Maastricht University

Maastricht Centre for European Law

Maastricht Centre for European Law

Master Working Paper

2018/9

Priscilla Tollini

The Era of Big Data

The Interplay Between Competition Law and Data Protection

The MCEL Master Working Paper series seeks to give excellent Master students the opportunity to publish their final theses and to make their work accessible to a wide audience. Those wishing to submit papers for consideration are invited to send work to: mcel@maastrichtuniversity.nl
Our submission guidelines and further information are available at:
https://www.maastrichtuniversity.nl/research/institutes/mcel/mcel-publications#master

Published in Maastricht, September 2018

Faculty of Law
Maastricht University
Postbox 616
6200 MD
Maastricht
The Netherlands

This paper is to be cited as MCEL Master Working Paper 2018/9

Table of Contents

1. Introduction

The ongoing technological revolution and the expansion of the digital economy have boosted the possibilities to collect, store and process data in unprecedented ways. This has led to the development and rise of Big Data, a phenomenon that is changing how companies compete with each other and how they gather and use data of individuals. Businesses nowadays perceive Big Data as an important asset that can bring significant competitive advantages over their rivals. Social network and search engine companies such as Facebook and Google are common examples of data-driven business models in which Big Data plays a crucial role in ensuring the businesses' success.

Notwithstanding the numerous positive impacts that the Big Data phenomenon can bring to society, ranging from transportation safety to improvements in healthcare to reduction in energy consumption, it is important to recognize that it also entails significant risks and implications to different fields of law.[1] Competition law and data protection are two examples of legal domains that can face several challenges in a Big Data world. Taking into account the relevance of these challenges to the economy and to the lives of individuals and consumers, governmental agencies and international organizations have been continuously studying about how Big Data can affect competition and data protection policies.

The topic has been one of the key focuses of the European Commission and it has been mentioned in various speeches delivered by Commissioner Margrethe Vestager.[2] National competition authorities including the German Bundeskartellamt and

[1] European Parliament, *Report on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI))* (Committee on Civil Liberties, Justice and Home Affairs, 20 February 2017) 5.

[2] Margrethe Vestager, "Big Data and Competition" (EDPS-BEUC Conference, Brussels 29 September 2016) <https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/big-data-and-competition_en> accessed 01 December 2017; Margrethe Vestager, "Competition in a big data world" (DLD 16, Munich 17 January 2016) <https://ec.europa.eu/commission/commissioners/2014- 2019/vestager/announcements/competition-big-data-world_en> accessed 04 December 2017; Margrethe Vestager, "Helping people cope with technological change" (Rencontres de Bercy, Paris 21 November 2017) <https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/helping-people-cope-technological-change_en> accessed 04 December 2017; Margrethe Vestager, "What competition can do – and what it can't" (Chilling Competition Conference, 25 October

the French Autorité de la Concurrence have also joined forces to research about the implications of data to competition law.[3] Likewise, the Organization for Economic Co-operation and Development (OECD) has published several papers which address this discussion.[4]

Apart from the competition law standpoint, the Big Data debate also extends to data protection concerns, as the use of data by companies frequently includes the processing of personal data, which in turn may trigger the application of the European Union's recently reformed data protection framework. Such framework – which now comprises the General Data Protection Regulation (GDPR) and the Directive 2016/680 – set a high standard of protection across the EU by restricting the conditions under which personal data can be gathered and imposing several obligations on companies that collect and process personal information.[5]

Against this backdrop, the present research paper aims to tackle the following question: what are the main competition law and data protection challenges stemming from the growing use of Big Data by businesses around the world? Additionally, what are the possible implications arising from the interplay between Big Data, competition law and data protection? The paper is structured as follows. **Part 2** scrutinizes the relationship between Big Data and competition law by identifying possible legal challenges regarding market power, mergers and acquisitions, exclusionary practices, price discrimination and market transparency. **Part 3** highlights potential areas of difficulty involving the application of the EU data protection framework to the Big Data

---

2017) <https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/what-competition-can-do-and-what-it-cant_en> accessed 04 December 2017.

[3] Autorité de la Concurrence and Bundeskartellamt, *Competition Law and Data* (10 May 2016) <https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.pdf?__blob=publicationFile&v=2> accessed 11 May 2018.

[4] OECD, *Big Data: Bringing Competition Policy to the Digital Era* (Executive Summary, 29-30 November 2016); OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being* (OECD Publishing, Paris, 2015); OECD, *Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data"* (OECD Digital Economy Papers, No. 222, OECD Publishing, Paris, 2013).

[5] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119/1; Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA [2016] OJ L 119/89.

world, including the applicability of the GDPR's material scope, the purpose limitation principle, the issue of consent and the data minimization principle. **Part 4** assesses possible overlapping legal concerns of competition law and data protection, including the right to data portability, the question of whether competition authorities should consider data protection concerns throughout their merger and abuse of dominant position analyses, and the discussion of whether there is a scope for cooperation between competition and data protection authorities when it comes to cases that touch upon both fields of law. At last, **Part 5** concludes by presenting an overview of the main competition and data protection law challenges encountered in the era of Big Data, as well as possible concerns arising from the interplay between these two legal fields. The article proposes that the rise of Big Data defies competition authorities, data protection supervisors, policymakers, academia, and practitioners to debate about possible solutions to the numerous legal challenges posed by the Big Data phenomenon.

2. Big Data and Competition Law

2.1. Preliminary Considerations: Understanding Big Data and the Big Data Value Chain

Before embarking upon the antitrust legal challenges encountered in the age of Big Data, it is necessary to first clarify what is meant with the terms 'Big Data' and 'Big Data value chain' since both terms will be recurring in the text that follows. Although the term 'Big Data' has been repeatedly used in the legal and business literature, there is still no clear-cut consensual definition of what Big Data means.[6] The concept of Big Data is known to be a notoriously difficult concept to find a common definition with widespread acceptance.[7] The European Data Protection Supervisor defined Big Data as 'large amounts of different types of data produced at high speed from multiple sources,

[6] OECD, *Data-Driven Innovation for Growth and Well-being* (Interim Synthesis Report, 2014) 11; OECD (2013) (n 4) 11-12.
[7] Yvonne McDermott, 'Conceptualising the right to data protection in an era of Big Data' (2017) 4:1 Big Data & Society 4.

requiring new and more powerful processors and algorithms to process and to analyze'.[8] Another definition, put forward by a report of the European Parliament, described Big Data as 'the collection, analysis and the recurring accumulation of large amounts of data, including personal data, from a variety of sources, which are subject to automatic processing by computer algorithms and advanced data-processing techniques'.[9]

The most common definition found in the literature indicates that the main difference between 'normal data' and 'Big Data' are the following four characteristics, also known as Big Data's four V's: (i) the *volume* of data; (ii) the *velocity* at which the data are generated, collected, processed and analyzed; (iii) the *variety* of the data and information gathered; and (iv) the *value* of the information extracted from the data.[10]

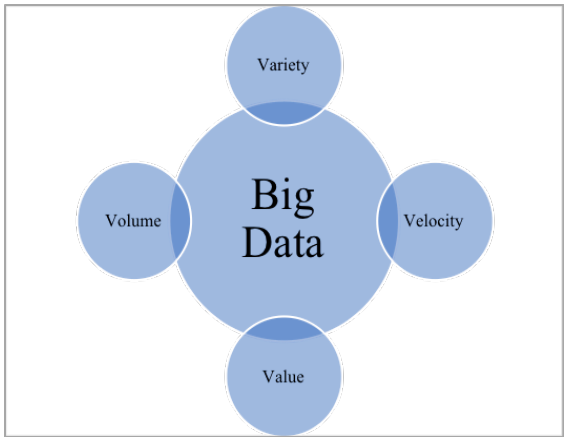

Figure 1: Big Data's Four V's

Over the past years, the *volume* of data collected worldwide has remarkably increased with digitization and the migration of social and economic activities to the Internet.[11] From smartphones and wearable fitness devices to smart meters and autonomous vehicles, the technological revolution underway in the 21st century is

[8] European Data Protection Supervisor, 'Big Data and Digital Clearing House' (*European Data Protection Supervisor*) <https://edps.europa.eu/node/3671> accessed 14 May 2018.
[9] European Parliament (n 1) 4.
[10] Daniel Rubinfeld and Michal Gal, 'Access Barriers to Big Data' (2017) 59:2 Arizona Law Review 346; Allen Grunes and Maurice Stucke, 'No Mistake About It: The Important Role of Antitrust in the Era of Big Data' [2015] Competition Policy Internatioanl Antitrust Chronicle <https://www.competitionpolicyinternational.com/assets/Uploads/StuckeGrunesMay-152.pdf> accessed 30 April 2018 2; Maurice Stucke and Allen Grunes, *Big Data and Competition Policy* (First edition, Oxford Competition Law 2016) 16; Marc Bourreau, Alexandre de Streel and Inge Graef, 'Big Data and Competition Policy: Market Power, Personalised Pricing and Advertising' [2017] SSRN <https://ssrn.com/abstract=2920301> accessed 01 May 2018 11; Autorité de la Concurrence and Bundeskartellamt (n 3) 4.
[11] Rubinfeld and Gal (n 10) 346.

changing not only the way how society communicates, travels and does business but also the amount of data that is produced by each individual.[12] With the rise of the so-called Internet of Things, the sources of generation of data will expand even more and ordinary household devices such as refrigerators, light bulbs and garbage cans will also be able to collect data. Businesses are increasingly developing alternative ways to collect data about its consumers, in particular what is their location, what they search for online, what they spend their money with and how and where they shop.[13]

In addition to the volume, the *velocity* at which data are generated, collected, processed and analyzed has also increased with technological enhancements.[14] As a consequence, many companies are already able to make real-time or contemporaneous 'nowcasts', which allows them to predict what is happening as it occurs.[15] This is useful particularly for cases in which the value of the data decreases as the data becomes older – for instance, the use of geo-location data to assist commuters in avoiding traffic.[16] Thus, velocity is related to the 'freshness' of data in the sense that new data (e.g. the current location of an individual) may render older data (e.g. the location of an individual one year ago) outdated, stale, and perhaps useless in economic terms.[17]

The *variety* of data may also have a direct link to its value, as the value of data can significantly increase when data from various sources are combined together and new information is obtained from the mixture – a phenomenon known as 'data fusion'.[18] In the words of Stucke and Grunes, data fusion is useful for companies to 'identify and improve their profiles of individuals; better track their activities, preferences, and vulnerabilities; and better target them with behavioral advertising'.[19] Finally, due to the

[12] Grunes and Stucke (n 10) 1.
[13] Grunes and Stucke (n 10) 2.
[14] Rubinfeld and Gal (n 10) 346-347.
[15] According to Rubinfeld and Gal (n 10) 353, 'nowcasting' is the capacity of a certain company to use the velocity of data collection to discern and track trends in users' conduct in real-time. See also Stucke and Grunes (n 10) 19.
[16] Stucke and Grunes (n 10) 19.
[17] As stated by Tucker and Wellford, *historical data* can be studied to extract possible trends, but compared to *new data* – which can be used to take real-time decisions, such as which advertisement to serve – it has little value. Darren Tucker and Hill Wellford, 'Big Mistakes Regarding Big Data' (2014) 14:2 The Antitrust Source 4. See also Rubinfeld and Gal (n 10) 346-347, 353.
[18] Rubinfeld and Gal (n 10) 347; Stucke and Grunes (n 10) 21.
[19] Stucke and Grunes (n 10) 21-22; Maurice Stucke and Allen Grunes, 'Debunking the Myths Over Big Data and Antitrust' [2015] Competition Policy International Antitrust Chronicle <https://ssrn.com/abstract=2612562> accessed 30 April 2018 2-3.

decreasing costs and time needed to collect and analyze the data through different data analytics mechanisms, an increase in *volume*, *velocity* and *variety* can consequently also produce an increase the *value* of data.[20]

Having clarified the core characteristics of Big Data, it is also opportune to explore what the term 'Big Data value chain' means. The Big Data value chain consists of three main stages: (i) data *collection*, (ii) data *storage* and (iii) data *analytics*.[21]

Figure 2: The Big Data Value Chain

At the first stage of the Big Data value chain, data is collected. Data can be collected *directly* – when there is a direct contact between the firm and the person from which data is collected – or *indirectly* – when the firm procures data from third parties, such as data brokers.[22] An example of data collection is when individuals voluntarily provide their personal data in exchange for a 'free' service, such as Facebook's social media platform or Google's search engine.[23] Secondly, data is stored – either internally or externally. When large amounts of data are involved, the storage requires either the investment in expensive data centers or the use of cloud computing.[24] At last, the third stage of the Big Data value chain consists of the analysis of data by applications and algorithms that are able to extract relevant information and identify correlation patterns.[25] It is usual for raw data to have a low value at the data *collection* stage, but the value subsequently increases at the data *analytics* stage when unstructured information is

[20] Stucke and Grunes (n 10) 24.
[21] For further information regarding the Big Data value chain, see Rubinfeld and Gal (n 10) 349. See also Bourreau, Streel and Graef (n 10) 11.
[22] Bourreau, Streel and Graef (n 10) 11.
[23] Manon Oostveen, 'Identifiability and the applicability of data protection to big data' (2016) 6:4 International Data Privacy Law 299, 301.
[24] Bourreau, Streel and Graef (n 10) 13.
[25] Bourreau, Streel and Graef (n 10) 14.

transformed into actionable and insightful information.[26] The information extracted can serve multiple purposes, such as improving products or services, personalizing prices or marketing strategies, better targeting consumers with tailor-made advertisements, and increasing productive efficiencies.[27]

From these preliminary considerations, it is important to keep in mind that the term 'Big Data' is distinguished by its four characteristics of *volume*, *velocity*, *variety* and *value*, and that the 'Big Data value chain' consists of the *collection*, *storage* and *analysis* stages of data.

2.2. Competition Law Challenges in the Era of Big Data

The reliance on data by businesses is not a recent phenomenon, given that companies have been using consumer data for marketing strategies long before the existence of Big Data.[28] However, technological developments have revolutionized the possibilities to collect and use such data, also in ways which may give rise to competition concerns. This has led competition authorities such as the German Bundeskartellamt, the French Autorité de la Concurrence, and the European Commission to focus in understanding the usages and implications of data for competition law.

Nowadays, considering the large potential of profitability with Big Data, numerous online companies have implemented data-driven business models in which personal data is a key strategic input.[29] Many of these data-driven business models involve two-sided markets – (i) the market between the company and consumers and (ii) the market between the company and advertisers.[30] Firstly, such companies offer 'free' technology-

[26] Rubinfeld and Gal (n 10) 342.
[27] Rubinfeld and Gal (n 10) 342; Bourreau, Streel and Graef (n 10) 14.
[28] Autorité and Bundeskartellamt (n 3) 8-9.
[29] Grunes and Stucke (n 10) p. 3; Stucke and Grunes (n 19) 2.
[30] Lapo Filistrucchi and Tobias Klein, 'Price Competition in Two-Sided Markets with Heterogeneous Consumers and Network Effects' [2013] NET Institute Working Paper N. 13-20 <https://ssrn.com/abstract=2336411> accessed 24 July 2018 2. According to Filistrucchi and Klein, markets can be considered two-sided when four conditions are met: (i) platforms sell two different types of products or services to two different groups of customers; (ii) the demand of at least one group of customers is dependent on the demand of the other; (iii) there are indirect network effects (see

based products and services to consumers. By making use of these products and services, consumers provide companies with valuable information and data about themselves. Secondly, after having obtained the consumer data, companies are contracted by advertisers to target advertising for the right audience. Thus, by offering consumers allegedly 'free' services, companies are able to acquire valuable personal data and to assist advertisers in better targeting their ads.[31] Although many of the services provided by such companies are presented as 'free', much criticism is voiced towards the fact that consumers in reality pay for these services by sharing their personal information.[32]

With an increasingly amount of companies adopting data-driven business models, data has a more significant influence on companies' strategic decision-making. As a consequence, Stucke and Grunes sustain that there is currently a competitive arms race amongst such companies to determine who will win the race of connecting the data bucket with the money bucket and, thus, who will ultimately be able to sustain a data-related competitive advantage over its rivals.[33]

Taking these matters into account, it is possible to identify at least five challenges that Big Data and data-driven business models can pose to the competition law framework, including (i) the establishment and perpetuation of market power in favor of a few players, (ii) strategic mergers and acquisitions to obtain better access to data, (iii) exclusionary conducts that deprive competitors from access to data, such as refusal to access data, discriminatory access to data and exclusive agreements, (iv) price discrimination between different customer groups, and (v) increased market transparency and risk of collusion between players. Such challenges will be addressed in the subsections below and are summarized in Annex 1.

---

below for a further explanation of 'network effects'); and (iv) any increase in price asked by the platform cannot be transferred from one customer group to the other.

[31] Damien Geradin and Monica Kuschewsky, 'Competition Law and Personal Data: Preliminary Thoughts on a Complex Issue' [2013] SSRN <https://ssrn.com/abstract=2216088> accessed 29 July 2018 2-3; Stucke and Grunes (n 19) 2.

[32] See, for instance, Maurice Stucke and Ariel Ezrachi, 'When Competition Fails to Optimize Quality: A Look at Search Engines' (2016) 18:1 Yale Journal of Law and Technology 72. See also Autorité and Bundeskartellamt (n 3) 3.

[33] Stucke and Grunes (n 10) 1.

2.2.1. The Establishment of Market Power in Favor of a Few Players

The first legal challenge that attracts attention in the competition law analysis is the assessment of data as a factor to establish market power.[34] Data can be a source of market power when, for instance, an entrant player needs to have access to a large amount of *volume* or *variety* of data to be able to compete on a certain market – also known as traditional scale and scope economies.[35] Entrant players or smaller companies can face more difficulty in *collecting* data directly from its users or customers compared to incumbent players due to the fact that they have a smaller number of users in their platforms.[36] With less users, there is less data to be collected. In these instances, entrant players could – at least in theory – have an indirect access to data by purchasing it from third-parties, such as data brokers.[37] If the costs of collecting alternative data through data brokers are low enough and not prohibitive, then such costs cannot be considered as entry barriers.[38] Nonetheless, entrant players could still encounter obstacles in accessing data, as perhaps third-parties are not willing to sell the data to their competitors or perhaps it is impracticable to match the *volume* or *variety* of the incumbent company's dataset.[39] Thus, when entrant players lack the possibility to access and collect sufficient *volume* or *variety* of data directly by themselves or indirectly through data brokers, they cannot ensure their competitiveness on the market vis-à-vis incumbent players because of the high entry barriers.[40]

In the EU, a 'barrier to entry' is commonly understood as 'any cost that must be borne by the operators in a given industry, even if that cost must be or must have been borne by the already-established or 'incumbent' operators'.[41] Barriers to entry can arise from various factors, including problems of access to data, the existence of essential

[34] Bourreau, Streel and Graef (n 10) 30-37; Autorité and Bundeskartellamt (n 3) 3, 11.
[35] See Rubinfeld and Gal (n 10) 349-357 for an in-depth analysis of access barriers in relation to each stage of the Big Data value chain.
[36] Autorité and Bundeskartellamt (n 3) 12. Nonetheless, for another point of view that does not perceive data as a barrier to entry, see Tucker and Wellford (n 17) 6-9.
[37] According to Tucker and Wellford (n 17) 7-8, third-party sourcing options (e.g. other online providers or data brokers) are expanding and falling in cost.
[38] Rubinfeld and Gal (n 10) 350-351.
[39] Autorité and Bundeskartellamt (n 3) 12.
[40] Autorité and Bundeskartellamt (n 3) 11.
[41] Luis Ortiz Blanco, *Market Power in EU Antitrust Law* (First edition, Hart Publishing 2012) 60.

facilities in the hands of dominant companies, economies of scale and network effects.[42] It is relevant to point out, however, that high entry barriers and their subsequent link to market power have to be assessed on a case-by-case basis.[43] This because some undertakings could intend to enter a certain market despite the existence of significant barriers to entry, while other undertakings may not show any interest in entering a certain market even though the market is free of entry barriers.[44] The important aspect to consider is whether, and to what extent, it is *probable* that other competitors enter the market and that the market power of the established undertaking is limited.[45]

To put this into context, take the example of the social network and search engine industries, in which market shares are highly concentrated in the hands of a few players, such as Facebook and Google. These data-driven online markets tend to be less competitive due to the existence of strong scale and network effects.[46] 'Network effects' occur when the use of a certain platform by a customer impacts the value of that platform for other customers.[47] Network effects may be *direct* (e.g. when the high number of Facebook users increases the value of the platform for other users, which benefit from having many 'friends' on the same social network) or *indirect* (e.g. when the high number of Facebook users increases the value of the platform for advertisers, which benefit from the possibility of reaching more customers).[48] Such network effects can ultimately strengthen or lessen competition – it depends on whether they can give innovative entrants the possibility to quickly enter the market and expand their consumer base, or on whether they favor market concentration and stand as a barrier to entry.[49]

---

[42] A further explanation of such factors that can influence entry barriers is given below.
[43] See, for instance, the conclusion of Rubinfeld and Gal (n 10) 354 of which barriers to entry caused by economies of scale, scope, and speed are not necessarily insurmountable. To illustrate, the author gave the examples of Google's displacement of Yahoo! as the main search engine used in the United States and Facebook's displacement of MySpace in the social network market.
[44] Blanco (n 41) 59.
[45] Blanco (n 41) 59.
[46] According to Rubinfeld and Gal (n 10) 352, barriers to entry can exist when dominant companies have achieved substantial economies of scale by means of large 'sunk' investments. In these cases, entrant companies which are unable to achieve a minimum viable scale to compete on the market are likely to switch to a more profitable alternative investment. See also Bourreau, Streel and Graef (n 10) 35-36 and Autorité and Bundeskartellamt (n 3) 13.
[47] Inge Graef, 'Market Definition and Market Power in Data: The Case of Online Platforms' (2015) 38:4 World Competition: Law & Economics Review 473; Stucke and Ezrachi (n 32) 81; Rubinfeld and Gal (n 10) 355.
[48] Daniel Sokol and Roisin Comerford, 'Antitrust and Regulating Big Data' (2016) 23:5 George Mason Law Review 1148; Graef (n 47) 476; Stucke and Ezrachi (n 32) 81-82; Bourreau, Streel and Graef (n 10) 7; Rubinfeld and Gal (n 10) 355.
[49] Autorité and Bundeskartellamt (n 3) 28.

When the latter situation occurs, it is relevant to recognize that these network effects act as a relevant barrier to entry to potential competition from other companies which may consider entering the market.

The existence of network effects in turn can lead to the so-called 'snowball effects' or 'positive feedback loops'.[50] Snowball effects or positive feedback loops run as follows: an incumbent company has a large number of users, which in turn allows it to collect more data about these users, which in turn leads to the provision of services with a better quality (e.g. better targeted advertisements), which in turn creates a qualitative comparative advantage, which in turn attracts even more users and enables the company to collect even more data.[51] Such self-reinforcing trend could potentially eliminate entrant companies and ultimately harm competition through the monopolization of data-driven markets.[52] The feedback loop phenomenon is not limited to online data-driven industries – it can also be observed, for example, in the World Cup. The more people watch the games, the more advertisers are attracted to the event and willing to sponsor the teams. As a result, the quality of football events raises, drawing even more people to watch the games.

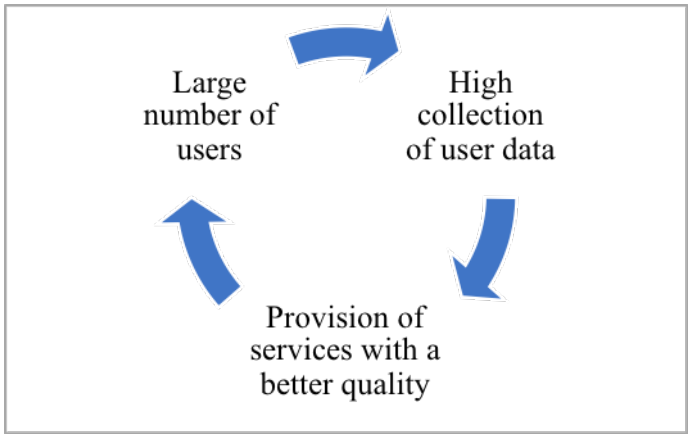

Figure 3: Positive Feedback Loops in Data-Driven Markets

[50] Rubinfeld and Gal (n 10) 356; Sokol and Comerford (n 48) 1147-1148; Stucke and Ezrachi (n 32) 87; Bourreau, Streel and Graef (n 10) 35-36; Autorité and Bundeskartellamt (n 3) 13. For a more economic perspective of feedback loops, see also Filistrucchi and Klein (n 30) 3.
[51] Rubinfeld and Gal (n 10) 356; Bourreau, Streel and Graef (n 10) 35-36; Stucke and Ezrachi (n 32) 87; Autorité and Bundeskartellamt (n 3) 13, 28.
[52] Autorité and Bundeskartellamt (n 3) 13, 28.

Moreover, economies of scale, scope and speed should also be considered when assessing data as a factor to establish market power.[53] In relation to the *value* of data, three different reflections should be carried out: first, is there a link between having *access to more data* and the quality of a certain product?; second, is there a link between the ability to *combine various types of data* and the quality of a certain product?; finally, is there a link between *the age of the data* and the quality of a certain product?[54] To exemplify, think of the Facebook platform. Can Facebook enhance the quality of its platform by having access to more data (e.g. constantly collecting data about which posts its users liked), or by combining different kinds of data (e.g. not only users' names, but also their location, workplace, studies, interests), or by having access to real-time data (e.g. when users 'check-in' in a nearby location)? The answer seems to be positive. With more *volume*, *variety*, and *velocity* of data, Facebook could enhance its platform and make it more tailored to the needs and interests of both users and advertisers. Consequently, this would attract more users and fall back into the concepts of 'network effects' and 'feedback loops'.

Besides network effects, feedback loops, and economies of scale, scope and speed, other aspects should be taken into consideration when assessing a company's market power in data-driven online industries. As explained above in Part 2.2, many online companies operate within 'multi-sided markets' and usually offer alleged 'free' services to its users while charging advertisers for targeted advertising services.[55] When analyzing possible abusive behaviors or proposed mergers, the first step taken by competition authorities is to define the relevant product markets. In the EU, the relevant product market comprises all products or services which are regarded as interchangeable or substitutable by consumers, depending on several factors such as product characteristics, prices and intended usages.[56] In the case of multi-sided markets, a common concern is *how* to define relevant markets.[57] A first question that

[53] Sokol and Comerford (n 48) 1147; Rubinfeld and Gal (n 10) 352.
[54] Bourreau, Streel and Graef (n 10) 7-8; Tucker and Wellford (n 17) 4.
[55] Graef (n 47) 476; Stucke and Ezrachi (n 32); Filistrucchi and Klein (n 30) 2; Rubinfeld and Gal (n 10) 357.
[56] Richard Whish and David Bailey, *Competition Law* (Eighth edition, Oxford University Press 2015); Commission Notice on the definition of relevant market for the purposes of Community competition law (97/C 372/03) OJ C 372/5 para 7.
[57] Graef (n 47) 489; Tucker and Wellford (n 17) 4; Autorité and Bundeskartellamt (n 3) 27.

comes to mind is: can the user side of the market be considered a relevant market even though it is theoretically free of charge? Some will argue that it can indeed be considered a relevant market given that in reality users pay for these platform services by sharing their personal information with the companies. Another question that arises is: can this interaction between users and platform providers be considered an economic exchange at all? Many online platforms like Facebook and Google do not sell user data to third parties, but merely use the collected information as an *input* for advertising services.[58] Thus, perhaps such user data cannot be seen as a 'payment', but rather an intermediary product that is not further traded on any market.[59] Graef supports that competition concerns related to datasets may not be sufficiently taken into account when competition authorities rely solely on relevant markets for the *end* products and services (e.g. market of social networks or market of search engines).[60] For this reason, Graef argues in favor of the definition of an *additional* relevant market for 'user data' even in cases where the data is not truly traded with third parties and would not be defined as a relevant market under traditional competition law standards. Tucker and Wellford defend otherwise by reasoning that data could only constitute a relevant market where it is actually sold to consumers, as the rationale for defining a relevant market for data in the absence of any sales of such data would be quite unclear.[61] However, this paper sides with Graef's forward-looking and dynamic stance towards market definition, since it seems to reflect a more up-to-date antitrust perspective that should be adopted by antitrust authorities in order for a more comprehensive and thorough assessment of competition in data-related markets.

Additionally, another aspect to be considered when assessing market power in data-driven markets is whether there exists the possibility of 'multi-homing' from the consumer side. 'Multi-homing' is the ability of consumers to use several platform providers for the same type of service. To illustrate, users multi-home if they use at the same time WhatsApp and Telegram, both providers of text messaging services. When

[58] Sokol and Comerford (n 48) 1155.
[59] Graef (n 47) 490.
[60] Graef (n 47) 493.
[61] Tucker and Wellford put forward the question of how it would be possible to define a product market if there were no 'product' and no 'market'. See Tucker and Wellford (n 17) 4-5.

there is evidence that users are multi-homing between platforms, this can suggest that the switching costs between service providers are relatively low.[62] However, caution is needed when using the multi-homing factor to assess market power. Competition authorities should not limit their analyses only to *potential* multi-homing.[63] While it is relevant to examine whether WhatsApp users could *potentially* switch to Telegram, it is undoubtfully more significant to consider whether WhatsApp users truly multi-home with the Telegram app at a certain frequency. In practice, the number of users that truly multi-home between WhatsApp and Telegram tends to be low, especially given the existence of 'network effects' and 'feedback loops' that magnetize consumers to use the platform which is also used by most other users. The likelihood that users switch to a platform that is barely used by his or her connections is extremely low, as it would go against the purpose of using a social network or a text messaging service in the first place.

In sum, as seen above, although the existing literature attempts to explore different competition law tools that can be used to analyze data as a factor to establish market power, there are still no comprehensive or fixed set of criteria to be used for the assessment of market power in data-driven markets. Nevertheless, entry barriers, network effects, feedback loops, multi-sided markets and multi-homing are all factors that should be taken into consideration throughout this assessment. The discussion is currently on-going and will be further developed with future key cases, such as Germany's probe over Facebook's alleged abuse of market power, reviewed in more details in Part 4 of the present paper.[64]

2.2.2. Mergers and Acquisitions to Gain Better Access to Data

A second legal challenge that Big Data can trigger from a competition law standpoint involves the area of mergers and acquisitions. To put it simply, since the

[62] Rubinfeld and Gal (n 10) 358; Tucker and Wellford (n 17) 3-4.
[63] Autorité and Bundeskartellamt (n 3) 29.
[64] Bundeskartellamt, 'Background information on the Facebook proceeding' (*Bundeskartellamt*, 19 December 2017) <https://www.bundeskartellamt.de/SharedDocs/Publikation/EN/Diskussions_Hintergrundpapiere/2017/Hintergrundpapier_Facebook.pdf?__blob=publicationFile&v=6> accessed 22 May 2018.

value of the data depends on the other 3 V's – *volume*, *variety* and *velocity* –, companies are aiming to acquire and sustain a data-related competitive advantage by merging with or acquiring other companies that previously owned large datasets.[65] By doing so, companies gain access to a greater *amount* of data that possibly carries more *diverse* information with which companies can extract *value*.[66]

When an incumbent company merges with an entrant company in markets that are not data-driven, there are normally no competition law concerns due to the fact that the entrant company has a low market share or that there are no horizontal overlaps between the activities of both companies.[67] Nonetheless, in data-driven markets the conclusions drawn can be different. Even if one of the companies has a low market share or there is no horizontal overlap, the merger can still give rise to a significant data-related competitive advantage to the post-merger company, which will have access to different sets of data and will have the possibility to combine them together to make new correlation patterns.[68] Such an increase in the concentration of data in the hands of fewer players could potentially raise competition concerns if the information extracted from the data fusion is so unique that it is impossible to be replicated by competitors.[69]

Moreover, it is not unusual for merging parties to raise data-driven efficiencies as a defense to justify why their potentially anticompetitive merger should be approved by competition authorities.[70] For instance, in *Microsoft/Yahoo!* the merging parties presented efficiency claims and attempted to argue that the *scale* of data resulting from the deal would allow them to produce better products, to be a more credible alternative to Google in the search advertising market, and to provide greater value for users as well as advertisers.[71] Likewise, in *TomTom/TeleAtlas* the merging parties argued that the deal would generate efficiencies by enabling the merged firm to produce better maps in a shorter period of time, especially due to the integration of TomTom's and

[65] Sokol and Comerford (n 48) 1145; Grunes and Stucke (n 10) 3.
[66] See above in Part 2.2.1 for the linkage between volume, variety and value of data.
[67] For an in-depth explanation of the substantive analysis carried out in a merger case, see Whish and Bailey (n 88) 209.
[68] Autorité and Bundeskartellamt (n 3) 16.
[69] Rubinfeld and Gal (n 10) 350-351.
[70] Grunes and Stucke (n 10) 3-4; Stucke and Grunes (n 19) 3; Autorité and Bundeskartellamt (n 3) 17.
[71] *Microsoft/Yahoo! Search Business* (Case COMP/M.5727) Commission Decision C(2010) 1077 paras 160-164.

TeleAtlas's customer feedback data.[72] In the latter case, the European Commission took a skeptical view by stating that although consumers would certainly benefit from the map database updates made possible by the merger, such efficiencies were 'difficult to quantify' and the estimates and calculations provided by the parties were 'not particularly convincing'.[73] Even so, the Commission did not attempt to estimate the magnitude of possible data-related efficiencies in *TomTom/TeleAtlas* as it considered the merger pro-competitive regardless of any efficiencies.

It is likely that similar arguments based on data-driven efficiencies arise again in future data-driven mergers that have the potential of being anticompetitive. Thus, competition authorities around the globe should be ready to meticulously analyze any claims made in this direction to secure that consumers are in fact benefitting from any claimed product improvements. If antitrust watchdogs simply start accepting blanket claims that data will enhance products and bring along other efficiencies, there is a high risk for the welfare of consumers as well as for the competitiveness of markets.

Taking this into account, critics such as Grunes and Stucke encourage competition authorities to reexamine previous data-driven mergers in order to understand whether the analytical tools used by them were adequate to scrutinize these mergers.[74] The authors support the importance of observing *ex post* whether certain mergers enabled companies to entrench or increase their market power, to impede others from entering the market, or to combine data in a privacy-unfriendly way.[75] Indeed, it is necessary for antitrust watchdogs to carry out retrospective studies of merger decisions that involved Big Data companies, as these studies have the potential of determining if their analytical tools are still suitable for the assessment of data-driven mergers. In this regard, Fidelis and Ortaç defend the need to incorporate a more dynamic approach into merger analysis of data-driven deals by placing weight not only

[72] *TomTom/TeleAtlas* (Case COMP/M.4854) Commission Decision C(2008) 1859 paras 245-250.
[73] *TomTom/TeleAtlas* (n 72) para 248.
[74] Grunes and Stucke (n 10) 9-10.
[75] Grunes and Stucke (n 10) 9-10.

on market share and concentration but also on potential competition, innovation process, and post-merger changes in the firm's incentives and behavior.[76]

Apart from reviewing their substantive merger control assessment tools, some competition authorities such as the German Bundeskartellamt also took the lead to update their merger control notification thresholds. The German reform – which entered into force in June 2017 – aimed to adapt the country's competition framework to the legal challenges of the digital economy, including innovation-driven and high-tech merger deals which were not caught by the previous regime.[77] The amendment was supposedly triggered by the *Facebook/WhatsApp* merger, in which neither German nor EU merger control notification thresholds were met since WhatsApp had a low global turnover at the time.[78] After the *Facebook/WhatsApp* merger, German legislators decided to introduce an additional 'size-of-transaction threshold', which was designed to capture mergers that may significantly harm competition despite the low turnover figures of the target company.[79] This new 'size-of-transaction' threshold makes the acquisition of target companies with low turnover but large business potential subject to merger control in Germany.[80] Following the German lead, Austria also introduced a deal-value threshold in November 2017.[81]

A similar reform is also being discussed at EU level. Fearing the risk of under-enforcement, in 2016 the European Commission invited interested shareholders to respond to a public consultation aimed at evaluating the effectiveness of the turnover-

[76] Fidelis and Ortaç also recognize that even if a merger does not raise competitive concerns from a 'static competition' perspective (for instance if post-merger there is no price increase or if the concentration in the relevant product market is not higher), the merger can however harm consumer welfare from a long-term perspective (e.g. less product innovation). See Andressa Lins Fidelis and Zeynep Ortaç, 'Data-Driven Mergers: A Call For Further Integration Of Dynamics Effects Into Competition Analysis' [2017] Barcelona Graduate School of Economics <https://repositori.upf.edu/bitstream/handle/10230/33467/FidelisOrtac%20TFM2017.pdf?sequence=1&isAllowed=y> accessed 29 July 2018.

[77] Silvio Cappellari and Stephanie Birmanns, 'Germany: Merger Control' (*Global Competition Review*, 14 August 2017) <https://globalcompetitionreview.com/insight/the-european-middle-eastern-and-african-antitrust-review-2018/1145587/germany-merger-control> accessed 16 May 2018.

[78] Cappellari and Birmanns (n 77).

[79] Werner Berg and Lisa Weinert, 'New Merger Control Threshold in Germany – Beware of Ongoing Transactions' (*Kluwer Competition Law Blog*, 7 June 2017) <http://competitionlawblog.kluwercompetitionlaw.com/2017/06/07/new-merger-control-threshold-germany-beware-ongoing-transactions/> accessed 16 May 2018.

[80] Cappellari and Birmanns (n 77).

[81] Michael Mayr, 'Austria to introduce Transaction Value Merger Notification Threshold' (*Kluwer Competition Law Blog*, 10 April 2017) <http://competitionlawblog.kluwercompetitionlaw.com/2017/04/10/austria-to-introduce-transaction-value-merger-notification-threshold/> accessed 22 May 2018.

based jurisdictional thresholds of EU merger control, particularly regarding whether such thresholds allow the EU to capture all transactions which can potentially impact the internal market.[82] Commissioner Vestager expressed in one of her speeches that turnovers are not always what make companies attractive for mergers, given that sometimes what actually matters are certain assets, such as a dataset.[83] In Vestager's view, a data-driven merger 'could clearly affect competition, even though the company's turnover might not be high enough to meet our thresholds'.[84] Indeed, in the digital economy many companies whose business models are based on the Big Data value chain do not yet generate significant turnovers. As a consequence, acquisitions of such companies are likely not captured under the current turnover-based thresholds, even though these companies hold commercially valuable data and considerable market potential. Yet, the results of the public consultation show that the majority of public and private stakeholder do not see any need for introducing complementary thresholds to solve the alleged enforcement gap.[85] Rather, the respondents argued, among other things, that there is insufficient empirical evidence for an 'enforcement gap' and that there already exists mechanisms that make the merger notification system more flexible (e.g. the referral system of Articles 4(4), 4(5) and 22 of the EU Merger Regulation).[86] Up to the date of publication, there have been no measures adopted by the Commission in relation to this matter.

To sum up, taking into account that the relevance of mergers and acquisitions is likely to grow with the expansion of the digital economy and that companies will attempt

[82] Note that currently the EU Merger Regulation applies only to concentrations of a Union dimension. A concentration is considered to have a Union dimension when the turnover thresholds of Article 1 of the Merger Regulation are met. Council Regulation (EC) No 139/2004 of 20 January 2004 on the control of concentrations between undertakings (the EC Merger Regulation) OJ L 24/1; European Commission, 'Consultation on Evaluation of procedural and jurisdictional aspects of EU merger control' (*European Commission Public Consultations*) <http://ec.europa.eu/competition/consultations/2016_merger_control/index_en.html> accessed 26 July 2018.
[83] Margrethe Vestager, "Refining the EU merger control system" (n 2).
[84] Margrethe Vestager, "Refining the EU merger control system" (n 2).
[85] European Commission, 'Summary of replies to the Public Consultation on Evaluation of procedural and jurisdictional aspects of EU merger control' <http://ec.europa.eu/competition/consultations/2016_merger_control/summary_of_replies_en.pdf> accessed 27 July 2018.
[86] EC Merger Regulation (n 82); European Commission, 'Summary of replies to the Public Consultation on Evaluation of procedural and jurisdictional aspects of EU merger control' (n 85). See also Davilla's eight points of criticism towards the introduction of a new 'size-of-transaction' merger threshold at the EU level. Marixenia Davilla, 'Is Big Data a Different Kind of Animal? The Treatment of Big Data Under the EU Competition Rules' (2017) 8:6 Journal of European Competition Law & Practice 377-379.

to sustain a data-related competitive advantage over their rivals by merging with other companies that own valuable datasets, competition authorities need to ensure that (i) they have the adequate analytical tools to scrutinize data-driven mergers and that (ii) their notification thresholds are able to capture these types of deals.

2.2.3. Exclusionary Conducts to Hinder Competitors' Access to Data

A third potential legal challenge that competition law can encounter in the era of Big Data is the use of exclusionary conducts by companies who wish to maintain their data-driven competitive advantage over their rivals.[87] In the words of Professor Richard Whish, an exclusionary conduct is a 'behavior by a dominant firm designed to, or which might have the effect of, preventing the development of competition'.[88] Such conducts are often considered an abuse of dominant position under Article 102 of the Treaty on the Functioning of the European Union (TFEU).[89]

When companies make significant investments in the Big Data value chain to *collect*, *store*, and *analyze* data, most likely they will have strong incentives to undertake data-driven – and possibly anticompetitive – strategies.[90] As a result, businesses can be tempted to implement certain exclusionary conducts, including (i) the refusal to access data, (ii) a discriminatory access to data, (iii) exclusive agreements and (iv) tied sales and cross-usage of datasets.[91]

A *refusal to access data* may give rise to anticompetitive concerns if an incumbent company (hereinafter Company A) who owns a dataset that is considered an 'essential facility' to the activity of another undertaking (hereinafter Company B) refuses to share the data with the latter company. The crux of the question is whether Company B could rely on Article 102 TFEU to gain access to the large quantity of data that

[87] Grunes and Stucke (n 10) 3, 10; Stucke and Grunes (n 19) 3; Autorité and Bundeskartellamt (n 3) 17-20.
[88] Whish and Bailey (n 56) 212.
[89] Whish and Bailey (n 56) 214; Consolidated Version of the Treaty on the Functioning of the European Union [2012] OJ C 326/47.
[90] Grunes and Stucke (n 10) 3; Stucke and Grunes (n 19) 3.
[91] Autorité and Bundeskartellamt (n 3) 17-20.

Company A owns.[92] EU courts have developed a rigorous view over the past several years under which dominant firms, only under *very specific circumstances*, may be in breach of Article 102 TFEU if they refuse to give access to an 'essential input'. The conditions on compulsory access to essential inputs set forth by EU courts are quite strict and, as a consequence, Company B would have a hard time in seeking to rely on Article 102 TFEU to obtain access to Company A's dataset.[93]

First, Company A would need to have a proven dominant position on the market for such essential facility – namely, the dataset. Note that it may be rather difficult and complex to define a market for data, especially if such data is not traded on the market.[94] Second, in order for Company B to carry out its business in a competitive way, it would need to have access to Company A's dataset, which would need to be considered 'indispensable' or 'essential'. In the *Bronner* case, the EU courts decided that a product or service is 'indispensable' if certain conditions are met: (i) there are no alternative products or services (in this case, no alternative datasets); and (ii) there are technical, legal or economic obstacles capable of making it impossible – or unreasonably difficult – for any other undertaking aiming to operate on the downstream market (in this case, Company B) to develop alternative products or services.[95] Moreover, in order to accept the existence of economic obstacles, Company B would have to establish that the creation of such dataset is not economically viable for production on a scale comparable to that of Company A.[96] Third, Company A's refusal of access to its dataset must be likely to prevent any competition on the downstream

[92] Geradin and Kuschewsky (n 31) 13-14.
[93] For a further approach on the question of whether a dataset could be considered an 'essential facility' within the meaning of Article 102 TFEU, see Geradin and Kuschewsky (n 31) 13-15.
[94] See above in Part 2.2.1 for the debate about the difficulties in defining a relevant market for data. However, it is also important to keep in mind that in the *IMS Health* case the ECJ considered that 'it is sufficient that a potential market or even a hypothetical market can be identified'. See C-418/01 *IMS Health* [2004] ECJ I-05039, para 44.
[95] Case C-7/97 *Bronner v Mediaprint* [1998] ECJ I-07791, paras 43-46. In the *Bronner* case, the matter under dispute was whether a press undertaking (Mediaprint) which held a very large share of the daily newspaper market and operated the only nationwide newspaper home-delivery scheme was abusing its dominant position under Article 102 TFEU if it refused to allow the publisher of a rival newspaper (Bronner) – which by reason of its small circulation was unable either alone or in cooperation with other publishers to set up and operate its own home-delivery scheme in economically reasonable conditions – to have access to that home-delivery scheme for appropriate remuneration. *In casu*, the European Court of Justice ruled that under those particular circumstances there was no abuse of dominant position under Article 102 TFEU. See also Case C-418/01 *IMS Health* [2004] ECJ I-05039, para 28, in which the discussion about 'essential facilities' arose once again.
[96] Case C-7/97 *Bronner v Mediaprint* [1998] ECJ I-07791, para 46.

market.[97] As seen from the three conditions set out above, Company B holds a somewhat challenging burden of proving that the dataset owned by Company A is truly unique and that there are no possibilities whatsoever to obtain such dataset elsewhere. This line of argumentation can be even more difficult to uphold given the non-rivalrous nature of data.[98] Therefore, the likelihood that data-driven companies succeed with this argument of 'essential facilities' in future cases is relatively low.

Another exclusionary conduct that can be considered detrimental to competition is a situation of *discriminatory access to data*.[99] To illustrate, a situation of discriminatory access to data can occur when an incumbent company refuses to share its dataset with one company but is willing to sell the dataset to other companies. This type of conduct would most likely fall under subparagraph (c) of Article 102 TFEU, according to which an abuse may consist in 'applying dissimilar conditions to equivalent transactions with other trading parties, thereby placing them at a competitive disadvantage'.[100]

Likewise, incumbent companies can also safeguard their data-related competitive advantage by gathering large volumes of data through *exclusive agreements* with third-party providers.[101] Exclusive agreements – also known as exclusive purchasing, single branding, requirement contracts or non-compete obligations – are typical examples of exclusionary abuses caught by Article 102 TFEU.[102] These agreements can block third-party data providers from doing business with anyone else other than the dominant firm, prevent rivals from accessing data, preclude competitors' possibilities of acquiring similar datasets, and foreclose competition in upstream or downstream markets.[103]

---

[97] Geradin and Kuschewsky (n 31) 15.
[98] According to Graef (n 47) 479, 'data is a so-called non-rivalrous good which means that the fact that a certain entity has collected a piece of data does not preclude others from gathering identical information'. Likewise, Rubinfeld and Gal (n 10) 369 state that 'data is nonrivalrous, and collecting it does not prevent others from collecting identical data by comparable or different means'. Similarly, Tucker and Wellford (n 17) 3 assert that 'big data is non-rivalrous. In other words, collecting a particular piece of data does not prevent other companies from collecting identical data by similar or other means'. See also Sokol and Comerford (n 48) 1137, 'Data is non-exclusive and non-rivalrous. No one firm can, or does, control all of the world's data. Collection of a piece of data by one firm does not occur at the expense of another firm.'
[99] Autorité and Bundeskartellamt (n 3) 18-19.
[100] TFEU (n 89) article 102.
[101] Geradin and Kuschewsky (n 31) 7-9 examined in their paper whether the acquisition of personal data through exclusivity agreements may breach EU competition law.
[102] Whish and Bailey (n 56) 221, 723.
[103] Geradin and Kuschewsky (n 31) 8; Grunes and Stucke (n 10) p. 3; Stucke and Grunes (n 19) 3.

An example of how exclusive agreements can come into play in a data-related situation is the *Google Search (AdSense)* case.[104] In July 2016, the European Commission decided to initiate antitrust proceedings against Google due to its preliminary view that 'the company has abused its dominant position by artificially restricting the possibility of third party websites to display search advertisements from Google's competitors'.[105] The Commission's concern is that Google was able to protect its dominant position in online search advertising by entering into exclusive agreements with third-parties – the so-called 'Direct Partners', which were required not to source search ads from Google's rivals.[106] Moreover, Google supposedly required third-parties to take a minimum number of search ads from Google, to reserve the most prominent space on their search results pages to Google search ads, and to refrain from placing competing search ads above or next to Google search ads. Lastly, Google purportedly required third-parties to obtain Google's approval before making any changes to the display of competing search ads. By doing so, Google allegedly prevented existing and potential competitors from entering the market, reduced choice in an artificial way, and stifled innovation in the particular market.[107] Up to the date of publication, the proceedings were still on-going, and no decisions had been taken by the Commission. It is relevant to mention, however, that during the course of the proceedings Google changed the conditions of AdSense contracts with its Direct Partners, giving them more freedom to display competing search ads.

---

[104] *Google Search (AdSense)* (Case AT.40411).

[105] European Commission, 'Antitrust: Commission takes further steps in investigations alleging Google's comparison shopping and advertising-related practices breach EU rules*' (*European Commission Press Release Database*, 14 July 2016) <http://europa.eu/rapid/press-release_IP-16-2532_en.htm> accessed 29 July 2018.

[106] See an explanation of how Google AdSense works in European Commission, 'Antitrust: Commission takes further steps in investigations alleging Google's comparison shopping and advertising-related practices breach EU rules*' (*European Commission Press Release Database*, 14 July 2016) <http://europa.eu/rapid/press-release_IP-16-2532_en.htm> accessed 29 July 2018: 'Google places search ads directly on the Google search website but also as an intermediary on third party websites through its "AdSense for Search" platform ("search advertising intermediation"). These include websites of online retailers, telecoms operators and newspapers. The websites offer a search box that allows users to search for information. Whenever a user enters a search query, in addition to the search results, also search ads are displayed. If the user clicks on the search ad, both Google and the third party receive a commission. A large proportion of Google's revenues from search advertising intermediation stems from its agreements with a limited number of large third parties, so-called "Direct Partners". The Commission has concerns that in these agreements with Direct Partners, Google has breached EU antitrust rules (…).'

[107] For more information on the effects of exclusive agreements, see Communication from the Commission – Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings [2009] OJ C 45/7, paras 32 et seq.

Lastly, if a dominant company makes an arrangement under which it only sells the access to its dataset on the condition that the buyer also purchases another one of its products or services (*tying*) or if a company uses the data it has collected about a certain market on adjacent markets (*cross-usage of data*), this could potentially reduce competition and enhance the given company's competitive advantage over its rivals.[108] To illustrate, think of the following data-related tying situation. A large firm has a strong market power due to the creation of a valuable dataset. If this firm attempts to enter the data analytics market by tying the purchase of its dataset with the use of its data analytics service, then the firm is engaging in a typical tying arrangement.[109] Article 102(d) TFEU specifically refers to tie-in agreements as a possible abuse which may consist in 'making the conclusion of contracts subject to acceptance by the other parties of supplementary obligations which, by their nature or according to commercial usage, have no connection with the subject of such contracts'.[110] Even so, to be considered an infringement, a series of questions would have to be answered beforehand, such as if the undertaking has a dominant position, if it is tying two distinct products, if the customer was coerced to buy both products, if the tie could be detrimental to competition by foreclosing access to the market and if there is an objective justification for the tie.[111]

A recent example of a tying conduct in a Big Data context is the European Commission's *Google Android* case.[112] The Commission had been closely investigating Google's conduct regarding its Android operating system since early 2015. In July 2018, the Commission decided to fine Google €4.34 billion for the practice of imposing three different types of illegal restrictions on Android device manufacturers and mobile network operators in order to strengthen its dominant position on the search engine

[108] Note that this type of conduct also has implications and limitations under the current GDPR framework, which will be further explored in Part 3 of the present paper. Communication from the Commission – Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings [2009] OJ C 45/7, para 48; Autorité and Bundeskartellamt (n 3) 20.
[109] Competition and Markets Authority, *The Commercial Use of Consumer Data: Report on the CMA's Call for Information* (June 2015) 90.
[110] TFEU (n 89).
[111] Whish and Bailey (n 56) 732.
[112] *Google Android* (Case AT.40099).

market.[113] One of these contractual restrictions was the conduct of tying the Google Search app and the Google Chrome browser to the Google Play Store. While users expect to find the Google Play Store pre-installed on their devices, the tying of such product to the Google Search app and to the Google Chrome browser resulted in the pre-installation of all three products as a bundle. Thus, it was impossible for manufacturers to pre-install certain apps (e.g. Google Play Store) without having to install other apps (e.g. Google Search). The Commission found that such pre-installation led to a '*status quo* bias', meaning that users were more likely to stick to apps that were pre-installed on their devices than to switch to apps that still had to be downloaded.[114] As a consequence of Google's illegal tying, rival companies had less chances of being able to effectively compete with Google on the merits and consumers had fewer options of browsers and search engine apps to choose from.

In sum, the above-mentioned conducts of (i) refusal to access data, (ii) discriminatory access to data, (iii) exclusive agreements and (iv) tied sales are exclusionary because their end effect is to limit competitors' access to data, to prevent others from sharing data and to impede rivals from achieving the minimum efficient scale to be able to compete in the market.[115]

2.2.4. Price Discrimination Between Different Customer Groups

A fourth potential antitrust challenge triggered by Big Data is the use of data for price discrimination among consumers based on their revealed preferences.[116] Price discrimination occurs either when a company charges its consumers different prices for the same products even though there are no differences in the costs of supplying them *or* when a company charges identical prices even though there are enough cost

[113] European Commission, 'Antitrust: Commission fines Google €4.34 billion for illegal practices regarding Android mobile devices to strengthen dominance of Google's search engine' (*European Commission Press Release Database*, 18 July 2018) <http://europa.eu/rapid/press-release_IP-18-4581_en.htm> accessed 23 July 2018.
[114] European Commission, 'Antitrust: Commission fines Google €4.34 billion for illegal practices regarding Android mobile devices to strengthen dominance of Google's search engine' (*European Commission Press Release Database*, 18 July 2018) <http://europa.eu/rapid/press-release_IP-18-4581_en.htm> accessed 23 July 2018.
[115] Grunes and Stucke (n 10) 3; Stucke and Grunes (n 19) 3.
[116] Monopolkomission, *Competition Policy: The Challenge of Digital Markets* (Special Report No 68, 2015) para 80; Rubinfeld and Gal (n 10) 342; Autorité and Bundeskartellamt (n 3) 21; Bourreau, Streel and Graef (n 10) 8.

differences that would justify their differentiation.[117] In the words of Bourreau, Streel and Graef, 'a firm price discriminates to extract as much as possible what the consumers are willing to pay for its products or services'.[118]

In this sense, when a Big Data company collects enough information about its consumers' purchasing habits and willingness to pay, it can infer which consumers are willing to pay a higher price for a given product or service and which consumers are not.[119] As a consequence, considering that the company has market power, it could then easily set individual prices and price discriminate based on its estimation of each consumer's willingness to pay.[120] Although the phenomenon of price differentiation is not new and in fact often constitutes a vital element of a company's revenue management (e.g. when companies charge different prices depending on the time of the day), what changed with Big Data is the possibility to set personalized prices on the internet based on the *characteristics* and *habits* of each consumer and their respective willingness to pay for a certain product or service.[121]

This particular use of Big Data to individualize products and prices can raise several social-welfare questions.[122] On one side, it may increase general welfare in economic terms by serving customer groups that would not have purchased a certain product or service in the absence of such price differentiation.[123] Indeed, price

[117] Whish and Bailey (n 56) 802.
[118] Bourreau, Streel and Graef (n 10) 39.
[119] Although it is out of the scope of research of the present paper, it is relevant to point out that the use of Big Data analytics and algorithms to price discriminate may also raise data protection concerns under Article 22(1) of the General Data Protection Regulation. In particular, the GDPR provides data subjects with the right not to be subject to a decision based solely on automated processing (e.g. algorithms) which produces legal effects concerning the data subject or similarly 'significantly affects' the data subject. According to Malgieri and Comandé, the envisaged 'significant effects' can encompass, for instance, price discrimination. For a further analysis regarding automated decision-making and the GDPR, see Gianclaudio Malgieri and Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) 7:4 International Data Privacy Law 243.
[120] Autorité and Bundeskartellamt (n 3) 21; Bourreau, Streel and Graef (n 10) 40.
[121] Monopolkomission (n 116) para 80. However, Bourreau, Streel and Graef (n 10) 41 sustain that although a personalized pricing strategy would not be too difficult to be implemented, it is nevertheless rarely observed in practice, as perhaps companies fear a negative consumer reaction. The authors maintain that there are cleverer ways for companies to achieve the same outcome. Companies could, for instance, employ indirect methods of setting personalized prices, such as (i) displaying the same uniform price to all consumers but offering personalized discounts only to part of them or (ii) steering searches and presenting different products to consumers from different groups.
[122] Rubinfeld and Gal (n 10) 348.
[123] Monopolkomission (n 116) para 80.

discrimination may improve social welfare by increasing the total number of transactions when comparing a uniform price scenario with a discriminated price scenario.[124] However, it may be the case that such price discrimination places at a disadvantage some consumers who are less aware of the ways in which their data is being utilized, which potentially also raises consumer and data protection law concerns.[125] While some authors sustain that price discrimination is not necessarily detrimental to social welfare or to consumer surplus and can even increase them vis-à-vis uniform pricing, others support the precautious view that an increase in social welfare does not automatically reflect an increase in consumer surplus.[126]

In any case, from an antitrust law standpoint, it is relevant to notice that price discrimination is not *per se* an exclusionary abuse under Article 102(c) TFEU.[127] Similarly to the tie-in situation explained above in Part 2.2.3, a situation of price discrimination is only considered an infringement to Article 102 TFEU if the undertaking has a dominant position, if it entered into equivalent transactions with other companies but applied dissimilar conditions, if this discrimination caused a competitive disadvantage, and if there is no objective justification for the discrimination.[128]

2.2.5. Market Transparency and Increased Risk of Collusion

Last but not least, a fifth competition law challenge that can arise in the era of Big Data is regarding market transparency. Big Data can lead to more transparent online markets, in the sense that players can easily have access to their competitors' prices, products, quality standards and, sometimes, business tactics. More market transparency in digital markets brings along both advantageous and disadvantageous consequences.

[124] Whish and Bailey (n 56) 803; Autorité and Bundeskartellamt (n 3) 21.
[125] For instance, a scenario of welfare loss would occur with the introduction of data-based insurance tariffs. In this regard, see Monopolkomission (n 116) 30-31.
[126] For the former viewpoint, see Bourreau, Streel and Graef (n 10) 8. For the latter viewpoint, see Autorité and Bundeskartellamt (n 3) 21 and Monopolkomission (n 116) 30-31.
[127] Whish and Bailey (n 56) 804.
[128] Whish and Bailey (n 56) 804.

By looking at one side of the coin, one can infer that the more transparent the market is, the more consumers can compare prices and characteristics of different goods or services.[129] For instance, platforms like Amazon, eBay and TripAdvisor allow online merchants to announce their products and consumers to compare the different options available, including their prices, characteristics, user reviews and ratings. Thus, a greater market transparency allows consumers to make informed and knowledgeable choices about their purchases.[130]

Yet, the other side of the coin shows that the more transparent a digital market is, the greater the availability of information about competitors' pricing and the higher the chances of having a more stable tacit or explicit collusion between different players (or between algorithms).[131] In other words, a more transparent online market can enhance the ability of firms to easily monitor how other economic operators are behaving and can increase the probability of detection if one is deviating from the common conduct, therefore contributing to the maintenance of a collusion between competitors.[132] As a result, a greater market transparency in a Big Data world may actually have the effect of increasing prices for consumers, given that the likelihood of collusion between firms increases.[133]

3. Big Data and Data Protection

3.1. The Right to Privacy, to Data Protection and the Novel EU Framework

In addition to the competition law challenges explained in Part 2, the fact that businesses in the era of Big Data are now able to rapidly collect, store and analyze more

[129] Autorité and Bundeskartellamt (n 3) 14.
[130] Autorité and Bundeskartellamt (n 3) 14.
[131] Tucker and Wellford (n 17) 11; Autorité and Bundeskartellamt (n 3) 14-15. See also OECD, *Algorithms and Collusion: Competition Policy in the Digital Age* (2017) for a detailed assessment regarding algorithms, the risk of collusion and the challenges for competition law enforcement: '(…) algorithms may work as a facilitating factor for collusion and may enable new forms of co-ordination that were not observed or even possible before. This is referred to as "algorithmic collusion".'
[132] Whish and Bailey (n 56) 598; Autorité and Bundeskartellamt (n 3) 14-15.
[133] Tucker and Wellford (n 17) 11.

data about its customers – in particular *personal data* – also raises several data protection concerns. Before exploring such concerns, it is crucial to grasp the background underpinning the rights to privacy and to data protection.

The *right to privacy*, recognized worldwide as a human and fundamental right, protects the private life of individuals and limits the access that others have to an individual's personal sphere, including their private and family life, their home and their correspondence.[134] It is enshrined in Article 12 of the Universal Declaration of Human Rights, Article 17 of the International Covenant on Civil and Political Rights, Article 7 of the EU Charter of Fundamental Rights and Article 8 of the European Convention on Human Rights.[135] Apart from international and regional treaties, most countries in the world also recognize the right to privacy in their constitutions or in other relevant laws.[136]

The *right to data protection* aims to protect individuals' fundamental rights and freedoms and safeguard their right to the protection of their *personal data*.[137] Article 16(1) TFEU provides that everyone has the right to the protection of personal data concerning him or her.[138] Moreover, Article 8(1) of the EU Charter of Fundamental Rights explicitly contains the right to protection of personal data, which 'must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law'.[139] The right to data protection is 'intrinsically linked' to the right to privacy.[140] Yet, although these two rights

[134] Oostveen (n 23) 302.
[135] European Data Protection Supervisor, 'Data Protection' (*European Data Protection Supervisor*) <https://edps.europa.eu/data-protection/data-protection_en> accessed 26 May 2018.
[136] European Data Protection Supervisor, 'Data Protection' (n 135).
[137] General Data Protection Regulation (n 5), art 1(2).
[138] TFEU (n 89).
[139] Charter of Fundamental Rights of the European Union [2000] OJ C 364/1.
[140] According to Brkan, the distinction between the right to privacy and the right to data protection is far from clear in the case law from the Court of Justice of the European Union (CJEU). Brkan distinguished the case law into three different categories: (i) cases in which the CJEU made a relatively clear distinction between the right to privacy and the right to data protection; (ii) cases in which the distinction between these two rights is rather blurred; and (iii) cases that treat the right to data protection as a sub-category of the right to privacy. Brkan argues that while it is undeniable that both rights are 'intrinsically linked', it is yet to be seen whether the CJEU will distinguish them in future cases, as until now the case law lacks consistency in this regard. For a thorough analysis of the distinction between these two rights in the CJEU's case law, see Brkan M, 'The Court of Justice of the EU, Privacy and Data Protection: Judge-Made Law as a Leitmotif in Fundamental Rights Protection' in Maja Brkan and Evangelia Psychogiopoulou (eds) *Courts, Privacy and Data Protection in the Digital Environment* (Edward Elgar Publishing, 2017) 13-17.

are closely related, their relationship at a EU level is not clear-cut.[141] Some authors such as McDermott argue that the right to data protection contains enough distinct elements that justify its framing as a separate right.[142] In the same line, Lynskey suggests that even with the significant overlaps between them, the two legal regimes are distinct and the protection offered by data protection legislation is broader than that offered by privacy rules insofar as it provides individuals with an enhanced control over personal data.[143]

Prior to May 2018, the main piece of EU legislation in the area of data protection was the Data Protection Directive of 1995. As every Directive, it only laid down the results that had to be achieved by member states but left each member state with discretion to decide on how to transpose the Directive into their national laws. Thus, until recently data protection legislation across the EU was not fully harmonized. Nonetheless, growing calls for an increased legal certainty and greater protection in the processing of data spurred the EU to begin a reform of its data protection rules in 2012. The reform culminated in a novel EU data protection framework consisting of (i) the GDPR, which entered into force on 25 May 2018, and (ii) the Directive on data protection, applicable since 06 May 2018.[144] The new framework set forth a uniform interpretation with a binding legal force throughout all member states.

3.2. Data Protection Challenges in the Era of Big Data

Despite the numerous benefits that Big Data can generate – saving resources and costs, optimizing processes, decreasing the risks associated with a decision, increasing sales, preventing fraud and discovering unexpected correlations –, there are also many implications that can pose a risk to individuals' right to data protection.[145] In the past

[141] For a comprehensive explanation of the link between data protection and privacy in the EU legal order, see Orla Lynskey, *The Foundations of EU Data Protection Law* (Oxford University Press 2015) 89-130. See also Brkan (n 140).
[142] McDermott (n 7) 2; European Data Protection Supervisor, 'Data Protection' (n 135).
[143] Lynskey (n 141) 129, 265.
[144] General Data Protection Regulation (n 5); Directive (EU) 2016/680 (n 5).
[145] For more illustrations of the opportunities and benefits of Big Data, see Dennis Broeders, Erik Schrijvers, Bart van der Sloot, Rosamunde van Brakel, Josta de Hoog and Ernst Ballin, 'Big Data and Security Policies: Towards a

years, consumers have become increasingly bothered with their sense of powerlessness and lack of control over who has access to their personal data, which of their data is being collected and sold, and how and when such data is being used.[146] Critics have emphasized the rather negative opinion that market forces are currently far from solving privacy issues without external intervention and that the traditional 'notice-and-consent' model is ineffective and inadequate to safeguard the privacy of individuals.[147] Although the GDPR framework attempted to modernize EU data protection legislation and boost the rights of individuals to better control their personal information, it still faces many limitations when it comes to the Big Data value chain. These challenges will be discussed in the subsections below.

3.2.1. Partial Applicability of the GDPR's Material Scope

The first potential data protection challenge in the era of Big Data is regarding the partial applicability of the GDPR's material scope. The GDPR framework is applicable to 'the processing of personal data wholly or partly by automated means'.[148] It is also applicable to 'the processing other than by automated means', but this part of the GDPR's material scope is not relevant for the Big Data value chain, as it is almost impossible to process Big Data with non-automated means.[149] In a nutshell, the material scope of EU data protection depends on two main elements: (i) processing and (ii) personal data.

The first element – *processing* – is given a broad definition by the GDPR and involves any operation or set of operations performed on personal data, including the

---

Framework for Regulating the Phases of Analytics and Use of Big Data' (2017) 33 Computer Law & Security Review 310. See also Oostveen (n 23) 302.

[146] Stucke and Grunes (n 19) 6; Stucke and Grunes (n 10) 326-327.

[147] Stucke and Grunes (n 19) 6; Stucke and Grunes (n 10) 326.

[148] General Data Protection Regulation (n 5), art 2(1).

[149] As explained in Part 2 of the present paper, one of the main characteristics of Big Data is *velocity*. The processing of data other than by automated means is not compatible with the notion of Big Data, as it would hinder the possibility of firms to collect, store and analyze data with sufficient velocity for it to be valuable. See General Data Protection Regulation (n 5), art 2(1): the GDPR is also applicable to 'to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system'. See also Oostveen (n 23) 304.

collection, storage, usage, disclosure, alteration and erasure of personal data.[150] Since the concept of processing comprises several types of operations, the threshold is easily met during any of the stages of the Big Data value chain. To put it simply, when companies involved in Big Data projects firstly *collect* data, then *store* data, and subsequently *analyze* data, they are most likely considered to be 'processing' data, thereby fulfilling the first criterion of the GDPR's material scope.

The second element – *personal data* – is given a narrower definition which is less easily met. According to the GDPR, personal data means 'any information relating to an identified or identifiable natural person'.[151] The term 'any information' is given a wide interpretation – it can include objective information (e.g. the presence of a substance in a person's blood), subjective information (e.g. opinions expressed by a person), and information in any sort of content or format.[152] Moreover, the term 'relating to' means that the information needs to be *related to* an identified or identifiable individual, in the sense that it needs to be *about* that individual, be it content-wise, purpose-wise or result-wise.[153]

Additionally, data is only considered personal if the *identifiability* element is present. The identifiability criterion serves as a differentiator and separates data that is subject to the GDPR from data that is not.[154] A natural person is considered 'identified' when he or she can be distinguished from all other persons and recognizable as an

[150] Bourreau, Streel and Graef (n 10) 23; General Data Protection Regulation (n 5), art 4(2). Processing is defined as 'any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction'.

[151] General Data Protection Regulation (n 5), art 4(1).

[152] For a breakdown of the constitutive elements of 'personal data', see Article 29 Data Protection Working Party, *Opinion 4/2007 on the concept of personal data* (20 June 2007) <https://www.clinicalstudydatarequest.com/Documents/Privacy-European-guidance.pdf> accessed 03 August 2018 6-24.

[153] To illustrate, Article 29 Data Protection Working Party gave the following examples of data *relating to* an individual: (i) data on the results of a patient's medical test contained in his medical records; (ii) data registered in an employee's individual file in the personnel office; and (iii) data about the value of a particular house, which is then used to determine the extent of a person's obligation to pay certain taxes. See Article 29 Data Protection Working Party, *Opinion 4/2007 on the concept of personal data* (n 152) 9-11.

[154] For a comprehensive analysis on the criterion of 'identifiability' under EU data protection law, see Worku Gedefa Urgessa, 'The Protective Capacity of the Criterion of 'Identifiability' under EU Data Protection Law' (2016) 4 European Data Protection Law Review 521.

individual.[155] On the other hand, an 'identifiable' natural person is one who has not yet been identified but even so can be identified directly or indirectly by reference to names, identification numbers, location data, online identifiers or factors specific to the 'physical, physiological, genetic, mental, economic, cultural or social identity of that natural person'.[156] *Identifiable data* can be either directly or indirectly identifiable, depending on whether an individual is immediately identified through the data (e.g. the name of a person) or whether it is only identifiable through a combination with other auxiliary data (e.g. social security number of a person).[157] Thus, the name of an individual may not be necessary to identify an individual, given that other identifiers that hold a close relationship with such individual can be used to single him or her out.[158]

The identifiability threshold is the most decisive yet most challenging element to be analyzed when demarcating the material scope of the GDPR, especially considering that it is highly context-dependent, meaning that data may be identifiable for one person but non-identifiable for another depending on the particular circumstances.[159] The debate on identifiability also raises the question of whether personal and non-personal data can be distinguished at all. Recital 26 of the GDPR attempted to shed some light on the troubling notion of identifiability by stating that to determine whether a person is identifiable, 'account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly'.[160] Recital 26 makes clear that 'all the means' should be taken into account to determine whether a certain data is identifiable or not.[161] In practice, this broadens the notion of identifiability, since identification can occur through

[155] Article 29 Data Protection Working Party*, Opinion 4/2007 on the concept of personal data* (n 152) 12; Urgessa (n 154) 521.
[156] General Data Protection Regulation (n 5), art 4(1).
[157] Urgessa (n 154) 521; Oostveen (n 23) 305.
[158] Article 29 Data Protection Working Party*, Opinion 4/2007 on the concept of personal data* (n 152) 14; Urgessa (n 154) 521.
[159] Oostveen (n 23) 306 cites the example of a medical journal article about an anonymous patient with a certain disease. While in general the information about the patient is non-identifiable, it can become identifiable if the disease is extremely rare and researchers in the field are able to recognize the identity of the patient.
[160] General Data Protection Regulation (n 5), recital 26.
[161] In the words of Urgessa (n 154) 522, 'the intent behind identification, the way the processing is structured, the advantage expected by the controller, the interests at stake for the individuals, as well as the risk of organizational dysfunctions, the state of the art in technology at the time of the processing and the possibilities for development during the period for which the data will be processed, the duration of processing and other relevant factors should be considered before a data is said to be 'identifiable''.

the combination with any additional knowledge, such as information obtained from another dataset. It is also relevant to notice that *non-identifiable data*, such as anonymous (or de-identified) data and non-personal data, are not within the material scope of the GDPR as long as they remain untraceable back to an individual.[162]

In relation to the identifiability threshold, an interesting debate is put forward by Urgessa. Urgessa uses the examples of profiling and Internet Protocol (IP) addresses[163] as 'data in grey area' – that is, data that may not be entirely classified as 'personal data' but even so continue to single out and target individuals. According to Urgessa, the arguments advanced by Internet companies such as Google against the 'identifying capacity' of data collected for profiling and data found in IP addresses is based on the fact that such data only identifies *machines* and not *data subjects*.[164] As such, these online advertising companies argue that to provide their targeted advertising services it is not necessary for them to know the real-world identify of users, but rather only the virtual identity based on profiles generated from data gathered that relates to them. Yet, skepticism is needed with any argument in this direction, given that nowadays with the abundant amount of data collected from individuals it is rather easy to find auxiliary data that makes such 'data in grey area' in fact identifiable.[165] Moreover, if the identifiability criterion starts having a very strict definition that is hardly met, then this high threshold may fail to qualify as 'personal' data that is indeed capable to single out individuals, going against the criterion's very objective of protecting fundamental rights of data protection. Yet, this controversy is yet to be settled, as the GDPR framework is not entirely clear on the question of online identifiers.[166]

---

[162] Instead of 'anonymous data', Oostveen (n 23) 306 uses the term 'de-identified data' due to the fact that there are several studies contesting the technical possibilities of anonymizing data. Indeed, if an alleged anonymized data is actually retraceable and can backtrack to a certain individual, then it will be considered *indirectly identifiable* data, which falls under the material scope of the GDPR. See also General Data Protection Regulation (n 5), recital 26: 'The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes.'

[163] Urgessa (n 154) 523 defines *profiling* as 'a process of processing and analysing data about individuals in search of patterns, sequences or relationships to generate a profile based on which those individuals will be treated in a certain way' and *IP addresses* as a 'unique string of numbers assigned to every device connected to the Internet for to be recognized for communication purposes'.

[164] Urgessa (n 154) 524-525.

[165] Urgessa (n 154) 524.

[166] The GDPR only recognizes the 'identifying capacity' of IP addresses and other online identifiers when combined with unique identifiers or other information received by the servers. See General Data Protection Regulation (n 5),

As seen above, the *processing* threshold is easily met during the different phases of the Big Data value chain. However, taking into account that the material scope of EU data protection law also depends on whether the *personal data* threshold is met (and, thus, the *identifiability* criterion as well), it can be argued that the GDPR will only be partially applicable to the Big Data value chain. For instance, in the first stage of the Big Data value chain – data *collection* – various kinds of data can be gathered. If the data is *directly identifiable*, such as the name of a natural person connected to its place of birth, then EU data protection law will probably be applicable.[167] If the data is *indirectly identifiable*, such as a zip code connected to a birthdate, the application of the GDPR will also likely be triggered.[168] However, in cases where anonymous data or non-personal data are collected, EU data protection legislation will not be applicable. The same conclusion of *partial* applicability of the GDPR's material scope holds true for the second and third stages of the Big data value chain (data *storage* and *analytics*). Therefore, since not every stage of the Big Data value chain will necessarily process *personal data*, the applicability of EU data protection legislation to Big Data projects will most likely be partial.[169]

While the material scope of the GDPR tends to be only partially applicable to the Big Data value chain, it is relevant to observe that the territorial scope of EU data protection has been extended with the GDPR. In practice, this means that more situations within the Big Data value chain are likely to be caught by the novel territorial reach, which applies (i) when companies have an establishment in the EU and process personal data in the context of their activities or (ii) when companies are not established in the EU but process personal data of individuals who are in the EU and such

---

recital 30: 'Natural persons may be associated with online identifiers provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers or other identifiers such as radio frequency identification tags. This may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them.'

[167] Oostveen (n 23) 306-307.

[168] As mentioned by Oostveen (n 23) 307, it is important to take into account that the controller makes a self-assessment of how the data are categorized and if they are identifiable or non-identifiable. Thus, in practice companies tend to characterize the data they process as non-identifiable in order to escape the material scope of data protection law.

[169] In the words of Oostveen (n 23) 309, 'it is important to realize that big data is not completely covered by data protection'.

processing is related to the offering of goods and services or to the monitoring of their behavior.[170] In other words, the territorial scope of EU data protection law is not limited anymore to companies established in the EU, as it can also reach companies outside the EU under certain circumstances. This expansion will most likely affect the Big Data value chain in the sense that the processing of data in the *collection*, *storage* and *analysis* phases by different companies worldwide is more likely to be caught by the GDPR's extensive territorial scope of application.

3.2.2. Purpose Limitation Principle, Repurposing and Unforeseen Purposes

The purpose limitation principle of Article 5(1)(b) GDPR can also pose a challenge to the relationship between data protection legislation and the Big Data value chain. According to this principle, personal data shall only be collected for 'specified, explicit and legitimate purposes' and it shall not be further processed in a way which is 'incompatible' with the original purposes.[171] The purpose limitation principle neither bars any new purposes for processing nor requires that the new purpose and the original purpose be the same.[172] Rather, it simply emphasizes that the new purpose for processing personal data needs to be 'compatible' with the original purpose, meaning that an assessment of compatibility of processing purposes must be made.[173]

The literature suggests that the purpose limitation principle is at odds with the prospect of Big Data analytics.[174] Big Data projects recurrently involve the use of data in ways that the controller did not even consider at the time of collection – a practice

[170] General Data Protection Regulation (n 5), art 3(1) and art 3(2).
[171] General Data Protection Regulation (n 5), art 5(1)(b).
[172] Richard Kemp, 'Big Data and Data Protection' [2014] Kemp IT Law <http://www.kempitlaw.com/wp-content/uploads/2014/10/Big-Data-and-Data-Protection-White-Paper-v1_0-November-2014.pdf> accessed 06 June 2018 12.
[173] Information Commissioner's Office, *Big data, artificial intelligence, machine learning and data protection* (Discussion Paper, September 2017) 37; Tal Zarsky, 'Incompatible: The GDPR in the Age of Big Data' (2017) 47:4(2) Seton Hall Law Review 995, 1007.
[174] According to Mantelero, 'the big data paradigm also undermines the very notion of "specified purpose", in terms of the scope of data processing which should be known and defined at the moment of data collection.' See Alessandro Mantelero, 'Regulating Big Data. The Guidelines of the Council of Europe in the Context of the European Data Protection Framework' (2017) 33 Computer Law & Security Review 586-587. See also Broeders, Schrijvers, van der Sloot, van Brakel, de Hoog and Ballin (n 145) 316; Zarsky (n 173) 1005.

known as 'repurposing'.[175] Analyzing data for different purposes is at the core of the Big Data phenomenon and can benefit society when, for example, firms use the location of mobile phones to predict traffic jams.[176] Issues may arise, however, when personal data obtained during the provision of a certain service is subsequently used for a purpose that is not necessarily compatible with the original specified, explicit and legitimate purpose.[177] To illustrate, if the Apple Health app collects users' personal health data with the purpose of using it to inform users about their health statistics, but subsequently uses this information to assess users' health risks (and possibly to sell this information to health insurance companies), this 'repurposing' is likely to be rendered *incompatible* and *unfair* unless data subjects are informed of such repurposing and freely give their consent to it. Accordingly, if a company active in Big Data analytics wishes to collect personal data in the first stage of the Big Data value chain by purchasing it from data brokers, it needs to ensure that the processing it intends to do is *compatible* with the original purpose for which the data was collected in the first place, or otherwise if it needs to obtain an additional consent from the data subjects.[178]

Due to the possibility of 'repurposing' and the risks that come with it, some argue that a solely consent-based model is not adequate anymore to protect one's personal data.[179] Moreover, taking into account that Big Data aims to find unexpected correlations and insights between different datasets and to unfold novel usages for data, companies may not foresee at the outset all potential purposes and uses for which data may be collected.[180] Such 'unforeseen purposes' also carry the risk of not being considered compatible with the original processing purposes and, henceforth, with the GDPR's purpose limitation principle.[181]

[175] Zarsky (n 173) 1006; McDermott (n 7) 4.
[176] Information Commissioner's Office (n 173) para 154.
[177] Kemp (n 172) 11; McDermott (n 7) 4.
[178] Information Commissioner's Office (n 173) para 83.
[179] McDermott (n 7) 4; Stucke and Grunes (n 19) 6; Stucke and Grunes (n 10) 326; Mantelero (n 174) 587.
[180] See Mantelero (n 174) 587, 'the descriptions of the purposes of data collection have become vague or extremely broad. Data is used on the basis of the inferences that arise in an unpredictable manner when analytics mine databases; as a result, the specific uses of data cannot always be known or expected at the moment of data collection'. See also Broeders, Schrijvers, van der Sloot, van Brakel, de Hoog and Ballin (n 145) 316.
[181] Broeders, Schrijvers, van der Sloot, van Brakel, de Hoog and Ballin (n 145) 316.

The line between what is a 'compatible' or an 'incompatible' purpose is not a simple one and should be drawn on a case-by-case basis.[182] In order to ascertain whether processing for a purpose other than that for which the personal data had been initially collected is compatible with the original purpose, the GDPR provides certain factors that the controller should take into consideration, including (i) any *link* between the original purposes and intended further purposes, (ii) the *context* in which the personal data was collected, (iii) the *nature* of the personal data, (iv) *possible consequences* of the intended additional processing, and (v) if there are any *appropriate safeguards* such as encryption or anonymization.[183] All of the above-mentioned factors come down to a notion of 'fairness' and, in essence, question whether the data subject had reasonable expectations that its personal data could be used for another purpose and how this new purpose could affect the data subject's privacy.[184] Yet, these factors have been criticized for being somewhat abstract and difficult to establish in a Big Data context.[185]

3.2.3. Complexity of Big Data as an Excuse for Not Obtaining Consent

The lawfulness of processing, particularly with regard to consent, can be a third possible challenge involving data protection and the Big Data world. The data subject's consent is one of the most common legal bases for the lawful processing of personal data.[186] According to Article 6(1)(a) GDPR, processing is considered to be lawful if and to the extent that 'the data subject has given consent to the processing of his or her personal data for one or more specific purposes'.[187] Consent is defined by EU data protection legislation as any freely given, specific, informed and unambiguous indication of agreement by the data subject to the processing of his or her personal data, be it

[182] Kemp (n 172) 11; Information Commissioner's Office (n 173) 38.
[183] General Data Protection Regulation (n 5), art 6(4)(a)-(e).
[184] Information Commissioner's Office (n 173) para 81.
[185] Zarsky (n 173) 1008.
[186] It is relevant to keep in mind that although consent is one of the possible legal bases under Article 6(1) GDPR for processing personal data, it is not the only one and it does not have a greater status than the other ones. See Information Commissioner's Office (n 173) para 65; Francisco Costa-Cabral and Orla Lynskey, 'Family ties: the intersection between data protection and competition in EU Law' (2017) 54:1 Common Market Law Review 16.
[187] General Data Protection Regulation, art 6(1)(a).

through a statement or by a clear affirmative action.[188] The logic behind consent is that data subjects need to be able to understand what the controller is going to do with their data and they should provide a clear indication of their agreement to it.[189]

The standards which need to be met in order for consent to be valid have increasingly risen in the past years.[190] Article 7 GDPR explicitly sets forth certain conditions for consent. For instance, when the legal basis for processing is consent, the controller needs to be able to prove that the data subject has consented to the processing of his or her personal data.[191] Also, when consent is given by a written declaration which also comprises content about other matters, the request for consent must be 'presented in a manner which is clearly distinguishable from the other matters, in an intelligible and easily accessible form, using clear and plain language'.[192] Moreover, it should be as easy for the data subject to withdraw its consent as it is for it to give its consent.[193]

The issue of consent can be problematic in the Big Data world.[194] It is acknowledged that the traditional approach to consent – the 'notice-and-consent' model – may not be very practical for Big Data situations and that there is a need to go beyond this simple model.[195] Apart from the fact that the notice-and-consent model is binary, as it only gives data subjects the choice of 'yes' or 'no', the non-transparent nature of data analytics can hinder a valid consent from being provided.[196] Even when companies have a notice and consent policy, a rather small percentage of consumers actually take their time to read them, while the majority of consumers usually tick the 'I agree' boxes without even reading all of the terms and conditions.[197] Such policies are usually long, detailed, and written in vague, opaque and legalistic terms, making it difficult for

[188] General Data Protection Regulation, art 4(11). According to Information Commissioner's Office (n 173) para 57, a clear affirmative action could be, for instance, ticking a box on a website.
[189] Information Commissioner's Office (n 173) para 56.
[190] Kemp (n 172) 11.
[191] General Data Protection Regulation (n 5), art 7(1).
[192] General Data Protection Regulation (n 5), art 7(2).
[193] General Data Protection Regulation (n 5), art 7(3).
[194] Mantelero (n 174) 598-599; Kemp (n 172) 13.
[195] Information Commissioner's Office (n 173) para 58; McDermott (n 7) 4; Stucke and Grunes (n 19) 6; Stucke and Grunes (n 10) 326.
[196] Mantelero (n 174) 598-599; Information Commissioner's Office (n 173) para 58.
[197] Grunes and Stucke (n 10) 12; Information Commissioner's Office (n 152) para 143.

consumers to understand them or to negotiate better terms and conditions of use, leading to a significant imbalance of power between users and service providers.[198]

Moreover, the predictive and self-learning power of the algorithms used in Big Data analytics can unfold types of personal data that an individual did not necessarily consent to be collected or processed – for example, data predicting an individual's willingness to pay for certain products or services.[199] In this sense, when Big Data projects come into play, the traditional notice-and-consent model becomes far from optimal, especially due to data analytics' experimental nature and its high tendency of finding new usages for data.[200]

Yet, any excuse in an attempt to justify the lack of users' consent should not succeed. Although Big Data projects tend to be complex, especially during the third phase of *data analytics*, caution is needed to avoid that complexity is used as an excuse for not obtaining consent whenever it is required.[201] In other words, businesses should not rely on the complexity of Big Data analytics as a justification for failing to comply with the consent requirement of the GDPR.[202] This can be particularly difficult for companies involved in Big Data processes, since whoever finds 'previously unobserved and unexpected correlations is at a premium'.[203]

In order to avoid potential problems with these excuses, the literature has already identified new forms of giving consent which could be more realistic for Big Data projects. For example, one report has suggested the following two alternatives: (i) a process of *graduated consent* that allows data subjects to give their consent to different usages of their personal data at various moments in time throughout their relationship with controllers, instead of doing so only at the beginning (the so-called 'just-in-time'

[198] Grunes and Stucke (n 10) 12; Stucke and Grunes (n 10) 328; European Data Protection Supervisor, *Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy* (Preliminary Opinion of the European Data Protection Supervisor, March 2014) para 79; Information Commissioner's Office (n 173) para 144.
[199] See above in Part 2.2.4 for a discussion on the notion of 'willingness to pay' and the risk of price discrimination between consumers. See also Oostveen (n 23) 302.
[200] Information Commissioner's Office (n 173) para 58.
[201] Kemp (n 172) 11.
[202] Kemp (n 172) 11.
[203] Kemp (n 172) 13.

notifications); and (ii) a *time-limited consent*, in which data cannot be further used once a certain time limit has lapsed.[204] Both options appear to be functional and feasible alternatives that adapt the 'notice-and-consent' model to the Big Data world, either by dividing consent into various parts or by limiting the consent timeframe. Nonetheless, the first alternative can still clash with the notion of *data analytics*, as it would be unpractical for firms to request user consent for *every* new distinct usage of personal data – which, in a Big Data world, can occur quite often. In the end, who would be willing to spend hours reading 'boring' privacy notices that pop up every month, week, or day? Perhaps a solution to this would be for companies to produce more user-friendly policy notices by making use of a combination of innovative approaches – such as videos, cartoons and standardized icons – in order to make the information easier (and faster) to understand.[205] In sum, these alternatives demonstrate that the complexity of Big Data should not pose as an obstacle for controllers to seek consent from data subjects, as there are many possibilities out there of *how* to seek consent in creative ways.

3.2.4. Data Minimization Principle and Big Data's Pursuit of Volume and Variety

A fourth possible challenge involving the Big Data value chain and EU data protection legislation is the data minimization principle. Pursuant Article 5(1)(c) GDPR, personal data shall be 'adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed'.[206] Simply put, data minimization limits the collection of personal data to that which is strictly adequate, relevant and necessary to achieve a specified purpose. Such requirement minimizes the risk of leakage by controllers and reduces the chances that controllers violate users' privacy by going beyond consented usage.[207] Nevertheless, two of the four V's of Big Data – *volume* and *variety* – can possibly clash with the data minimization principle.

[204] Information Commissioner's Office (n 173) paras 59, 61, 149.
[205] Mantelero (n 174) 599; Information Commissioner's Office (n 173) 62, para 145.
[206] General Data Protection Regulation (n 5), art 5(1)(c).
[207] Zarsky (n 173) 1010.

As exposed in Part 2.1, one of the main characteristics of Big Data is *volume*. The logic behind the Big Data value chain clearly incentivizes companies to *collect*, *store* and *process* as much data as possible for as long as possible.[208] The more data are collected and processed about a person, the more knowledge there is about intimate details of a person's life and the more correlations a Big Data firm can potentially make.[209] The abundant amount of personal data collected in Big Data projects can, however, *exceed what is necessary* in relation to the processing purposes and thereby infringe the data minimization principle set forth by the GDPR.[210]

Likewise, *variety* is also another relevant characteristic of Big Data. With a greater variety of data combined from different sources, new unforeseeable and unpredictable inferences can be made and new data about individuals can be created.[211] Yet, a greater variety of data sources can also raise questions regarding whether such data is actually *relevant* to accomplish the specific purposes for processing or whether it goes beyond the data minimization principle.[212]

Therefore, Big Data's rush to collect massive amounts of data from a variety of different sources can be in conflict with the GDPR's data minimization principle. According to Broeders et al., there are 'inherent tensions' between Big Data and legal principles such as data minimization.[213] Since the notion of data minimization can limit the success of Big Data initiatives and undermine their potential utility, some advocates of the incompatibility of the GDPR vis-à-vis Big Data argue that data minimization requirements should be loosened and that further privacy concerns should be dealt with by *ex post* regulation.[214] Nevertheless, this opinion seems to represent only a minority of the existent literature.

[208] Kemp (n 172) 11; Zarsky (n 173) 1010-1011.
[209] Oostveen (n 23) 302.
[210] Information Commissioner's Office (n 173) para 85.
[211] Oostveen (n 23) 302.
[212] Kemp (n 172) 12.
[213] Broeders, Schrijvers, van der Sloot, van Brakel, de Hoog and Ballin (n 145) 316-317.
[214] Zarsky (n 173) 1011.

4. The Interplay Between Big Data, Competition Law and Data Protection

Parts 2 and 3 scrutinized some of the competition law and data protection challenges in the era of Big Data. In its turn, Part 4 aims to assess possible overlapping legal concerns of competition law and data protection, including the right to data portability and the question of whether competition authorities should consider data protection concerns in their merger and abuse of dominant position analyses. To understand how competition authorities have dealt with the latter question, a brief overview of two recent cases involving Facebook is presented. Lastly, it is discussed whether there is a scope for cooperation between EU competition and data protection authorities when it comes to cases that touch upon both fields of law.

4.1. The Double Scope of Application of the Right to Data Portability

Data portability is a concern that can arise in both competition law and data protection domains. In the realm of competition law, data portability can be conceived as a competition law remedy.[215] Imagine, for instance, that a dominant company refuses to port data. This could potentially constitute an abuse of dominant position, which would call for an intervention by the competent antitrust authority in order to impose the practice of data portability on the dominant company.[216] In fact, possible antitrust issues involving data portability arose in the European Commission's *Google Search (Shopping)* case.[217] In 2010, the Commission opened investigations into allegations that Google had, among other practices, restricted 'the portability of online advertising campaign data to competing online advertising platforms'.[218]

[215] Orla Lynskey, 'Aligning Data Protection Rights with Competition Law Remedies? The GDPR Right to Data Portability' (2017) 42:6 European Law Review 795.
[216] Bourreau, Streel and Graef (n 10) 25.
[217] *Google Search (Shopping)* (Case AT.39740).
[218] European Commission, 'Antitrust: Commission probes allegations of antitrust violations by Google' (*European Commission Press Release Database*, 30 November 2010) <http://europa.eu/rapid/press-release_IP-10-1624_en.htm?locale=en> accessed 01 August 2018.

From a competition law standpoint, data portability can theoretically reduce future barriers to entry and weaken the market position of a dominant company, as it facilitates the process of individuals switching from one service provider to another and moving their data elsewhere.[219] Thus, data portability empowers individuals to choose providers that match their privacy preferences, lowers switching costs, and incentivizes companies to enter the market.[220]

In the field of data protection, the novel GDPR framework further strengthened data subjects' control over their own data and transformed the concept of data portability into law.[221] Where the data subject allowed the processing of his or her personal data based on consent or where the processing is necessary for the performance of a contract, Article 20 GDPR provides the data subject the right to receive personal data concerning him or her in a structured, commonly used, machine-readable and interoperable format and the right to transmit that data to another controller without interference from the previous controller, provided that the processing is carried out by automated means.[222] Therefore, the contours of the GDPR's right to data portability are defined by four criteria: (i) only *personal data*; (ii) only data *provided by* the data subject; (iii) only data processed pursuant *consent* or *contract*; and (iv) only data processed by *automated means*.[223] As an example of the right to data portability post-GDPR, Facebook now gives its users the possibility to download a copy of their Facebook account information at any time.[224] The social network company allows its users to download the information all at once or to select only certain types of information and

[219] According to Geradin and Kuschewsky (n 31) 9, data portability is 'key to market entry'. See also European Data Protection Supervisor, *Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy* (n 198) 36.
[220] Grunes and Stucke (n 10) 13.
[221] Note that in comparison to the previous data protection Directive of 1995, the right to data portability is one of the only 'brand new' rights introduced by the GDPR. It was seemingly inspired by the concept of number portability, which allows consumers to switch from one mobile phone provider to another without having to change numbers. In this regard, see Lynskey (n 215) 794, 796 and Geradin and Kuschewsky (n 31) 9. See also General Data Protection Regulation (n 5), recital 68; Information Commissioner's Office (n 173) para 187.
[222] General Data Protection Regulation (n 5), art 20(1), recital 68.
[223] Lynskey (n 215) 799.
[224] Since the GDPR came into force, Facebook provides its users a 'Download Your Information' tool: 'You can download a copy of your Facebook information at any time. You can download all of it at once, or you can select only the types of information and date ranges you want. You can choose to receive your information in an HTML format that is easy to view, or a JSON format, which could allow another service to more easily import it.' Facebook, 'Accessing and Downloading Your Facebook Information' <https://www.facebook.com/help/contact/2032834846972583> accessed 16 June 2018.

data ranges. Moreover, users can download the information in an HTML format that is easy to view or in a JSON format, which allows other service providers to easily import the data.

As the European Data Protection Supervisor reported in a 2014 study, data portability could release synergies between data protection and competition law in at least two ways: (i) it could prevent abuse of dominance situations and avoid consumers from being locked into certain services and (ii) it could empower consumers to take advantage of third-party value-added services while facilitating competitors' greater access to the market.[225] Yet, it is relevant to notice that the scope of application and the objectives of data portability are not exactly identical in these two fields of law. According to Bourreau, Streel and Graef, the main differences between data portability in competition law cases and in GDPR-related cases is that in the former the portability would be applicable to all types of data (both *personal* and *non-personal* data) but only dominant firms would be covered, while in the latter the portability would be applicable only to what the GDPR considers *personal data* but all firms (dominant and non-dominant) would be covered.[226] Lynskey further complements such analysis by comparing data portability as a competition law *remedy* and data portability as a data protection law *right*. Lynskey's comparison is summarized in the table below:

	Competition Law Remedy	**Data Protection Law Right**
Rationale	- Fostering competition by making competing options available to consumers and ensuring that they have the ability to choose among these options	- Empowering data subjects through individual control over their personal data
Personal Scope	- Applies only to 'undertakings' (entities engaged in economic activity)	- Applies to 'data controllers' (including undertakings but also entities or individuals not engaged in

[225] European Data Protection Supervisor, *Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy* (n 198) 36; Grunes and Stucke (n 10) 13.
[226] Bourreau, Streel and Graef (n 10) 25.

		economic activity)
Material Scope	- Limited to a finding that an undertaking is dominant on a relevant market and that its conduct was abusive - Applies to any type of data, 'personal' or 'non-personal'	- Limited to (i) 'personal data' (ii) 'provided by' the data subject and (iii) processed pursuant 'consent' or 'contract' and (iv) only by 'automated means'
Extent of Application	- Confined to the facts of a particular case	- General right available to all
Substantive Dissimilarities	- A dominant undertaking is only required to make the data available to third parties, but not to ease its transfer by directly transmitting it to another undertaking	- If technically feasible, data subject has the right to have its personal data transmitted directly from one controller to another

Figure 4: Comparison Between Data Portability as a Competition Law Remedy and as a Data Protection Right[227]

4.2. Should Competition Authorities Consider Data Protection Concerns in their Analyses?

In addition to the right of data portability, competition law and data protection can also intersect when competition authorities analyze merger deals or abuse of dominant position cases that involve data-driven companies. A question that has gained momentum in the last years has been whether competition authorities should consider data protection and privacy concerns throughout their competition analyses and, if so, how this should be done. Two prominent cases involving the social networking company Facebook are explored below to better illustrate this debate and to show how the approaches of competition authorities can be divergent.

[227] This table was elaborated based on the comparison made by Lynskey (n 215).

4.2.1. *Facebook/WhatsApp* Merger Decision by the European Commission

In October 2014, the European Commission authorized the acquisition of WhatsApp by Facebook.[228] The Commission identified three distinctive relevant product markets – (i) consumer communication services, (ii) social networking services and (iii) online advertising services – and concluded after analyzing each market that the concentration between the two data-driven companies would not give rise to any anticompetitive concerns.[229]

Although for the purposes of this paper there is no need to go into details about the Commission's competition law analysis, it is nevertheless relevant to mention the Commission's understanding that 'any privacy-related concerns flowing from the increased concentration of data within the control of Facebook as a result of the Transaction do not fall within the scope of the EU competition law rules but within the scope of the EU data protection rules'.[230] Accordingly, the Commission explicitly excluded any privacy-related concerns from its analysis of the *Facebook/WhatsApp* merger and only considered the issue of data concentration to the extent that it would be likely to strengthen Facebook's position in the online advertising services market.[231]

Nonetheless, much criticism was made towards the possible use of data after the transaction.[232] Critics argued that the Commission may have underestimated the true value of data in this deal, especially taking into account that Facebook would benefit from a higher velocity of data collection and a greater capability for real-time analysis, reason why it was willing to pay the high price of US$21.8 billion for the acquisition of a company with low revenues and high net losses.[233] Moreover, concerns were raised over the fact that the Commission did not analyze whether the merger was a defensive mechanism aimed at depriving Facebook's competitor of the scale necessary to compete on the market of consumer communication services. Another point brought to

[228] *Facebook/WhatsApp* (Case COMP/M.7217) Commission Decision C(2014) 7239 final.
[229] *Facebook/WhatsApp* 35.
[230] *Facebook/WhatsApp* para 164.
[231] *Facebook/WhatsApp* para 164; Whish and Bailey (n 56) 81.
[232] Ariel Ezrachi, *EU Competition Law: An Analytical Guide to the Leading Cases* (Fifth edition, Hart Publishing 2016) 454.
[233] Ezrachi (n 232) 454; Whish and Bailey (n 56) 80, 82.

discussion was that the merger would serve to prevent privacy-focused texting apps such as WhatsApp from gaining a strong market position and overtaking privacy-intrusive texting services such as Facebook Messenger.[234] All things considered, critics reasoned that at least three different groups could be potentially harmed by the merger: (i) competitors, by being foreclosed from achieving sufficient scale to compete; (ii) users of texting apps, by enjoying less privacy protection, quality and innovation; and (iii) advertisers, by facing higher service rates.[235]

In the words of Whish and Bailey, 'the Commission simply erred in stating that the concerns of one firm controlling so much data were strictly a privacy issue, not a competition issue'.[236] In their view, the Commission failed to see the whole picture by analyzing the potential impacts of data concentration only on one side of the multi-sided market – namely the side of the advertising market –, while data concentration can actually touch upon and affect multiple sides of the market.

Indeed, some of the voiced concerns became reality when WhatsApp announced in late 2016 updates to its terms of service and privacy policy indicating the possibility of linking its users' phone numbers with their Facebook accounts, which consequently led the Commission to reopen the case and assess if there were any 'incorrect' or 'misleading' information provided by the merging companies.[237] Since Facebook had informed the Commission back in 2014 that it would be unable to automatically link users' WhatsApp and Facebook accounts, but later it was discovered that the company was already exploring ways to do so at the time of the merger investigation and that it in fact implemented such measures in 2016, the Commission considered that the information provided by Facebook during the merger was incorrect and misleading and, thus, decided to fine the company in EUR 110 million in May 2017.[238] Following the Commission's fining decision – which did not have any impact on its previous decision to

[234] Whish and Bailey (n 56) 83.
[235] Whish and Bailey (n 56) 76.
[236] Whish and Bailey (n 56) 81.
[237] *Facebook/WhatsApp* (Case M.8228) Commission Decision 2017/C 286/06 [2017] OJ C 286/6.
[238] *Facebook/WhatsApp* (Case M.8228) Commission Decision 2017/C 286/06 [2017] OJ C 286/6.

authorize the merger –, national data protection regulators in Germany[239], France[240], and the United Kingdom[241] have ordered WhatsApp to stop sharing user data with Facebook without previous user consent.

4.2.2. Facebook Investigation by Germany's Bundeskartellamt

Amidst the growing concerns over the implications of the *Facebook/WhatsApp* merger, the German competition authority started proceedings against Facebook in March 2016 for an alleged abuse of dominant position in the market for social networks.[242] The Bundeskartellamt suspected that the company's specific terms and conditions regarding the collection and use of user data were in violation of data protection legislation.[243] In an early press release, the German watchdog cautioned that it is difficult for social network users to fully comprehend the scope of the agreement accepted by them in the terms of service and that, depending on the circumstances, this type of behavior could constitute an abusive practice under German competition law.[244]

After approximately two years of investigation, the Bundeskartellamt released a preliminary legal assessment under which it held the view that Facebook was abusing its dominant position on the German market for social networks by making the use of its social network service conditional upon the user granting the company permission to limitlessly collect every type of data generated by third-party websites and to merge it

[239] The Hamburg Commissioner for Data Protection and Freedom of Information, 'Administrative Order Against the Mass Synchronisation of Data Between Facebook and WhatsApp' (27 September 2016) <https://datenschutz-hamburg.de/assets/pdf/Press_Release_2016-09-27_Adminstrative_Order_Facebook_WhatsApp.pdf> accessed 02 August 2018.
[240] Commission Nationale de l'Informatique et des Libertés, 'Data Transfer from WhatsApp to Facebook: CNIL Publicly Serves Formal Notice for Lack of Legal Basis' (18 December 2017) <https://www.cnil.fr/en/data-transfer-whatsapp-facebook-cnil-publicly-serves-formal-notice-lack-legal-basis> accessed 02 August 2018.
[241] Information Commissioner's Office, 'Blog: A Win for the Data Protection of UK Consumers' (14 March 2018) <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2018/03/blog-a-win-for-the-data-protection-of-uk-consumers/> accessed 02 August 2018.
[242] Bundeskartellamt, 'Bundeskartellamt initiates proceeding against Facebook on suspicion of having abused its market power by infringing data protection rules' (*Bundeskartellamt*, 02 March 2016) <https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2016/02_03_2016_Facebook.html> accessed 18 June 2018.
[243] Bundeskartellamt, 'Bundeskartellamt initiates proceeding against Facebook on suspicion of having abused its market power by infringing data protection rules' (n 242).
[244] Bundeskartellamt, 'Bundeskartellamt initiates proceeding against Facebook on suspicion of having abused its market power by infringing data protection rules' (n 242).

with users' Facebook account.[245] The Bundeskartellamt's current concern in this investigation is limited to the collection of data *outside* Facebook's social network – that is, data gathered via third-party websites and apps with embedded Facebook Application Programming Interfaces (so-called 'APIs') – and the merging of this data to the users' Facebook account.[246] Such data collection outside of Facebook can take place, for instance, when users make use of services owned by the Facebook company – such as WhatsApp or Instagram – or of third-party websites that have a Facebook 'like' button or a Facebook 'login' option.[247]

The German competition authority stated that users are generally unaware that their data can be collected and transmitted to Facebook even when they are visiting other websites and warned that this raises questions regarding the validity and effectiveness of users' consent to Facebook's data processing activities.[248] Therefore, although the case is in essence a competition law case involving an alleged abuse of dominant position, it has strong ties to two other fields of law – data protection and consumer protection. For this reason, the Bundeskartellamt has worked in close cooperation with data protection authorities throughout the investigation.[249] The preliminary legal assessment also clarified why a competition authority is dealing with this case: 'where access to the personal data of users is essential for the market position of a company, the question of how that company handles the personal data of its users is no longer only relevant for data protection authorities. It becomes a relevant question for the competition authorities, too'.[250]

Therefore, unlike in the European Commission's *Facebook/WhatsApp* merger analysis, the German Bundeskartellamt has up to the present date taken the view that

[245] Bundeskartellamt, 'Preliminary assessment in Facebook proceeding: Facebook's collection and use of data from third-party sources is abusive' (*Bundeskartellamt*, 19 December 2017) <https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2017/19_12_2017_Facebook.html> accessed 18 June 2018; Bundeskartellamt, 'Background information on the Facebook proceeding' (n 64) 4.
[246] Bundeskartellamt, 'Preliminary assessment in Facebook proceeding: Facebook's collection and use of data from third-party sources is abusive' (n 245).
[247] Bundeskartellamt, 'Background information on the Facebook proceeding' (n 64) 2.
[248] Bundeskartellamt, 'Preliminary assessment in Facebook proceeding: Facebook's collection and use of data from third-party sources is abusive' (n 245).
[249] Bundeskartellamt, 'Preliminary assessment in Facebook proceeding: Facebook's collection and use of data from third-party sources is abusive' (n 245).
[250] Bundeskartellamt, 'Background information on the Facebook proceeding' (n 64) 1-2.

data protection principles can and should be used throughout antitrust assessment in order to determine whether Facebook is abusing its dominant position. It has relied on the case law from the German Federal Court of Justice to justify the application of EU data protection legislation – including the GDPR – to its assessment of admissibility of Facebook's terms and conditions.[251] The German approach seems to be much more comprehensive, progressive and forward-looking in comparison to the approach taken by the European Commission, as it encompasses common concerns that touch upon competition, data protection and consumer protection legislations.

4.2.3. Data Protection and Privacy as Non-Price Dimensions of Competition

At a first glance, it can be held that privacy concerns are not in and of themselves within the general scope of competition law. The European Commission embraced this view in the *Facebook/WhatsApp* merger by stating that any privacy-related concerns resulting from an increased concentration of data in a post-merger scenario do not fall within the scope of EU competition law rules but rather within the scope of EU data protection rules.[252] However, does this mean that data protection concerns are completely irrelevant to competition law assessment?

As illustrated above, the German competition authority argued otherwise in the preliminary view of its Facebook investigation by upholding that, during its competition assessment, it can consider principles of EU data protection laws. Prior to the Facebook investigation, the German Bundeskartellamt, together with the French Autorité de la Concurrence, had already expressed the opinion that privacy issues should not be excluded from consideration during a competition law assessment just because of their nature and that they can actually be taken into account whenever a company's collection and use of personal data has competition implications in parallel to data protection concerns.[253] In this sense, the joint report reasoned that the fact that some specific legal instruments – such as data protection legislation – exist to resolve

[251] Bundeskartellamt, 'Background information on the Facebook proceeding' (n 64) 4-5.
[252] *Facebook/WhatsApp* para. 164.
[253] Autorité and Bundeskartellamt (n 3) 23.

sensitive issues involving personal data does not entail that competition law will be irrelevant to personal data. Rather, the consideration of data protection policies in competition proceedings could be justified by a 'close link between the dominance of the company, its data collection processes and competition on the relevant markets'.[254]

Likewise, many authors defend that data protection can indeed play an important role in competition law assessment.[255] In their view, data protection can be considered an aspect of non-price competition, meaning that if a dominant undertaking reduces the level of privacy protection it offers then there is consequently a reduction of the product quality itself.[256] The core of the argument lies on the understanding that standards of data protection and privacy are parts of the 'quality' parameter of a product or service. Such authors also maintain that the recognition of privacy as a non-price dimension of competition entails that firms can compete to offer more or less privacy protection to its customers.[257] Thus, similar to quality, variety, and innovation, 'privacy preferences' cannot be measured in the same way as price, given that different customers have different privacy preferences. Nonetheless, a certain degree of caution is needed when considering these arguments, given that competition authorities have not yet comprehensively adopted privacy or data protection as significant parameters of competition. Even so, as will be argued below, this paper endorses the view that data protection can be considered during a competition law assessment, and this can be facilitated by a close cooperation between regulators.

4.3. Scope for Cooperation Between Competition and Data Protection Authorities

Taking into account the example of the two Facebook cases examined above, a discussion that arises is if there is a scope for cooperation between EU competition and

[254] Autorité and Bundeskartellamt (n 3) 24.
[255] Costa-Cabral and Lynskey (n 186) 11; Sokol and Comerford (n 48) 1144.
[256] Sokol and Comerford (n 48) 1144; Grunes and Stucke (n 10) 4; Autorité and Bundeskartellamt (n 3) 24-25; European Data Protection Supervisor, *EDPS Opinion on coherent enforcement of fundamental rights in the age of big data* (Opinion 8/2016, 23 September 2016) 13.
[257] Sokol and Comerford (n 48) 1144; Grunes and Stucke (n 10) 4.

data protection authorities when it comes to Big Data-related cases that touch upon both fields of law.

In the EU, competition and data protection legislation have common goals, such as the aim of promoting 'fairness'.[258] The notion of 'fairness' is deeply rooted in both competition law and data protection fields.[259] For instance, European Commissioner Margrethe Vestager has used the concept of fairness in several speeches.[260] Likewise, the GDPR has adopted the fairness of personal data processing as a core principle of its framework.[261] Nevertheless, while the notion of fairness pervades both legal areas, there is still a limited cooperation between authorities at a European level.[262] Although there are significant overlaps in terms of substance, which in theory could facilitate the collaboration and teamwork between the competent EU authorities, the enforceability of EU rules is in fact still quite fragmented.[263] This may be due to the fact that regulators have separate jurisdictions and tend to respect the powers and competences of other authorities by not enforcing laws in other legal fields.[264]

Yet, as seen in the recent cases involving Facebook, regulatory jurisdictions of competition and data protection authorities may have points of intersection, and in those scenarios the synergies between the different fields of law could lead to a closer cooperation between authorities in order to synchronize enforcement policies, especially where one of them lacks expertise in a certain area.[265] In this sense, the European Data

[258] For a more detailed analysis of the concept of fairness and its link to EU competition and data protection law, see Harri Kalimo and Klaudia Majcher, 'The Concept of Fairness: Linking EU Competition and Data Protection Law in the Digital Marketplace' (2017) 42:2 European Law Review 210.
[259] Kalimo and Majcher (n 258).
[260] For instance, in a speech delivered in 2016, Margrethe Vestager stated: 'We all care about fairness. And the rules on data protection, on competition and on consumer protection all play a part in making that fairness a reality.' Margrethe Vestager, "Big Data and Competition" (n 2).
[261] General Data Protection Regulation (n 5), art 5(1)(a).
[262] Bourreau, Streel and Graef (n 10) 10.
[263] European Data Protection Supervisor, *EDPS Opinion on coherent enforcement of fundamental rights in the age of big data* (n 256) 9.
[264] European Data Protection Supervisor, *EDPS Opinion on coherent enforcement of fundamental rights in the age of big data* (n 256) 9.
[265] Bourreau, Streel and Graef (n 10) 10; European Data Protection Supervisor, *EDPS Opinion on coherent enforcement of fundamental rights in the age of big data* (n 256) 10.

Protection Supervisor has recognized the need for a joined-up enforcement by EU regulators to overcome any 'regulatory fragmentation'.[266]

Indeed, there is a scope for cooperation between competition and data protection regulators, particularly in data-driven cases where the expertise of one authority can be of use to the analysis of the other. In fact, cooperation between competition authorities and other regulators (e.g. financial, telecommunications, energy, and aviation supervisory agencies) already exists in several jurisdictions, as many times antitrust watchdogs do not have the technical expertise to deal with sector-specific issues.[267] It is relevant to keep in mind, however, that there is no one-size-fits-all solution for dealing with Big Data and the consequences of this phenomenon.[268] Even so, EU competition and data protection agencies should join forces in the following years to better understand the functioning of the Big Data value chain, to increase their expertise in data science, and to ultimately cooperate to debate common problems and delineate aligned strategies.[269]

5. Conclusion

The rise of the Big Data phenomenon has increased the possibilities to collect, store and process data in unprecedented ways. This development has posed several challenges to different fields of law, including competition law and data protection. The present paper aimed to identify the main competition law and data protection challenges stemming from the growing use of Big Data, as well as the possible implications arising from the interplay between Big Data, competition law and data protection. The solution to the research question can be summarized as follows.

---

[266] European Data Protection Supervisor, *EDPS Opinion on coherent enforcement of fundamental rights in the age of big data* (n 256) 11.
[267] For instance, see Priscilla Tollini, 'Complementaridade entre Agente Regulador e Autoridade da Concorrência: O Caso Do Sistema Financeiro' [Complementarity Between Regulatory Agency and Competition Authority: The Case of the Financial Sector] (2014) 2:2 Revista de Defesa da Concorrência 23 for a closer look at the cooperation between the Brazilian competition authority ('CADE') and Brazilian financial supervisors in merger cases involving banking institutions.
[268] Broeders, Schrijvers, van der Sloot, van Brakel, de Hoog and Ballin (n 145) 316.
[269] Bourreau, Streel and Graef (n 10) 10.

From a competition law standpoint, the era of Big Data brings along many concerns. First, Big Data can facilitate the establishment and perpetuation of market power in favor of a few players. In assessing data as a factor to establish market power, competition authorities should look at various factors, including entry barriers, network effects, feedback loops, economies of scale, scope, and speed, multi-sided markets, and multi-homing. Second, data-driven companies can make use of strategic mergers and acquisitions to obtain better access to data. For this reason, competition authorities should ensure that they have the adequate analytical tools to scrutinize data-driven mergers and that their notification thresholds are able to capture these types of deals. Third, incumbent companies can adopt exclusionary conducts that deprive competitors from access to data, such as the refusal to access data, a discriminatory access to data, exclusive agreements, and tying arrangements. Fourth, Big Data increases the risk of price discrimination between different customer groups, as companies are able infer which consumers are willing to pay a higher price for certain products or services and which consumers are not. Fifth, Big Data intensifies market transparency, which increases the risks of having a more stable collusion between players.

Moreover, from a data protection point of view, Big Data can also present various issues. First, EU data protection legislation is only partially applicable to the Big Data value chain, as not every data processing activity in the Big Data value chain involves the processing of *personal data*. Second, a core characteristic of Big Data – finding unexpected correlations between different datasets and unfolding new usages for data – clashes with the purpose limitation principle, known to be one of the central pillars of the GDPR. Third, the complex and experimental nature of data analytics challenges the traditional 'notice-and-consent' model, which in turn may instigate data-driven companies to use this as an excuse for not obtaining user consent. Fourth, while the logic of Big Data is to collect, store and process as much data as possible for as long as possible, the GDPR's data minimization principle limits the collection of data only to what is *strictly necessary* in relation to the purposes for which they are processed.

Additionally, it was seen that some legal concerns can be overlapping to both competition and data protection law. One example is the right to data portability, which has a double scope of application. Data portability can be imposed by competition authorities on companies that abuse their dominant position by refusing to port data. Data portability is also now required under the GDPR framework, which provides data subjects the right to receive their personal data in a structured, commonly used, machine-readable and interoperable format and to transfer such data to another company. Although the right to data portability exists in both fields of law, it was seen that its scope of application differs in each one of them.

Another example of the interplay between competition law and data protection is regarding antitrust cases that involve data-driven companies. As seen with two recent cases involving Facebook, the European Commission and the German competition authority apparently adopted diverging views about the question of whether competition authorities should consider data protection and privacy issues in their antitrust analyses. While the former seems to be more reluctant in taking into consideration data protection concerns, the latter presented solid jurisprudential bases supporting their expansive and forward-looking view. In any case, there is strong support in the literature in favor of considering data protection and privacy as non-price dimensions of competition.

At last, it was seen that although competition authorities and data protection supervisors in practice seldom cooperate, there is certainly a scope for cooperation between such regulators at EU level, given that both fields of law have common goals, including the aim of promoting ‘fairness’. In Big Data cases where these areas intersect, regulators have much to gain from a close cooperation, especially given that sometimes antitrust watchdogs do not have sufficient technical expertise to deal with sector-specific issues. Therefore, a joint enforcement action from competition and data protection authorities is needed to solve together mutual cases and overcome any regulatory fragmentation. All in all, the rise of Big Data defies competition authorities, data protection supervisors, policymakers, academia, and practitioners to debate about possible solutions to the numerous legal challenges posed by the Big Data phenomenon.

6. Reference List

Books

- Blanco L, *Market Power in EU Antitrust Law* (First edition, Hart Publishing 2012)
- Ezrachi A, *EU Competition Law: An Analytical Guide to the Leading Cases* (Fifth edition, Hart Publishing 2016)
- Lynskey O, *The Foundations of EU Data Protection Law* (Oxford University Press 2015)
- Stucke M and Grunes A, *Big Data and Competition Policy* (First edition, Oxford Competition Law 2016)
- Whish R and Bailey D, *Competition Law* (Eighth edition, Oxford University Press 2015)

Case Law

- Case C-7/97 *Bronner v Mediaprint* [1998] ECJ I-07791
- Case C-418/01 *IMS Health* [2004] ECJ I-0503
- *Google Android* (Case AT.40099)
- *Google Search (AdSense)* (Case AT.40411)
- *Google Search (Shopping)* (Case AT.39740)
- *Facebook/WhatsApp* (Case M.8228) Commission Decision 2017/C 286/06 [2017] OJ C 286/6
- *Facebook/WhatsApp* (Case COMP/M.7217) Commission Decision C(2014) 7239 final
- *Microsoft/Yahoo! Search Business* (Case COMP/M.5727) Commission Decision C(2010) 1077
- *TomTom/TeleAtlas* (Case COMP/M.4854) Commission Decision C(2008) 1859

Contributions to Edited Books

- Brkan M, 'The Court of Justice of the EU, Privacy and Data Protection: Judge-Made Law as a Leitmotif in Fundamental Rights Protection' in Maja Brkan and Evangelia Psychogiopoulou (eds) *Courts, Privacy and Data Protection in the Digital Environment* (Edward Elgar Publishing, 2017)

EU Legislation

- Charter of Fundamental Rights of the European Union [2000] OJ C 364/1
- Commission Notice on the definition of relevant market for the purposes of Community competition law (97/C 372/03) OJ C 372/5
- Communication from the Commission – Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings [2009] OJ C 45/7
- Consolidated Version of the Treaty on the Functioning of the European Union [2012] OJ C 326/47
- Council Regulation (EC) No 139/2004 of 20 January 2004 on the control of concentrations between undertakings (the EC Merger Regulation) OJ L 24/1
- Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA [2016] OJ L 119/89
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119/1

Hard Copy Journals

- Broeders D, Schrijvers E, van der Sloot B, van Brakel R, de Hoog J and Ballin E, 'Big Data and Security Policies: Towards a Framework for Regulating the Phases of Analytics and Use of Big Data' (2017) 33 Computer Law & Security Review 309
- Costa-Cabral F and Lynskey O, 'Family Ties: The Intersection Between Data Protection and Competition in EU Law' (2017) 54:1 Common Market Law Review 11
- Davilla M, 'Is Big Data a Different Kind of Animal? The Treatment of Big Data Under the EU Competition Rules' (2017) 8:6 Journal of European Competition Law & Practice 370
- Graef I, 'Market Definition and Market Power in Data: The Case of Online Platforms' (2015) 38:4 World Competition: Law & Economics Review 473
- Kalimo H and Majcher K, 'The Concept of Fairness: Linking EU Competition and Data Protection Law in the Digital Marketplace' (2017) 42:2 European Law Review 210
- Lynskey O, 'Aligning Data Protection Rights with Competition Law Remedies? The GDPR Right to Data Portability' (2017) 42:6 European Law Review 793
- Malgieri G and Comandé G, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) 7:4 International Data Privacy Law 243
- Mantelero A, 'Regulating Big Data. The Guidelines of the Council of Europe in the Context of the European Data Protection Framework' (2017) 33 Computer Law & Security Review 584
- McDermott Y, 'Conceptualising the Right to Data Protection in an Era of Big Data' (2017) 4:1 Big Data & Society 1
- Oostveen M, 'Identifiability and the Applicability of Data Protection to Big Data' (2016) 6:4 International Data Privacy Law 299
- Rubinfeld D and Gal M, 'Access Barriers to Big Data' (2017) 59:2 Arizona Law Review 339

- Sokol D and Comerford R, 'Antitrust and Regulating Big Data' (2016) 23:5 George Mason Law Review 1129
- Stucke M and Ezrachi A, 'When Competition Fails to Optimize Quality: A Look at Search Engines' (2016) 18:1 Yale Journal of Law and Technology 72
- Tollini P, 'Complementaridade entre Agente Regulador e Autoridade da Concorrência: O Caso Do Sistema Financeiro' [Complementarity Between Regulatory Agency and Competition Authority: The Case of the Financial Sector] (2014) 2:2 Revista de Defesa da Concorrência 23
- Tucker D and Wellford H, 'Big Mistakes Regarding Big Data' (2014) 14:2 The Antitrust Source 1
- Urgessa W, 'The Protective Capacity of the Criterion of 'Identifiability' under EU Data Protection Law' (2016) 4 European Data Protection Law Review 521
- Zarsky T, 'Incompatible: The GDPR in the Age of Big Data' (2017) 47:4(2) Seton Hall Law Review 995

Online Articles

- Bourreau M, de Streel A and Graef I, 'Big Data and Competition Policy: Market Power, Personalised Pricing and Advertising' [2017] SSRN <https://ssrn.com/abstract=2920301> accessed 01 May 2018
- Fidelis A and Ortaç Z, 'Data-Driven Mergers: A Call For Further Integration Of Dynamics Effects Into Competition Analysis' [2017] Barcelona Graduate School of Economics <https://repositori.upf.edu/bitstream/handle/10230/33467/FidelisOrtac%20TFM2017.pdf?sequence=1&isAllowed=y> accessed 29 July 2018
- Filistrucchi L and Klein T, 'Price Competition in Two-Sided Markets with Heterogeneous Consumers and Network Effects' [2013] NET Institute Working Paper N. 13-20 <https://ssrn.com/abstract=2336411> accessed 24 July 2018

- Geradin D and Kuschewsky M, 'Competition Law and Personal Data : Preliminary Thoughts on a Complex Issue' [2013] SSRN <https://ssrn.com/abstract=2216088> accessed 29 July 2018
- Grunes A and Stucke M, 'No Mistake About It: The Important Role of Antitrust in the Era of Big Data' [2015] Competition Policy International Antitrust Chronicle <https://www.competitionpolicyinternational.com/assets/Uploads/StuckeGrunesMay-152.pdf> accessed 30 April 2018
- Kemp R, 'Big Data and Data Protection' [2014] Kemp IT Law <http://www.kempitlaw.com/wp-content/uploads/2014/10/Big-Data-and-Data-Protection-White-Paper-v1_0-November-2014.pdf> accessed 06 June 2018
- Stucke M and Grunes A, 'Debunking the Myths Over Big Data and Antitrust' [2015] Competition Policy International Antitrust Chronicle <https://ssrn.com/abstract=2612562> accessed 30 April 2018

Reports

- Article 29 Data Protection Working Party, *Opinion 4/2007 on the concept of personal data* (20 June 2007) <https://www.clinicalstudydatarequest.com/Documents/Privacy-European-guidance.pdf> accessed 03 August 2018
- Autorité de la Concurrence and Bundeskartellamt, *Competition Law and Data* (10 May 2016) <https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.pdf?__blob=publicationFile&v=2> accessed 11 May 2018
- Competition and Markets Authority, *The Commercial Use of Consumer Data: Report on the CMA's Call for Information* (June 2015)
- European Data Protection Supervisor, *EDPS Opinion on Coherent Enforcement of Fundamental Rights in the Age of Big Data* (Opinion 8/2016, 23 September 2016)

- European Data Protection Supervisor, *Privacy and Competitiveness in the Age of Big Data: The Interplay Between Data Protection, Competition Law and Consumer Protection in the Digital Economy* (Preliminary Opinion of the European Data Protection Supervisor, March 2014)
- European Parliament, *Report on Fundamental Rights Implications of Big Data: Privacy, Data Protection, Non- Discrimination, Security and Law-Enforcement (2016/2225(INI))* (Committee on Civil Liberties, Justice and Home Affairs, 20 February 2017)
- Information Commissioner's Office, *Big Data, Artificial Intelligence, Machine Learning and Data Protection* (Discussion Paper, September 2017)
- Monopolkomission, *Competition Policy: The Challenge of Digital Markets* (Special Report No 68, 2015)
- OECD, *Algorithms and Collusion: Competition Policy in the Digital Age* (2017)
- OECD, *Big Data: Bringing Competition Policy to the Digital Era* (Executive Summary, 29-30 November 2016)
- OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being* (OECD Publishing, Paris, 2015)
- OECD, *Data-Driven Innovation for Growth and Well-being* (Interim Synthesis Report, 2014)
- OECD, *Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data"* (OECD Digital Economy Papers, No. 222, OECD Publishing, Paris, 2013)

Speeches

- Margrethe Vestager, "Big Data and Competition" (EDPS-BEUC Conference, Brussels 29 September 2016) <https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/big-data-and-competition_en> accessed 01 December 2017

- Margrethe Vestager, “Competition in a big data world” (DLD 16, Munich 17 January 2016) <https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/competition-big-data-world_en> accessed 04 December 2017
- Margrethe Vestager, “Helping people cope with technological change” (Rencontres de Bercy, Paris 21 November 2017) <https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/helping-people-cope-technological-change_en> accessed 04 December 2017
- Margrethe Vestager, “Refining the EU merger control system” (Studienvereinigung Kartellrecht, Brussels, 10 March 2016) <https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/refining-eu-merger-control-system_en> accessed 22 May 2018
- Margrethe Vestager, “What competition can do – and what it can’t” (Chilling Competition Conference, 25 October 2017) <https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/what-competition-can-do-and-what-it-cant_en> accessed 04 December 2017

Websites / Blog Posts / Press Releases

- Berg W and Weinert L, ‘New Merger Control Threshold in Germany – Beware of Ongoing Transactions’ (*Kluwer Competition Law Blog*, 7 June 2017) <http://competitionlawblog.kluwercompetitionlaw.com/2017/06/07/new-merger-control-threshold-germany-beware-ongoing-transactions/> accessed 16 May 2018
- Bundeskartellamt, ‘Background information on the Facebook proceeding’ (*Bundeskartellamt*, 19 December 2017) <https://www.bundeskartellamt.de/SharedDocs/Publikation/EN/Diskussions_Hinte

rgrundpapiere/2017/Hintergrundpapier_Facebook.pdf?__blob=publicationFile&v=6> accessed 22 May 2018

- Bundeskartellamt, 'Bundeskartellamt initiates proceeding against Facebook on suspicion of having abused its market power by infringing data protection rules' (*Bundeskartellamt*, 02 March 2016) <https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2016/02_03_2016_Facebook.html> accessed 18 June 2018
- Bundeskartellamt, 'Preliminary assessment in Facebook proceeding: Facebook's collection and use of data from third-party sources is abusive' (*Bundeskartellamt*, 19 December 2017) <https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2017/19_12_2017_Facebook.html> accessed 18 June 2018
- Cappellari S and Birmanns S, 'Germany: Merger Control' (*Global Competition Review*, 14 August 2017) <https://globalcompetitionreview.com/insight/the-european-middle-eastern-and-african-antitrust-review-2018/1145587/germany-merger-control> accessed 16 May 2018
- Commission Nationale de l'Informatique et des Libertés, 'Data Transfer from WhatsApp to Facebook: CNIL Publicly Serves Formal Notice for Lack of Legal Basis' (18 December 2017) <https://www.cnil.fr/en/data-transfer-whatsapp-facebook-cnil-publicly-serves-formal-notice-lack-legal-basis> accessed 02 August 2018
- European Commission, 'Antitrust: Commission fines Google €4.34 billion for illegal practices regarding Android mobile devices to strengthen dominance of Google's search engine' (*European Commission Press Release Database*, 18 July 2018) <http://europa.eu/rapid/press-release_IP-18-4581_en.htm> accessed 23 July 2018
- European Commission, 'Antitrust: Commission probes allegations of antitrust violations by Google' (*European Commission Press Release Database*, 30 November 2010) <http://europa.eu/rapid/press-release_IP-10-1624_en.htm?locale=en> accessed 01 August 2018

- European Commission, ‘Antitrust: Commission takes further steps in investigations alleging Google's comparison shopping and advertising-related practices breach EU rules*’ (*European Commission Press Release Database*, 14 July 2016) <http://europa.eu/rapid/press-release_IP-16-2532_en.htm> accessed 29 July 2018
- European Commission, ‘Consultation on Evaluation of procedural and jurisdictional aspects of EU merger control’ (*European Commission Public Consultations*) <http://ec.europa.eu/competition/consultations/2016_merger_control/index_en.html> accessed 26 July 2018
- European Commission, ‘Summary of replies to the Public Consultation on Evaluation of procedural and jurisdictional aspects of EU merger control’ <http://ec.europa.eu/competition/consultations/2016_merger_control/summary_of_replies_en.pdf> accessed 27 July 2018
- European Data Protection Supervisor, ‘Big Data and Digital Clearing House’ (*European Data Protection Supervisor*) <https://edps.europa.eu/node/3671> accessed 14 May 2018
- European Data Protection Supervisor, ‘Data Protection’ (*European Data Protection Supervisor*) <https://edps.europa.eu/data-protection/data-protection_en> accessed 26 May 2018
- Facebook, ‘Accessing and Downloading Your Facebook Information’ <https://www.facebook.com/help/contact/2032834846972583> accessed 16 June 2018
- Information Commissioner’s Office, ‘Blog: A Win for the Data Protection of UK Consumers’ (14 March 2018) <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2018/03/blog-a-win-for-the-data-protection-of-uk-consumers/> accessed 02 August 2018
- Mayr M, ‘Austria to introduce Transaction Value Merger Notification Threshold’ (*Kluwer Competition Law Blog*, 10 April 2017) <http://competitionlawblog.kluwercompetitionlaw.com/2017/04/10/austria-to-

introduce-transaction-value-merger-notification-threshold/> accessed 22 May 2018

- The Hamburg Commissioner for Data Protection and Freedom of Information, ‘Administrative Order Against the Mass Synchronisation of Data Between Facebook and WhatsApp’ (27 September 2016) <https://datenschutz-hamburg.de/assets/pdf/Press_Release_2016-09-27_Adminstrative_Order_Facebook_WhatsApp.pdf> accessed 02 August 2018

Annex 1 – Competition Law Challenges in the Era of Big Data

Establishment of market power in favor of a few players	Strategic mergers and acquisitions	Exclusionary conducts	Price discrimination between different customer groups	Increased market transparency
•**High entry barriers**: entrants need to access large amounts of *volume* or *variety* of data to be able to compete • **Feedback loops**: more users lead to more data collection, which results in better services and even more users •Aspects to be considered: •**Multi-sided markets**: how to define them? •**Network effects**: higher number of users increases the value of the platform for other users and advertisers •**Multi-homing**: do consumers truly use several providers for the same type of service?	•The ***value*** of data depends on the ***volume, variety*** and ***velocity*** of data • Aim? To obtain **better access to data** • How? By **merging** with companies that previously own large datasets • Consequence? Companies have access to a **greater *amount* of data** that possibly carries more ***diverse*** information with which they can extract ***value*** •The use of data-driven efficiencies as a defense •Updates to merger notification thresholds: •Germany/Austria introduced **'size-of-transaction' threshold** •Discussion at EU level	•Aim? **To deprive competitors from access to data** •How? •**Refusal to access data**: dominant company refuses to share dataset with other company, and such dataset is considered an 'essential facility' •**Discriminatory access to data**: dominant company refuses to share its dataset with one company but is willing to sell it to other companies •**Exclusive agreements**: dominant company enters into agreement with third-party data providers to prevent them from doing business with other firms •**Tying**: dominant company only sells the access to its dataset on the condition that the buyer also purchases another product or service	• Two possibilities: • (i) **Different prices for the same products** even though there are **no differences in the costs** of supplying them • (ii) **Identical prices** even though there are **enough cost differences** that would justify their differentiation • The more a company collects data of its customers, the more information it has about consumer's **purchasing habits and willingness to pay**, and the easier it is for such company to **set individual prices** and **price discriminate**	•Pros? Consumers are able to **compare prices and characteristics** of different goods or services and make more **informed choices** about their purchases •Cons? More information available about competitors' pricing and **higher chances of a more stable collusion between different players**