de Dados**

Alguns autores argumentam que dados, por si só, não são valiosos e, portanto, o seu acúmulo não poderia gerar vantagem competitiva. Nesse sentido, por exemplo, o professor Hal Varian (que, diga-se de passagem, também é *chief-economist* do Google) entende que dados são inúteis, a menos que possam ser transformados em conhecimento e ação usando *data analytics* e expertise, o que requer um investimento pesado em *hardware*, *software* e *know-how*. Para ele, portanto, o que seria mais importante para o sucesso do negócio seria a habilidade de desenvolver novos algoritmos e atrair expertise (OCDE, 2016b, p. 3). Com a mesma perspectiva, Boutin e Clemens (2017, p. 6) apontam para evidências em mercados baseados no uso de dados, que indicariam que largos volumes, por si só, raramente causariam um problema competitivo, sendo a incapacidade de coletar e analisar esse largo e variado conjunto de forma tempestiva que poderia levar um agente econômico ao insucesso no mercado.<sup>30</sup>

Em se tratando da exploração de *big data*, tão importante quanto o acesso a dados em volume e variedade é o acesso a tecnologias de coleta, de armazenamento<sup>31</sup> e, principalmente, de processamento e análise desses dados em alta velocidade para a extração de informações relevantes (FRAZÃO, 2017). Os elos da cadeia de valor do *big data* são bastante imbricados, de modo que de nada adianta coletar e armazenar um grande volume de dados, se não for possível processá-los e analisa-los oportunamente (sendo válido o caminho inverso também).<sup>32</sup>

Essa interdependência é reforçada pelo fato de que os *softwares* de *data analytics* e os algoritmos podem ser aprimorados a partir do próprio volume e da qualidade dos dados a que têm acesso tendo em vista sua capacidade autodidata (*self-learning*) e de aprendizado com a experiência (*learning by doing*). Rubinfeld e Gal (2016, p. 29) apontam para a existência de potencial correlação entre os dados e a qualidade dos algoritmos em função da presença dos *feedback loops* já abordados neste trabalho e das mudanças que podem ser verificadas nos algoritmos que evoluem a partir da própria observação dos dados que processam e analisam. Dessa forma, a dificuldade de acesso a

---

<sup>30</sup> Com relação a esse aspecto, os autores concluem que dados não devem ser comparados ao “petróleo do século XXI”, mas a vento: “dados circulam e são em grande medida disponíveis. Assim como vento, eles precisam ser capturados e transformados em algo de valor. Ambos moinhos de vento e tecnologias de processamento de dados certamente melhoram quando testados em condições reais. Contudo, nesse contexto, é a experiência acumulada que é valiosa, e não o acúmulo de vento, nem de dados históricos” (BOUTIN; CLEMENS, 2017, p. 6, tradução nossa).

<sup>31</sup> Conquanto o enfoque da presente seção seja o acesso a tecnologias de processamento e análise de dados, é importante que, no que couber, as autoridades concorrenciais também examinem a existência de tais barreiras para o acesso a tecnologias e infraestruturas para a coleta e o armazenamento de dados.

<sup>32</sup> A habilidade de extrair informação de dados, portanto, não depende exclusivamente das quantidades de dados disponíveis, mas também dos algoritmos que analisam esses dados. Como esses algoritmos não são todos da mesma qualidade, este atributo dos serviços prestados pelas plataformas *online* em mercados relacionados a dados não é exclusivamente atribuível a um maior ou menor conjunto de dados pelos concorrentes. (ALEMANHA; FRANÇA, 2016, p. 49). No mesmo sentido, Graef (2015, p. 14) observa que “*at the same time, it has to be recognized that in addition to relevant and recent data, engineering resources and a well-functioning underlying technology including an algorithm are required to successfully operate an online platform. But even though access to a large and up-to-date database is in itself no guarantee for the success of an online platform, data remains a necessary input of production for the delivery of services to users and advertisers that are of the quality and relevance they expect*”.

um volume suficiente de dados de qualidade, ao cabo, pode comprometer a qualidade da tecnologia de processamento aplicada por um agente econômico.

Com o desenvolvimento de algoritmos autodidatas, a curva de experiência se tornou mais acentuada, o que, em determinadas circunstâncias, pode aumentar a vantagem competitiva de o agente econômico ter sido o primeiro a se instalar em um mercado (*first-mover advantage*), tornando a entrada de concorrentes “seguidores” mais difícil, já que terão que tentar adquirir esses algoritmos no mercado ou investir em outros que consigam fornecer serviços similares para usuários, ainda que de forma diferente (CRE, 2017, P. 35).<sup>33</sup>

Conforme Rubinfeld e Gal (2017, p. 28), a disponibilidade e a qualidade de algoritmos usados para o processamento e análise de *big data* podem criar uma barreira tecnológica para potenciais entrantes. Conquanto algumas ferramentas analíticas estejam disponíveis de forma gratuita e abertas ao uso, isso, por si só, pode não ser suficiente para resolver o problema da disponibilidade e do acesso. É necessário avaliar, ainda, a sofisticação da análise que o volume de dados variados coletados requer vis-à-vis a capacidade de essas ferramentas atingirem esse grau de sofisticação. Como muitos dos dados coletados são não-estruturados, a utilização de *softwares* capazes de promover a fusão desses dados e a leitura das informações relevantes que podem ser reveladas a partir dessa combinação se torna ainda mais importante, assim como o desenvolvimento de capacidades de *self-learning*, entre elas o reconhecimento de padrões e o *machine learning*.

Rubinfeld e Gal (2017, p. 29) também observam que a diferença nas ferramentas de análise de dados detidas pelas empresas pode representar vantagens comparativas para algumas delas. Por esse motivo, como apontam os autores, algumas empresas que exploram *big data*, como o Google, por exemplo, investem quantias significativas no desenvolvimento ou na aquisição de algoritmos sofisticados. É importante ponderar, contudo, que o declínio nos custos de capacidade computacional (incluindo de armazenamento e memória, de processamento e de acesso à banda larga) têm resultado em crescente acessibilidade a ferramentas analíticas (RUBINFELD E GAL, 2017, p. 29). Como já exposto, a expansão da computação em nuvem e de ferramentas de fonte aberta permitiu que alguns custos fixos desses investimentos sejam convertidos em custos variáveis, o que possibilita que pequenos negócios também consigam ter sucesso na utilização de dados para entrar no mercado de forma mais fácil.<sup>34</sup>

Não obstante, ainda é possível que os potenciais entrantes e concorrentes menores enfrentem barreiras de ordem técnica e legal no acesso às tecnologias aplicadas, inclusive, durante toda a cadeia de valor de *big data* e, em especial, na atividade de processamento e análise de dados. Embora o cerne deste artigo seja a questão do acesso a dados – assunto pouco aprofundado na

---

<sup>33</sup> Além disso, como apontado pela OCDE (2016a, p. 10/11), a ausência de limitações físicas para a quantidade e variedade de dados coletados nas plataformas digitais e o conhecimento ilimitado que pode ser extraído da fusão, processamento e análise desses dados por meio de algoritmos aumentaram a inclinação da curva de aprendizagem dos negócios, permitindo um desenvolvimento mais acentuado e duradouro durante a fase de crescimento da empresa incumbente. O relatório ainda cita os efeitos multiplicadores de *databases*, que fazem com que os retornos desses dados sejam difíceis de exaurir. Ao chegar à fase de maturidade, a dimensão de conhecimento acumulado pelo detentor dos dados pode ser tão grande a ponto de impedir um *player* menor de efetivamente exercer pressão competitiva ou eventual *catch up*. Assim, cria-se um potencial para resultados de “*winner-takes-all*” e “*market tipping*”.

<sup>34</sup> A OCDE (2015, p. 144/145) reporta que *data analytics* está se tornando mais acessível para *startups* e pequenas e médias empresas, estimando que a adoção dessas ferramentas se intensificará conforme o volume de dados gerados continue a crescer. Ainda de acordo com essa organização, enquanto no passado as ferramentas de coleta, armazenamento e processamento de dados eram custosas e, na maioria das vezes, apenas acessíveis a grandes empresas, órgãos governamentais e universidades, atualmente a adoção de *data analytics* tem sido bastante facilitada pelo declínio nos custos de armazenamento e de processamento de dados.

experiência e doutrina pátrias<sup>35</sup> –, cumpre ainda destacar que determinadas tecnologias podem estar protegidas por direitos de propriedade intelectual. Como o desenvolvimento de *softwares* demanda esforços intelectuais e investimentos substanciais, mas o custo marginal de sua reprodução é baixo, a apropriação desses esforços iniciais por meio de imitações é passível de ocorrer. Por isso, as regras de propriedade intelectual conferem aos titulares de *softwares* a exclusividade temporária do direito de uso e de exploração desse produto, como recompensa pelos investimentos, criatividade e esforços intelectuais aplicados no seu desenvolvimento (ABDURAZZAKOV, 2009, p. 122).

O embate entre a proteção da propriedade intelectual e da livre concorrência não é novo. Aqueles que conciliam essas proteções vêm entre elas uma relação de complementariedade. Nessa perspectiva, a propriedade intelectual não seria repressora da concorrência no longo prazo, em função de servir de estímulo para que agentes econômicos em um mercado invistam em pesquisa e desenvolvimento com o objetivo de alcançar inovações na forma de aperfeiçoamento ou criação de novos produtos, serviços ou processos de produção que acirrarão a concorrência, gerando, ao cabo, efeitos pró-competitivos (ARIBONI; ROCHA; MARTINO, 2012, p. 210). Entretanto, há uma preocupação das autoridades antitruste com comportamentos abusivos nos registros e licenciamentos de direitos de uso de propriedades intelectuais, como os *softwares* de processamento de dados. Esses comportamentos podem prejudicar o desenvolvimento dos negócios de plataformas *online* de concorrentes e de potenciais entrantes, incrementando seus custos e funcionando como barreiras à exploração de *big data* e, ao cabo, à própria entrada no mercado, o que justificaria a análise e intervenção antitruste.

#### **4. A Variável Competitiva *Big Data* na Análise Concorrencial**

Não se pode assumir, de pronto, que o acesso a dados ou a tecnologias é invariavelmente difícil. É necessário analisar se as circunstâncias fáticas permitem concluir que os ativos detidos pela plataforma *online* incumbente funcionam como barreiras à entrada ou à expansão de concorrentes, que, sendo reforçadas pela presença de efeitos de rede decorrentes dos *feedback loops* e pela exibição de economias de escala e de escopo, podem impedir o ingresso de novos competidores, marginalizar rivais menores e até mesmo convergir para a monopolização do mercado, em evidente prejuízo da concorrência (LASSERRE; MUND'T, 2017, p. 91). Ou seja, o que as discussões já apresentadas indicam é que não há uma resposta universal e invariável para todo e qualquer caso. A exploração de *big data* pode ser ou não uma vantagem competitiva relevante e possibilitar o exercício de poder de mercado a depender de diversas circunstâncias do caso concreto, que precisam ser consideradas nas análises realizadas pelas autoridades de defesa da concorrência em mercados que se baseiam em dados.

Essa capacidade pode não ser alcançável por potenciais concorrentes e rivais por diversos fatores (ou uma combinação deles) que devem ser considerados pela autoridade concorrencial no caso concreto, tendo em vista as seguintes premissas básicas: (i) por definição, *big data* não consiste apenas em um amplo conjunto de dados, mas em uma capacidade de processar um largo volume de dados variados em alta velocidade, extraindo valor de sua conversão em informações relevantes, (ii) os principais atributos do *big data* são volume, variedade, velocidade e valor, mas a intensidade e a relevância de cada um deles pode variar, conforme o mercado relevante em que a capacidade

---

<sup>35</sup> O tema da relação entre a defesa da concorrência e a proteção da propriedade intelectual já foi fértilmente abordado em diversos trabalhos no Brasil e, por isso, não se pretende tecer profundos comentários a esse respeito neste trabalho, que se limita a indicar que a proteção intelectual de tecnologias de coleta, armazenamento e processamento e análise de dados também pode funcionar como uma barreira para a exploração de *big data*, de modo a conferir vantagem competitiva significativa e duradoura para aqueles agentes econômicos que possuam essa capacidade.

seja explorada (RUBINFELD; GAL, 2017, p. 17), e (iii) para que seja possível extrair valor dos dados em estratégias de *big data*, é necessário percorrer toda a sua cadeia de valor, ainda que por meio de intermediários.

Dito isto, para melhor avaliar o poder de mercado em plataformas *online* cujos modelos de negócios sejam baseados na exploração de *big data*, devem ser examinadas as condições para a realização das atividades de cada elo da sua cadeia de valor, e não simplesmente na coleta de dados. Ou seja, essa análise deve ser feita a partir da verificação dos elementos necessários para se alcançar a capacidade de exploração de *big data*, o que requer o exame de cada anel da sua cadeia de valor, a fim de se verificar se um rival ou potencial concorrente teria condições de atingir o último anel, retirando valor do uso de informações relevantes extraídas pelo processamento e análise de um grande e variado volume de dados em alta velocidade (senão, em tempo real).

Não é possível assumir a premissa de que dados invariavelmente têm uma natureza não-rival e estão sempre disponíveis, pois, em algumas situações, o acesso a dados por um agente econômico não será possível. Como já apresentado, nesse elo da cadeia de valor do *big data*, diversos fatores podem fazer com que determinados dados não estejam (facilmente) ao alcance de um agente econômico que pretenda explorá-los, assim como podem existir barreiras que atingem o acesso às próprias tecnologias utilizadas para a coleta direta desses dados. Por isso, no caso concreto, é relevante avaliar a disponibilidade de dados e a possibilidade de acesso a suas fontes ou a fontes alternativas pelo agente econômico que pretenda coletá-los. As barreiras à coleta direta de dados serão baixas ou inexistentes quando, por exemplo, os dados estiverem publicamente disponíveis. No entanto, ainda que este não seja o caso, as barreiras à coleta de dados também poderão ser baixas, se eles puderem ser acessados e coletados direta e simultaneamente (ou, ao menos, tempestivamente) da mesma fonte e de forma gratuita ou a custos não proibitivos. Nessa análise, também é relevante considerar se os mesmos dados ou dados substitutos (*i.e.*, que, ao cabo, permitiriam a extração das informações igualmente relevantes) ainda podem ser coletados de fontes alternativas, de forma tempestiva.

Se forem verificadas elevadas barreiras à coleta direta de dados nesse elo da cadeia de valor do *big data*, como visto, uma alternativa de que dispõem os rivais ou os potenciais concorrentes é a obtenção de dados por meio de intermediários, como órgãos do governo, bancos de dados públicos, *data brokers* ou mesmo plataformas *online* que também se engajem em atividades de comercialização de dados. Não obstante, como também já endereçado, para que essa alternativa seja viável e suficiente, é necessário analisar fatores como a quantidade, a qualidade e a velocidade (atualidade) dos dados que podem ser coletados e ofertados pelos intermediários, bem como os custos envolvidos na sua aquisição (preço dos dados, custos de transação, etc.) – o que é particularmente relevante quando os dados exibem taxas de depreciação elevadas –, e a disponibilidade de tecnologias de transferência dos dados coletados pelo intermediário para o agente econômico, incluindo-se neste quesito aspectos contratuais, legais e tecnológicos.

Como visto, o armazenamento de dados que ocorre no terceiro elo da cadeia de valor vem passando por reduções significativas nas restrições à sua realização. Apesar disso, em algumas situações, os agentes econômicos ainda podem se deparar com fatores que podem embaraçar o desenvolvimento dessa atividade, como a existência de custos de transferência (*switching costs*) e de efeitos *lock-in* na troca de programas de gerenciamento de base de dados, e a presença de barreiras

legais ou regulatórias que limitem o armazenamento de dados (RUBINDFELD; GAL, 2017, p. 27).<sup>36</sup>

O processamento e a análise de dados, no terceiro elo, tem um papel central na cadeia de valor do *big data*, uma vez que é a sintetização dos dados coletados e armazenados o processo que lhes confere valor ao convertê-los em informação por meio de variadas técnicas da ciência da computação e dos dados, como, por exemplo, o processamento de linguagem natural que converte conjuntos de dados em linguagem compreensível ao ser humano, o reconhecimento de padrões e as técnicas de *machine-learning* (capacidade computacional de aprendizado sem programação explícita) e *deep learning* (uso de algoritmos de aprendizagem profunda que buscam correlações no conjunto de dados sem orientações específicas). No entanto, sobretudo nesse elo, um agente econômico pode se deparar com barreiras que limitam o seu acesso a tecnologias capazes de sintetizar em alta velocidade o grande volume de dados variados coletados e armazenados. Como exemplo, se pode citar os elevados custos para a aquisição de tecnologias e infraestruturas de processamento e análise de dados, menor qualidade da tecnologia empregada (o que pode ser potencializado em função dos fenômenos de *self-learning* e de *learning by doing*), e barreiras comportamentais, como o abuso de propriedade intelectual.

Por fim, ainda que um agente econômico alcance a capacidade de explorar *big data*, isso não significa que ele poderá efetivamente utilizar as informações extraídas dos dados. Com relação ao ponto, Rubinfeld e Gal (2017, p. 29) apontam que a plataforma pode ter dificuldades de procurar, encontrar ou alcançar consumidores relevantes, o que pode acontecer, por exemplo, em contextos de *single-homing*. Além disso, o agente econômico pode se ver limitado a utilizar determinadas informações, como as de caráter pessoal, por exemplo, o que pode inclusive fazer parte de uma estratégia competitiva da empresa para criar reputação em mercados em que consumidores valorizam a sua privacidade ou evitar pressões para o compartilhamento de dados.

## 5. Conclusão

Como exposto, conquanto ainda não haja uma definição consensual para *big data*, em linhas gerais, o termo se refere a um grande volume de dados digitais variados, que são coletados, armazenados e processados analiticamente em alta velocidade, particularmente com o objetivo de identificar padrões e comportamentos e fazer correlações, extraíndo-se, dessa forma, informações relevantes que lhes conferem valor e podem ser utilizadas em processos de decisão informada nas mais variadas áreas.

É um consenso na literatura antitruste e econômica a existência de diversos benefícios econômicos oriundos da exploração de *big data* em plataformas *online* para as empresas e consumidores. No entanto, cada vez mais, são suscitados argumentos na literatura no sentido de que a exploração dessa capacidade pode efetivamente conferir a um agente econômico uma vantagem competitiva duradoura e sustentável, que funcionaria como uma barreira à entrada de novos concorrentes e poderia levar à manutenção no mercado de poucas grandes empresas incumbentes com poder econômico (particularmente nos setores *online*).

Como esse debate ainda tem recebido pouca atenção sistemática da doutrina, particularmente no Brasil, o presente trabalho objetivou avaliar se e em quais circunstâncias a

---

<sup>36</sup> Algumas normas podem proibir o armazenamento de determinados dados. Aqui no Brasil, por exemplo, o § 1º do artigo 43 do Código de Defesa do Consumidor – CDC proíbe os cadastros e bancos de dados de conter informações negativas do consumidor referentes a período superior a 5 (cinco) anos.

capacidade de *big data* pode ser considerada uma vantagem competitiva para as empresas que a exploram em mercados digitais, exigindo a sua inserção nas análises concorrenciais, especialmente no que concerne à avaliação de barreiras à entrada e rivalidade.

Como visto, não há uma resposta única sobre se a exploração de *big data* pode ser ou não uma vantagem competitiva relevante e possibilitar o exercício de poder de mercado. Na verdade, em mercados digitais onde o uso de dados é intensivo, a exploração da capacidade de *big data* poderá representar uma variável competitiva – e que deve ser levada em consideração na análise concorrencial –, se determinados fatores estiverem presentes.

Além de circunstâncias inerentes ao próprio funcionamento das plataformas *online*, a doutrina tem destacado certas características econômicas dos dados e de fenômenos do ecossistema de sua utilização nas plataformas *online* que podem fazer com que, em alguns casos, a exploração dessa capacidade represente uma verdadeira vantagem competitiva sustentável e duradoura. Entre esses fatores, destacam-se a necessidade de se avaliar se:

- (i) Os efeitos de rede tradicionalmente identificados nas plataformas de múltiplos lados são potencializados pela presença de efeitos de rede derivados do uso de dados (o “*user feedback loop*” e o “*monetisation feedback loop*”);
- (ii) Existem economias de escala (volume) e/ou de escopo (variedade) associadas à exploração de *big data* nas atividades da plataforma;
- (iii) Os dados – e, conseqüentemente, as tecnologias correlatas para o seu processamento e análise – não estão efetivamente disponíveis ou não são acessíveis para os demais *players* e potenciais entrantes no mercado; e
- (iv) Apesar do decurso do tempo, é mantido o valor dos dados coletados pela plataforma.

A partir da verificação da presença desses fatores, a autoridade concorrencial poderá avaliar se, no caso concreto, a variável *big data* efetivamente representa, para a plataforma incumbente, uma vantagem competitiva capaz de erguer barreiras à entrada, impedir o ingresso de novos competidores, marginalizar rivais menores e até mesmo convergir para a monopolização do mercado.

É importante, no entanto, que o exame desses fatores seja feito sob a ótica da cadeia de valor do *big data*. Ou seja, deve ser analisado como cada um deles impacta cada elo, impedindo que um determinado *player* se mova ao longo da cadeia e alcance o último elo (ainda que por meio de intermediários), onde teria condições de retirar valor do uso de informações relevantes extraídas do processamento e análise de um grande e variado volume de dados em alta velocidade dos dados a partir da verificação dos elementos necessários.

## 6. Bibliografia

ABDURAZZAKOV, Bekhzod A. Condenados ao conflito? Uma análise do papel da proteção da propriedade intelectual na licença de tecnologia de software e a análise antitruste. In: TIMM, Luciano Benetti; PARANAGUÁ, Pedro (Org.). **Propriedade Intelectual, antitruste e desenvolvimento**. Rio de Janeiro: FGV, 2009. p. 121-180.

ALEMANHA. BUNDESKARTELLAMT; FRANÇA. AUTORITÉ DE LA CONCURRENCE. **Competition Law and Data**. 2016. Disponível em: <<http://www.autoritedelaconcurrence.fr/doc/reportcompetitionlawanddatafinal.pdf>>. Acesso em: 31 de maio de 2017.

ARIBONI, Pietro; ROCHA, Lígia Ferreira Marcondes; MARTINO, Fernanda Dalla Valle. A concorrência e a propriedade intelectual: especial atenção ao sistema dos desenhos industriais à luz do SBDC e INPI. **Revista do IBRAC - Direito da Concorrência, Consumo e Comércio Internacional**, São Paulo, v.21, p.209-225, jan./jun 2012. Disponível em: <<https://www.ibrac.org.br/revistas.htm>>. Acesso em: 30 maio 2017.

ASHTON, Kevin. That 'Internet of Things' Thing. **RFID Journal**, Hauppauge, jun. 2009. Disponível em: <<http://www.rfidjournal.com/articles/view?4986>>. Acesso em: 30 de maio de 2017.

BAGNOLI, Vicente. The Big Data Relevant Market. **Concorrenza e Mercato - Antitrust, Regulation, Consumer Welfare, Intellectual Property**, Milão, vol.23, 2016. Disponível em: <<https://ssrn.com/abstract=3064792>>. Acesso em: 22 de novembro de 2017.

BOUTIN, Xavier; CLEMENS, Georg. Defining 'Big Data' in Antitrust. **Competition Policy International (CPI) – The Global Resource for Antitrust and Competition Policy: Antitrust Chronicle**, v.1, n.2, 7 p., 2017. Disponível em: <<https://ssrn.com/abstract=2938397>>. Acesso em: 30 de maio de 2017.

BROWN, Brad; CHUI, Michael; MANYIKA, James. Are you ready for the era of “big data”? **McKinsey Quarterly**, out. 2011. Disponível em: <<https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/are-you-ready-for-the-era-of-big-data>>. Acesso em: 25 de novembro de 2017.

CANADÁ. The Competition Bureau of Government of Canadá. Big data and Innovation: Implications for Competition Policy in Canada. **Competition Bureau**. Preliminary Version, 2017. Disponível em: <<http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/04304.html>>. Acesso. Em: 23 de novembro de 2017.

CASTRO, Daniel. Benefits and Limitations of Industry Self-Regulation for Online Behavioral Advertising. **The Information Technology & Innovation Foundation**, dez. 2011. Disponível em: <https://www.ntia.doc.gov/files/ntia/2011-self-regulation-online-behavioral-advertising.pdf>. Acesso em: 05 de novembro de 2017.

CENTER ON REGULATION IN EUROPE (CRE). Big Data and Competition Policy: Market Power, Personalized Pricing and Advertising. **SSRN's eLibrary**, 21 fev. 2017. Disponível em: <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2920301](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2920301)>. Acesso em: 31 de maio de 2017.

COLE, Miranda. **Data in EU merger control**. CPI Antitrust Chronicle February 2018. Winter 2018, Volume 1, N. 2. Disponível em: <<https://www.competitionpolicyinternational.com/category/antitrust-chronicle/>>. Acesso em: 03 de março

ESTADOS UNIDOS DA AMÉRICA. Federal Trade Commission. Data Brokers: A Call for Transparency and Accountability. 2014a. Disponível em: <



<https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf>.

Acesso em: 29 de novembro de 2017.

FEIJÓ, Bruno Vieira. O que a revolução dos dados pode fazer por sua empresa? **Revista Exame**, 29 out. 2013. Disponível em: <<http://exame.abril.com.br/pme/a-revolucao-dos-dados/>>. Acesso em: 18 de abril de 2017.

FRANÇA. **Autorité de la concurrence. Décision n° 14-MC-02 du 9 septembre 2014**, relative à une demande de mesures conservatoires présentée par la société Direct Energie dans les secteurs du gaz et de l'électricité. Disponível em: <[https:// Autorité de la concurrence. Décision n° 14-MC-02 du 9 septembre 2014](https://Autorité de la concurrence. Décision n° 14-MC-02 du 9 septembre 2014)>. Acesso em: 20 de novembro de 2017.

FRAZÃO, Ana. **Direito da Concorrência: pressupostos e perspectivas**. São Paulo: Saraiva, 2017.

GRAEF, Inge, Market Definition and Market Power in Data: The Case of Online Platforms. **World Competition: Law and Economics Review**, v. 38, n. 4, p. 473-506, 2015. Disponível em: <<https://ssrn.com/abstract=2657732>>. Acesso em: 31 de outubro de 2017.

GERADIN, Damien; KUSCHEWSKY, Monika, **Competition Law and Personal Data: Preliminary Thoughts on a Complex Issue**. **Discussion Papers Tilburg Law and Economics Center**, 2013. Disponível em: <<https://ssrn.com/abstract=2216088> ou <http://dx.doi.org/10.2139/ssrn.2216088>>. Acesso em: 31 de maio de 2017

HU, Han; WEN, Yonggang; CHUA, Tat-Seng; LI, Xuelong. Toward scalable systems for big data analytics: A technology tutorial. **IEEE Access**, v. 2, p. 652–687, maio 2014. Disponível em: <[https:// http://ieeexplore.ieee.org/document/6842585/?reload=true](https://http://ieeexplore.ieee.org/document/6842585/?reload=true)>. Acesso em: 01 de dezembro de 2017.

JAPKOWICZ, Nathalie; STEFANOWSKI, Jerzy. Big data analysis: new algorithms for a new society. *Sutudies in Big Data*, v. 16, p. 302-329, 2016.

KAROLCZY, Pola; LE CROY, Kyle. Navigating the Digital Age: The European Commission's Differing Approaches to Merger Control and Abuse of Dominance in the Digital Sector. **CPI Antitrust Chronicle February 2018. Winter 2018, Volume 1, N. 2**. Disponível em: <<https://www.competitionpolicyinternational.com/navigating-the-digital-age-the-european-commissions-differing-approaches-to-merger-control-and-abuse-of-dominance-in-the-digital-sector/>>. Acesso em: 03 de março de 2018.

KESAN, Jay P.; SHAH, Rajiv C. Setting Software Defaults: Perspectives from Law,

Computer Science and Behavioral Economics. **Notre Dame Law Review**, v. 82, n. 2, p. 583-634, 2006. Disponível em: <<http://scholarship.law.nd.edu/cgi/viewcontent.cgi?article=1293&context=ndlr>>. Acesso em 28 de outubro de 2017.

KUILWIJK, Kees Jan. Big data, the internet of things, and competition law. **Westlaw**, 2016. Disponível em: <<https://www.akd.nl/Downloads/PublicatiesPDF-NL/07-06-2016%20DATA%20COMPETITION%20LAW%20KJkuilwijk.pdf>>. Acesso em: 06 de novembro de 2017.

- LASSERRE, Bruno; MUNDT, Andreas. Competition Law and Big Data: The Enforcers' View. *Italian Antitrust Review*. [S. l.], vol. 4, n. 1, 2017, 17 p. Disponível em: <https://http://iar.agcm.it/article/view/12607/11417>. Acesso em: 08 de dezembro de 2017.
- LAMBRECHT, Anja; TUCKER, Catherine E. **Can Big Data Protect A Firm From Competition? Competition Policy International**, 2017. Disponível em: <https://www.competitionpolicyinternational.com/can-big-data-protect-a-firm-from-competition/>>. Acesso em: 06 de novembro de 2017.
- MAYER-SCHÖNBERGER, Viktor e CUKIER Kenneth. **Big data: a revolution that will transform how we live, work and think**. London: John Murray, 2014, 272 p.
- MENDES, Laura Schertel. **Privacidade, proteção de dados e defesa do consumidor: linhas gerais de um novo direito fundamental**. São Paulo. Editora: Saraiva. 2014.
- MOTTA, Massimo; SALGADO, Lucia Helena. **Política de concorrência: teoria e prática e sua aplicação no Brasil**. 1. ed. Rio de Janeiro: Elsevier. 2015.
- NING, Susan; WU, Han; ZHAO, Yangdi. Whether “Big Data” Could Facilitate Monopoly in the Platform Economy and How We Shall Step In. **CPI's Asia Column**, 2016. Disponível em: <https://www.competitionpolicyinternational.com/wp-content/uploads/2017/06/Asia-Column-June-Full-2.pdf>. Acesso em 24 de novembro de 2017
- ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO (OCDE). Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data. **OECD Digital Economy Papers**, n. 222, 2013. Disponível em: [https://< http://www.oecd-ilibrary.org/exploring-data-driven-innovation-as-a-new-source-of-growth\\_5k47zw3fcp43.pdf?contentType=%2fns%2fWorkingPaper&itemId=%2fcontent%2fworkingpaper%2f5k47zw3fcp43-en&mimeType=application%2fpdf&containerItemId=%2fcontent%2fworkingpaperseries%2f20716826&accessItemIds=&option6=imprint&value6=http%3a%2f%2foecd.metastore.ingenta.com%2fcontent%2fimprint%2foecd](https://http://www.oecd-ilibrary.org/exploring-data-driven-innovation-as-a-new-source-of-growth_5k47zw3fcp43.pdf?contentType=%2fns%2fWorkingPaper&itemId=%2fcontent%2fworkingpaper%2f5k47zw3fcp43-en&mimeType=application%2fpdf&containerItemId=%2fcontent%2fworkingpaperseries%2f20716826&accessItemIds=&option6=imprint&value6=http%3a%2f%2foecd.metastore.ingenta.com%2fcontent%2fimprint%2foecd)>. Acesso em: 10 de outubro de 2017.
- \_\_\_\_\_. Data-driven Innovation for Growth and Well-being. 2014. Documento disponível em: <https://www.oecd.org/sti/innno/data-driven-innovation-interim-synthesis.pdf>>. Acesso em: 29 de outubro de 2017.
- \_\_\_\_\_. Data-Driven Innovation: Big Data for Growth and Well-Being. Paris: OECD, 2015. Disponível em: [https://http://www.keepeek.com/Digital-Asset-Management/oecd/science-and-technology/data-driven-innovation\\_9789264229358-en#page1](https://http://www.keepeek.com/Digital-Asset-Management/oecd/science-and-technology/data-driven-innovation_9789264229358-en#page1)>. Acesso em: 04 nov. 2017.
- \_\_\_\_\_. Big data: bringing competition policy to the digital era. **Background note by the Secretariat**, 2016a. Disponível em: <http://www.oecd.org/competition/big-data-bringing-competition-policy-to-the-digital-era.htm>>. Acesso em: 24 de novembro de 2017.
- \_\_\_\_\_. Big Data: Bringing Competition Policy to the Digital Era. **Executive Summary by the OCDE Secretariat**, 2016b. Disponível em: [www.oecd.org/daf/competition/big-data-bringing-competition-policy-to-the-digital-era.htm](http://www.oecd.org/daf/competition/big-data-bringing-competition-policy-to-the-digital-era.htm). Acesso em: 23 de novembro de 2017.

OXFORD LIVING DICTIONARY. **BIG DATA**. Oxford, 2017. Disponível em: <[https://en.oxforddictionaries.com/definition/big\\_data](https://en.oxforddictionaries.com/definition/big_data)>. Acesso em: 27 de abril de 2017.

\_\_\_\_\_. **DATA**. Oxford, 2017. Disponível em: <<https://en.oxforddictionaries.com/definition/data>>. Acesso em: 27 de abril de 2017.

WORLD BANK. **World Development Report 2016: Digital Dividends**. Washington: World Bank, 2016. Disponível em: <[http://documents.worldbank.org/curated/en/961621467994698644/pdf/102724-WDR-WDR2016Overview-ENGLISH-WebResBox-394840B-OUO-9.pdf](http://documents.worldbank.org/curated/en/961621467994698644/pdf/102724WDR-WDR2016Overview-ENGLISH-WebResBox-394840B-OUO-9.pdf)>. Acesso em: 20 de maio de 2017.

PRESS, Gill. 12 Big Data Definitions: What's Yours? **Forbes**, 2014. Disponível em: <<https://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/#7cca0c1813ae>>. Acesso em: 12 de abril de 2017.

\_\_\_\_\_. Very Short History Of Big Data. **Forbes**. 2013. Disponível em: <<https://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/#44e6411a65a1>>. Acesso em: 12 de abril de 2017.

RAMIREZ, Edith. **Deconstructing the Antitrust Implications of Big Data**. Keynote Remarks of FTC Chairwoman for 43rd Annual Conference on International Antitrust Law and Policy Fordham Competition Law Institute. 2016. Disponível em: <[https://www.ftc.gov/system/files/documents/public\\_statements/1000913/ramirez\\_for\\_dham\\_speech\\_2016.pdf](https://www.ftc.gov/system/files/documents/public_statements/1000913/ramirez_for_dham_speech_2016.pdf)>. Acesso em: 23 de novembro de 2017.

RIFFE, Ronald. Big Data precisa de armazenamento definido por software. **CIO From IDG**, 2016. Disponível em: <<http://cio.com.br/tecnologia/2016/03/07/big-data-precisa-de-armazenamento-definido-por-software/>>. Acesso em: 31 de maio de 2017.

RUBINFELD, Daniel L.; GAL, Michal S. Access Barriers to Big Data (August 26, 2016). 59 **Arizona Law Review**, v. 59, p. 339-381, 2017. Disponível em: <<https://ssrn.com/abstract=2830586>>. Acesso em: 20 de novembro de 2017.

SOKOL, D. Daniel, COMERFORD, Roisin E. Does Antitrust Have a Role to Play in Regulating Big Data? 2016. **Cambridge Handbook of Antitrust, Intellectual Property and High Tech**. Cambridge University Press, Forthcoming, 2016. Disponível em: <<https://ssrn.com/abstract=2723693>>. Acesso em: 31 de outubro de 2017.

STUCKE, Maurice; GRUNES, Allen. Debunking the Myths Over Big Data and Antitrust. **CPI Antitrust Chronicle**, n. 276, mai. 2015, Disponível em: <<https://ssrn.com/abstract=2612562>>. Acesso em: 31 de maio de 2017.

\_\_\_\_\_. **Big data and Competition Policy**. Nova York: Oxford University Press, 2016, 371 p.

SULLIVAN, E. Thomas; HARRISON, Jeffrey L. **Understanding Antitrust and Its Economic Implications**. 6.ed. New Providence, NJ: LexisNexis. 2014. (The Understanding Series).

THE ECONOMIST. **Data, data everywhere**. 2010. Disponível em: <<http://www.economist.com/node/15557443>>. Acesso em 20 de abril de 2017.

- \_\_\_\_\_. **The world's most valuable resource is no longer oil, but data: the data economy demands a new approach to antitrust rules.** 2017. Disponível em: <https://www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource>. Acesso em 15 de novembro de 2017
- TUCKER, Darren S.; WELLFORD, Hill B. Big mistakes regarding big data. **Antitrust Source, American Bar Association**, 2014. Disponível em: <https://ssrn.com/abstract=2549044>>. Acesso em: 06 de novembro de 2017.
- VEZZOSO, Simonetta. Competition policy in a world of big data. In: OLLEROS, Xavier; ZHEGU, Majlinda (editores). **Research Handbook on Digital Transformations.** Cheltenham: Edward Elgar, 2016, Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2717497](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2717497)>. Acesso em: 30 de maio de 2017.
- VISCUSI, W. Kip; VERNON, John M.; HARRINGTON, Joseph E. **Economics of Regulation and Antitrust.** 4ª ed. Cambridge: The MIT Press, 2005.