

The Impact of Data Ownership Rights on Competition in Big Data Markets: Reflections in the Context of the EU and Global Data Economy

Maria José Schmidt-Kessen*

Abstract

Big data and machine learning pose both, great promises for the development of innovative products and services, but also considerable risks ranging from the creation of bottlenecks if vast amounts of data are concentrated in the hands of the few to mass-scale infringements of privacy. Options for regulating big data rank high on the agenda of policy makers all around the world, including in Europe. This paper looks at the debate from a very specific angle: the possibility of creating an IP-akin right in data, by some scholars referred to as “data producer’s right”, which is currently discussed in European policy circles. Which are possible consequences for Big Tech, big data markets and competition from the creation of such a right? This question is discussed first against the backdrop of the EU Digital Single Market Project and various initiatives to foster the EU data economy. Secondly, the impact of such a right for EU trade relations and the global data economy is assessed.

Introduction

We live in a world that is increasingly run by data. As markets go digital, data is becoming the currency for economic success. In this context, governments and policy makers all around the world are puzzling and brooding over how to organize, shape, and foster the data economy. How should good data governance look like? Concerns such as availability and access to data, ownership of data, the impact of data on competition and trade, are just a few of the many issues that need to be addressed.

In a European Data Market Study carried out by the European Commission,

* Ph.D. (European University Institute), Assistant Professor in EU Commercial Law (Copenhagen Business School).

the following definition of “data economy” has been provided:

“The data economy measures the overall impacts of the data market – i.e. the marketplace where digital data is exchanged as products or services derived from raw data – on the economy as a whole. It involves the generation, collection, storage, processing, distribution, analysis, elaboration, delivery, and exploitation of data enabled by digital technologies¹”

One of the widely discussed issues about the data economy is if and how to regulate big data. This paper sheds light upon this debate through the lens of the considerations that have shaped the EU Commission’s policies in this area during the last five years, in particular when it comes to big data and competition policy, and the question of ownership rights, as well as access to data.

After giving an overview over big data and current conceptions of competition on big data markets, the first part of the paper provides an overview over the EU Commission’s main policy documents on the data economy over the last five years (2015-2020). In the second part, the paper discusses the EU legal landscape on the issue of ownership in data, and discusses the impact of the creation of a data producer’s right that has been in the minds of EU policy makers. In the third part, the paper provides reflections on the impact of a data producer’s right on competition in the EU’s internal market. The last part provides reflections on the impact of the introduction of a data producer’s right on the EU’s international trade relations.

The issues discussed in this paper should be considered in a state of flux. As Hugenholtz wrote in 2018 he was aiming at a “moving target” when criticizing the idea to introduce a property right in data.² This situation is no different in 2020. This paper should be read in light of the moment it was written, i.e. in spring 2020, while a global pandemic has raised yet a new host of questions regarding the legal framework and use of big data, and which might trigger a host of new debates and changes in policies regarding the data economy.

Context: Big Data and European Policies for the Data Economy

“Data is the new oil”³ is just one of many metaphors trying to encapsulate

¹ European Data Market Study, SMART 2013/0063, IDC, 2016.

² P. Bernt Hugenholtz, *Data Property: Unwelcome Guest in the House of IP*, in PETER DRAHOS ET. AL., *KRIKA: ESSAYS ON INTELLECTUAL PROPERTY*, VOL. III (Edward Elgar 2018).

³ This metaphor is frequently used, e.g. in *The World’s Most Valuable Resource is No Longer Oil but Data*, THE ECONOMIST (May 6, 2017, 12:16 PM), <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>

the importance that data has gained for businesses, markets, and economic growth globally. It connotes that data has become a raw material, a critical resource, an asset, a tradable commodity, among others.⁴ The development of new technologies, including high-speed internet, the rise of online platform businesses, the improvement of computer processing power, and digital storage capacities, have all made the collection, processing and analysis of massive amounts of data easier and cheaper.

Big data is a commonly used term for large, often unstructured, fast-moving data sets⁵ held by corporations, public authorities and other organizations, which can be analysed with the help of algorithms.⁶ Big data allows for the discovery of previously invisible patterns, correlations, or trends.⁷ It also enables the training of machine learning algorithms, as well as the development of a host of new services and products.⁸

The characteristics that differentiate big data from traditional data are often explained through the 3V,⁹ 4V,¹⁰ or 5V¹¹ model. The Vs refer to big data's volume, variety, velocity (3V), value (4V) and veracity (5V),¹² that set big data

⁴ See, on the other hand, on all the shortcomings of likening data to new oil, e.g. Alec Stapp, *Why Data is Not the New Oil*. TRUTH ON THE MARKET (Oct. 8, 2019, 10:50), <https://truthonthemarket.com/2019/10/08/why-data-is-not-the-new-oil/>

⁵ These three attributes differentiate Big Data from previous approaches of using and analysing data, according to THOMAS H. DAVENPORT, *BIG DATA AT WORK: DISPELLING THE MYTHS, UNCOVERING THE OPPORTUNITIES* 10 (Harvard Business Press 2014).

⁶ There seem to be an endless amount of definitions of big data. This definition is taken from Bart Van Der Sloot & Sascha Van Schendel, *Ten Questions for Future Regulation of Big Data: A Comparative and Empirical Legal Study*, 7 JIPITEC 110-145 (2016).

⁷ Emilio Calvano & Michele Polo, *Market Power, Competition and Innovation in Digital Markets: A Survey*, 54 INFORMATION ECONOMICS AND POLICY 1-18 (2020), <https://doi.org/10.1016/j.infoecopol.2020.100853>

⁸ *Id.*

⁹ This is the oldest model that characterizes big data by volume, velocity and variety is attributed to the research and advisory firm META Group's (now Gartner) analyst Doug Laney who published a research report on "3D Data Management: Controlling Data Volume, Velocity and Variety" in 2001, <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>

¹⁰ Han Hu et al. *Toward Scalable Systems for Big Data Analytics: A Technology Tutorial*. 2 IEEE ACCESS 652-687 (2014).

¹¹ Bernard Marr, *Why One of the 5Vs matters*. IBM BIG DATA & ANALYTICS HUB BLOG (Oct.. 17, 2021, 08:27 PM), <https://www.ibmbigdatahub.com/blog/why-only-one-5-vs-big-data-really-matters>

¹² Some even add a sixth V, viability, which boils down to correctly picking the variables that you will use to build your models out of big data sets. See Niel Biehn, *The Missing Vs in Big*

apart from traditional data. Volume refers to the sheer mass of data gathered, e.g. Facebook registering billions of likes and comments of its users each day.¹³ Variety refers to the “three flavours”¹⁴ in which data can come, namely structured, semi-structured, and unstructured,¹⁵ as well as the possibility of combining many different datasets.¹⁶ Velocity connotes that the data needs to be analysed at the speed of its production, often in real-time, such as the data streams generated by users or by smart sensors, in order to generate value. Value derives from the insights generated from big data sets that can confer commercial benefits and competitive advantage to businesses. This presupposes, however, that businesses and organizations have the expertise and infrastructure to extract this value. Lastly, following computer science wisdom of “rubbish-in-rubbish-out”, the insights and value generated from big data set are directly related to the quality of the data contained in the dataset. Problems relating to the trustworthiness, consistency and completeness in big data require data cleaning and data governance to unlock the potential of big data,¹⁷ which again requires resources and expertise.

When it comes to data and market competition, big data both enables the creation of new business models, and it can enhance the performance of traditional businesses. New business models that make their revenue from online advertising, such as Google’s search engine and social media sites like Facebook and Twitter were the first to understand how to capitalize on big data through combining user profiling, personalization of content, and targeted advertising.¹⁸ Big data is also a key ingredient for the development of technologies that will shape our future, such as cloud computing, AI, smart objects, or the Internet of Things.¹⁹ At the same time, big data is revolutionizing

Data: Viability and Value, WIRED MAGAZINE, (May 20, 2013, 07:24 PM), <https://www.wired.com/insights/2013/05/the-missing-vs-in-big-data-viability-and-value/>

¹³ According to statistics from 2015, Facebook users produce 250 million posts per hour and like around 4 million posts per minute, <https://wersm.com/how-much-data-is-generated-every-minute-on-social-media/>

¹⁴ Han Hu et al. *Toward Scalable Systems for Big Data Analytics: A Technology Tutorial*. 2 IEEE ACCESS 652-687 (2014).

¹⁵ *Id.*

¹⁶ Niel Biehn, *The Missing V's in Big Data: Viability and Value*, WIRED MAGAZINE, (May 20, 2013, 07:24 PM), <https://www.wired.com/insights/2013/05/the-missing-vs-in-big-data-viability-and-value/>

¹⁷ *Id.*

¹⁸ OECD (2015). *Data-Driven Innovation: Big Data for Growth and Well-Being*, available at <https://www.oecd.org/sti/data-driven-innovation-9789264229358-en.htm> , 23-24.

¹⁹ *Id.* at 27-29.

traditional industries, such as logistics, retail and agriculture.²⁰

Competing based on big data appears to be easier for larger market players due to data exhibiting economies of scale.²¹ According to the positive-feedback loop hypothesis,²² firms with a larger installed base of data-generating units (users, consumers, sensors, products...) can amass more data in a shorter period. More data enables better predictions and the discovery of more unexpected patterns and correlations, leading to the development of better products and improved services, and generating a higher rate of innovation. This, in turn, attracts more consumers, which again increases the amount of data available to the firm. Apart from giving data-savvy players an innovation advantage, this virtuous circle also leads to increased expertise and know-how,²³ which helps to attract engineering talent to the company in question.²⁴

These competition dynamics on big data markets have led policy makers, antitrust authorities, and competition scholars to worry that a lot of data in the hands of a few dominant players can create entry barriers for smaller competitors. The idea is that data constitutes an essential facility or input that new and small players lack and which inhibits them from competing effectively. The assumption that data constitutes a barrier to entry hinges upon the question of data substitutability. To which extent is it impossible or prohibitively costly for competitors to replicate or purchase the data held by a large player? It is not possible to answer this question in general terms. The answer is highly case sensitive, depending on the type of data, the industry, and possible legal barriers to sharing certain types of data, among other factors.²⁵ In addition, some studies have challenged the belief that these large returns to scale in data actually exist,²⁶ thus questioning the positive-feedback loop hypothesis and the theory of harm to competition that follows from it.

²⁰ *Id.*

²¹ Gary Biglaiser et al., *Incumbency Advantage and its Value*. 28 J. OF ECONOMICS, MANAGEMENT AND STRATEGY, 41-48 (2019).

²² Emilio Calvano & Michele Polo, *Market Power, Competition and Innovation in Digital Markets: A Survey*, 54 INFORMATION ECONOMICS AND POLICY 1-18 (2020), <https://doi.org/10.1016/j.infoecopol.2020.100853>.

²³ *Id.*

²⁴ Gary Biglaiser et al., *Incumbency Advantage and its Value*. 28 J. OF ECONOMICS, MANAGEMENT AND STRATEGY 41-48 (2019).

²⁵ Daniel L. Rubinfeld & Michal S. Gal, *Access Barriers to Big Data*. 59 ARIZONA L. REV. 339-381 (2017)

²⁶ Patrick Bajari et al., *The Impact of Big Data on Firm Performance: An Empirical Investigation*. 109 AEA PAPERS AND PROCEEDINGS 33-37 (2019).

The possible problem of data as an entry barrier is just one among many policy concerns that policy makers have when it comes to the data economy. Many other questions, such as data mobility and access, data security and integrity, data ownership and control, the protection of personal data and privacy, and data justice, need careful assessment.²⁷

In May 2015, when the EU Commission (the Commission) set out its Digital Single Market Strategy (DSMS) for the European Union, it planned to maximize the growth potential of the EU digital economy.²⁸ One of the core ingredients of the DSMS was to build a data economy.²⁹ The central concern was to remove technical and legislative barriers in order to facilitate the flow of data across borders in the EU. According to the Commission, data location and storage requirements, fragmented legislative frameworks on rights in data, and the lack of interoperable systems for data portability were obstacles that prevent the development of cloud computing, data-driven science, and IoT projects in Europe.³⁰ Furthermore, full harmonization of personal data protection rules among all EU Member States was necessary to allow for the free flow of data while providing a high level of data protection for individuals in the EU.³¹

The next section surveys how the EU's policy towards the data economy has evolved in the years following the announcement of the DSMS. It highlights the issues the Commission has identified as particularly relevant to getting the regulatory approach to the data economy "right". In light of the focus of this paper, particular attention is paid to competition, trade, and data ownership issues in the Commission's policy documents.

The European Data Economy – General Policy Developments Since 2016

The evolution of the policies and legislative initiatives for the data economy in the EU can be classified into two broad categories: (i) measures for personal

²⁷ OECD (2019). Data in the Digital Age, March 2019 available at <https://www.oecd.org/going-digital/data-in-the-digital-age.pdf>.

²⁸ EU COMMISSION (2015). A DIGITAL SINGLE MARKET STRATEGY FOR EUROPE COM (2015) 192 final.

²⁹ *Id.* at § 4.1. There had already been prior policy and legislative initiatives in relation to data that are summarized on the EU Commission's website, see <https://ec.europa.eu/digital-single-market/en/data-policies-and-legislation>.

³⁰ EU COMMISSION (2015). A DIGITAL SINGLE MARKET STRATEGY FOR EUROPE COM (2015) 192 final.

³¹ *Id.*

data and (ii) measures for non-personal data. In relation to personal data, developments can be summarized quickly. The legislative process leading up to the General Data Protection Regulation (GDPR)³² was already moving at full speed when the Commission published its DSMS. In December 2015, the EU legislature³³ reached consensus on the final wording of the GDPR, and its rules became fully applicable on 25 May 2018. The GDPR aims at a high level of protection for data subjects and contains the principles that govern personal data collection and processing by entities based in the EU.³⁴ The GDPR's geographic scope also extends beyond the borders of the EU to any entity processing the data of individuals in the EU.³⁵

When it comes to non-personal data, there have been some legislative developments, but many important aspects are still part of ongoing policy developments. When surveying announcements and communications by the Commission, it quickly becomes evident that the focus of the discussion has been essentially the same over time, albeit with different emphases. The discussion boils down to how to strike a balance between letting market forces work out data governance schemes, and in how far regulatory intervention is necessary to ensure an optimal amount of data access.

In 2016, the EU Commission announced its intention to launch a “European Free Flow of Data Initiative”³⁶ and identified problems that currently hampered the development of the nascent European data economy. It set out to look into a range of related issues, from eliminating restrictions on the free movement of data (other than for the purposes of protection personal data) by national data location requirements to data ownership, data interoperability, usability, and access to data. Regarding data ownership, the EU Commission pointed to the difference between personal and non-personal data. Personal data could not be ‘owned’ under the GDPR regime, but strict rules on access and use by any other person or entity than the data subject applied.³⁷ In respect of non-personal data (i.e. anonymized/non-identifiable personal data

³² Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data.

³³ The proposal for the reform of data protection rules was already introduced by the EU Commission in 2012, which led to the draft of the GDPR. The final text of the GDPR was agreed upon by the European Parliament and the European Council in December 2015.

³⁴ GDPR, art. 3 (1).

³⁵ GDPR, art. 3(2).

³⁶ See, e.g. Inception Impact Assessment by DG CNECT from November 2016.

³⁷ I will elaborate on this point below.

and machine-generated data), there existed a gap in EU law: there was no universal legal regime determining ownership and access to non-personal data. Instead, data access was left to freedom of contract and technological protection measures, and thus determined by and between businesses. The EU Commission noted that leaving data access to contractual arrangements bore the risk of contractual lock-in, potentially hampering the free flow of data between sectors, companies, and within companies.

The 2017 Communication “Building a European Data Economy”

The EU Commission followed up with a Communication on “Building a European Data Economy”³⁸ in 2017, which laid out the Commission’s main considerations and next steps regarding the data economy. The key objective continued to be the facilitation and promotion of access for more players³⁹ to larger and more diverse data sets in order to allow for a higher rate of innovation and the development of data-intensive technologies.

According to the Commission, access to data was hampered by a variety of factors, for example, differences in national laws when it comes to establishing rights over data. While in some EU Member States property claims to data meeting specific conditions were possible, there was no legal instrument at European level that enabled any kind of claims over raw non-personal data (machine-generated or anonymized). While the Commission noted that general contract law and competition law instruments at EU level might offer a sufficient response, these instruments could be insufficient where bargaining power between parties was unequal. In this case, market-based solutions for data access and transfer might not ensure “fair and innovation-friendly results”,⁴⁰ as stronger market players might be able to take control over data and take advantage by imposing unfair standard terms on users or through technological means (e.g. by employing proprietary formats).

Mandating access to data could be a solution to contractual or technological lock-ins. When imposing any access measures, however, the Commission recognized that a balance needed to be struck. Access to anonymous or machine-generated data should be fostered while at the same time ensuring continued protection of personal data (e.g. by preventing anonymised data

³⁸ EU COMMISSION (2017) COMMUNICATION ON BUILDING A EUROPEAN DATA ECONOMY COM (2017) 9 final of 10 January 2017.

³⁹ The EU Commission considers that more access to data would not only be beneficial for businesses, including SMEs, but also for the public sector and for researchers.

⁴⁰ EU COMMISSION (2017) COMMUNICATION ON BUILDING A EUROPEAN DATA ECONOMY COM (2017) 9 final of 10 January 2017.

from being de-anonymized). Access to such data would help overcome differences in bargaining power between market players, but at the same time protection of investment and assets had to be guaranteed. A fair sharing of benefits between data holders, processors and application providers within value chains was desirable. In addition, any disclosure of confidential or sensitive data between competitors had to be avoided, as it could ultimately become a competition problem.

The Commission concluded that a range of measures could help strike the right balance when it came to regulatory approaches ensuring data access. These included fostering the adoption of standards to ensure interoperability of data formats, the creation of a set of default contract rules to avoid unfair terms when it comes to data production and sharing, the development of a FRAND framework for data access, and access for public interest and scientific purposes.

In the context of value chains, the Commission floated the idea of a data producer's right. This right given to long-term users of a device (the "data producer") would allow them to use and authorize the use of non-personal data produced with the device.⁴¹ Such a right could reduce the possibilities of contractual lock-in and allow the user to make use of the data, but it would need to come with clearly specified exceptions, such as non-exclusive access to the data by the manufacturer or by public authorities, for example for traffic management or environmental reasons.

As a 'light' version of the data producer's right, the Commission proposed to create a portability right for non-personal data.⁴² This would enable the data producer to take the data from one system to a competing system. A portability right would reduce switching costs for the user and allow for multi-homing, thereby reducing market entry barriers.⁴³

The Commission committed to gathering more evidence to evaluate these policy suggestions and carried out a public consultation on the ideas presented in the Communication. The result of the public consultation led to a second Communication in 2018.

⁴¹ One could think of examples such as the driver of a connected car being the data producer that has rights over her driving data.

⁴² A portability right for personal data already exists in Article 20 of the GDPR and is discussed below.

⁴³ EU COMMISSION (2017) COMMUNICATION ON BUILDING A EUROPEAN DATA ECONOMY, COM (2017) 9 final.

The 2018 Communication “Towards Building a European Data Space”

In April 2018, the Commission published a second Communication entitled “Towards Building a European Data Space.”⁴⁴ In this Communication, the Commission reflected upon the responses received during the public consultation on its 2017 Communication. It reformulated its policies for the data economy around two corner stones for a Common European Data Space: (i) access and reuse of public sector data and (ii) access and reuse of private sector data.

In respect of facilitating access to public sector data, the Commission proposed substantial amendments to the Public Sector Information (PSI) Directive,⁴⁵ a piece of secondary EU legislation that governs access to public sector information in the EU. The new PSI Directive would reduce or even eliminate charges for access to public-sector data, especially to facilitate access for SMEs. Furthermore, the scope of the Directive was to be widened to bring in new types of data (not only data by public authorities, but also by public undertakings and publicly-funded research) within its scope. Lastly, the new Directive would specify high value data sets to be shared via APIs in real time by public authorities.

In respect of private sector data, the Commission first noted that the industry responses to its public consultation showed that industry stakeholders were not supportive of an ownership right in data, but that it wanted full control over access via freedom of contract. In response, due to the risks posed by e.g. asymmetric bargaining power and contractual lock-in mentioned in the 2017 Communication, the Commission suggested a range of voluntary key principles for contractual arrangements over data. These were in particular applicable to arrangements for IoT objects, and products and services that rely on non-personal machine-generated data created by such objects.

These key contractual principles included transparency, recognition of shared value creation, respect for each other’s commercial interests, undistorted competition, and the minimization of data lock-in. The principle of transparency would require companies to stipulate in advance who will have access to which kind of data, at which level of detail, and for which

⁴⁴ EU COMMISSION (2018). COMMUNICATION TOWARDS A COMMON EUROPEAN DATA SPACE, COM (2018) 232 final.

⁴⁵ Directive 2003/98 on the re-use of public sector information, now replaced by Directive (EU) 2019/1024 on open data and the re-use of public sector information.

purposes it can be used. The recognition of shared value creation would entail the recognition that several parties have contributed to creating the data. Undistorted competition would require parties to be mindful not to exchange commercially sensitive data. Lastly, companies that offer a product or service that generates data as a by-product should allow and enable portability as much as possible, thereby minimizing data lock-in.

Lastly, the Commission also put emphasis on the importance of business-to-government (B2G) data sharing, i.e. enabling access to private sector data for public interest purposes. The EU Commission found that such access could help the public sector in a range of areas such as targeted responses to epidemics, better urban planning, road and traffic control, and market monitoring. In return, private companies should be compensated to allow them to recover investments made in collecting and making the data useful.

The 2017 and 2018 Commission Communications resulted in several pieces of legislation. In November 2018, Regulation 2018/1807 on the free flow of non-personal data in the EU entered into force and prohibited all data localization requirements in the EU other than for purposes of public security.⁴⁶ In 2019, the Open Data or PSI II Directive⁴⁷ entered into force with all the suggestions put forward by the EU Commission in 2018 to enhance access and reuse of public-sector information.

The 2020 European Strategy for Data

In the fall of 2019, after elections of the European Parliament, a newly composed Commission took over the agenda on the European data economy. It announced its approach in the Communication “A European Strategy for Data” on 19 February 2020.⁴⁸ According to the Commission, certain problems in the EU data economy persisted: data sharing between companies had still not taken off, possibly due to lack of trust, imbalances in bargaining power, and fear of misappropriation of data by third parties.⁴⁹ The Commission explicitly mentioned power imbalances between SMEs and large online platforms in terms of access to data. Certain online platforms were in a position to accumulate large amount of data relating to user behaviour and interactions

⁴⁶ Regulation 2018/1807, art. 4(1).

⁴⁷ Directive 2019/2024 on open data and the re-use of public sector information.

⁴⁸ EU COMMISSION (2020). A EUROPEAN STRATEGY FOR DATA COM(2020) 66 final.

⁴⁹ Note how these factors would militate in favour of creating some form of ownership rights in data, at least these were traditionally raised as arguments in favour of creating (intellectual) property rights.

between users of the platform. Such “data advantage” could result in a high degree of market power, which could affect the contestability of the markets on which the platform operated, and could be used by the platform to leverage its power to other markets. In addition, not much progress had been made on interoperability and standards for data quality.

A further challenge that the Commission noticed was that several EU Member States were adapting their national legislation to the data economy, which lead to a fragmentation of regulatory approaches. France and Finland, for example, introduced laws obliging private parties to give access to their data to government authorities.⁵⁰ Finland also introduced a data access regime for scientific research purposes.⁵¹ The German parliament passed an amendment to its national competition law, which clearly includes access to data as a measure for an undertaking’s dominance and introduces new provisions under which the German competition authority can order access to data as a competition remedy.⁵² Such regulatory patchwork puts the vision of a Common European Data Space at risk.

In its strategy, the Commission discusses the possibility of enacting a general piece of legislation in the form of a Data Act (2021) to incentivize horizontal, cross-sectoral data sharing. The Data Act would address a range of business-to-business (B2B) issues when it comes to the sharing of data, such as devising usage and portability rights for co-generated data, and a re-evaluation of the intellectual property framework to further enhance data access and use (especially in the area of the EU database right and trade secrets protection). In principle, data sharing should occur on a voluntary basis. Only in very specific circumstances could data access be mandated,

⁵⁰ French Law no. 2016-1321 for a Digital Republic (*LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique*) allows the public sector to access certain private sector data for public interest purposes. The Finnish Forest Act obliges forest owners to share data on forest management with the public sector.

⁵¹ Finnish Act on the Secondary Use of Health and Social Data, which also created a specific institution, the data permit authority (Findata), which started operating in 2020. See <https://thl.fi/en/web/thlfi-en/statistics/data-and-services/data-permit-authority-findata>.

⁵² The 10th amendment of the German Act on Restraints of Competition (*Gesetz gegen Wettbewerbsbeschränkungen*, short GWB) was passed by the German Parliament in January 2021. It contains provisions which will explicitly allow for an evaluation of market power based on a firm’s access to “competition relevant data” (new §18 (2) new GWB), and a refusal to give access to data where such data is indispensable to compete in an upstream or downstream market will be added as an abuse type to §19 (2) point 4 new GWB. Data advantage will also be an issue for the newly introduced category of dominance for undertakings with “highest relevance across markets for competition” (*Unternehmen mit überragender, marktübergreifender Bedeutung für den Wettbewerb*).

where appropriate under transparent FRAND conditions.

In addition, the Commission set out a large-scale project on nine Common European Data Spaces in strategic sectors and domains of public interest, namely manufacturing, the Green Deal, mobility, health data, energy, agriculture, finance, public administration and skills. These specific Data Spaces will create large data pools with data from the respective areas, combined with the technical tools, infrastructures, and governance mechanisms necessary to enable widespread data sharing within and across data spaces.

Lastly, the Commission also addresses international aspects of its Strategy for Data. It intends to foster international data flows, and to shape global standards for the free flow of data, especially when it comes to the security of data and the protection of personal data. In addition, the Commission plans to address unjustified barriers and digital restrictions imposed by third countries on data flows on a bilateral basis and within the World Trade Organization.

The legislative process has yielded a less ambitious European Data Governance Act, published in November 2020, which governs re-use of certain public sector data, the establishment of trusted data intermediaries, and data altruism.⁵³ The issue of the accumulation of vast amounts of data and resulting data advantages for Big Tech and respective data access obligations are dealt with in a different framework, the Digital Markets Act tabled in December 2020.⁵⁴

The European Data Economy – Specific EU Competition Policy Developments

The Commission also has engaged with the data economy within more specific policy realms, including competition policy. Since the Commission also acts as competition enforcer through its Directorate General for Competition (DG COMP), it commissioned a Report on Competition Policy in the Digital Era in 2019 that attracted a lot of public attention.⁵⁵ The report singles out three

⁵³ Proposal for a regulation on European Data Governance (Data Governance Act), COM (2020) 767 final.

⁵⁴ Proposal for a Regulation on contestable and fair markets in the digital sector (Digital Markets Act) - 2020/0374 (COD).

⁵⁵ Many other competition enforcers, government committees and research centres published competition policy reports for the data and digital economy in 2019, including the UK Competition and Markets Authority's Unlocking Digital Competition Report ("Furman Report"), the German ministry for the Economy's Wettbewerb 4.0 (Competition 4.0) Report, and the Stigler Center's Report on Digital Platforms.

key characteristics of the digital era that affect market competition: (i) the role of data, (ii) the emergence of digital platforms due to economies of scale and network effects, (iii) the rise of digital ecosystems due to economies of scope.⁵⁶

According to the report, any competition assessment of market behaviour involving big data needs to take into consideration that data is highly heterogeneous. Data can be personal or non-personal, individual (one person or machine) or aggregated (several persons or machines). Individual-level data can be volunteered (by users), observed (e.g. by sensors or algorithms tracking user behaviour), or inferred (applying analytics).⁵⁷ Aggregated data is usually standardized, anonymous and thus non-personal, such as statistics about the national economy or the figures in companies' annual reports.⁵⁸ Access to data can be given at single-user, bundled single-user, or aggregate level. Lastly, data can be historical, and can thus be transferred once, or real-time, i.e. it needs to be accessed continuously, e.g. via an API.

The report notes that data is very cheap to reproduce and can be shared easily. In addition, it is valuable to many economic players. This would militate for the broadest possible access and dissemination of data.⁵⁹ At the same time, firms have an incentive to monopolize data as it might give them a certain degree of market power,⁶⁰ and such incentives might be beneficial as they will lead firms to collect data and process it in the first place. These competing efficiencies need to be taken into account when deciding whether to impose access obligations.⁶¹

Whether and how data is accessed through market mechanisms depends heavily on the applicable legal frameworks.⁶² The GDPR already governs access to personal data extensively, and gives individuals a degree of control

⁵⁶ JACQUES CRÉMER, YVES ALEXANDRE DE MONTJOYE & HEIKE SCHWEIZER, COMPETITION POLICY FOR THE DIGITAL ERA 15 (Publications Office of the European Union 2019), available at <https://op.europa.eu/en/publication-detail/-/publication/21dc175c-7b76-11e9-9f05-01aa75ed71a1/language-en>

⁵⁷ The classification of individual-level data is taken from the World Economic Forum. *See* World Economic Forum (2011), *Personal Data: The Emergence of a New Asset Class*, January 2011.

⁵⁸ CRÉMER, ET AL., *supra* note 56 at 26.

⁵⁹ *Id.* at 27.

⁶⁰ *Id.*

⁶¹ *Id.* at 76.

⁶² *Id.* at 73.

over their own data. The legal framework governing non-personal data is still evolving. Since the allocation of rights of data control are determined by contractual arrangements, it is not clear when and in how far companies have access to non-personal data.⁶³ Questions of access under EU competition law will necessarily have to take other existing legal frameworks into account, including the GDPR, IP regimes, etc.⁶⁴

The authors of the report make clear that competition law is not well-equipped to guarantee access or interoperability on an ongoing basis.⁶⁵ Such access rules should be left to sector-specific regulation. Nonetheless, competition law will deal with access questions in specific situations, for example when it comes to data pools and when it comes to dominant firms that abuse their market power.

Data pools to which many firms contribute and have access likely yield considerable efficiencies in scope and scale, as they allow participants access to larger and more diverse data sets as raw material for analysis, new goods and services.⁶⁶ Data pools can thus improve the functioning of markets. On the other hand, data pools might allow for easier sharing of sensitive information and thus lead to collusion between participants.⁶⁷ They might also discourage investment in differentiating or improving data collection techniques if a pool governance framework sets a certain standard.⁶⁸ Furthermore, pool participants may exclude other market players wishing to access the pool, which would potentially call for access based on a FRAND system. These issues would need careful assessment under competition rules.

When it comes to an abuse by a dominant undertaking under Article 102 of the Treaty on the Functioning of the European Union (TFEU), access obligations can be imposed under the so-called essential facilities doctrine under certain circumstances.⁶⁹ The doctrine initially developed in the area of infrastructures and was later extended to determine access obligations to IP-protected

⁶³ *Id.* at 76.

⁶⁴ *Id.* at 73.

⁶⁵ *Id.* at 74.

⁶⁶ *Id.* at 94. In Case C-238/05 *Asnef-Equifax* that concerned data sharing in the insurance sector in the form of registries about individuals' creditworthiness that were accessible to credit providers, the European Court of Justice held that such data sharing yielded important efficiencies.

⁶⁷ CRÉMER, ET AL., *supra* note 56 at 96.

⁶⁸ *Id.* at 97.

⁶⁹ See extensively on this issue, Inge Graef, *EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility*, PhD Thesis KU Leuven (2016), available at https://limo.libis.be/primo-explore/fulldisplay?docid=LIRIAS1711644&context=L&vid=Lirias&search_scope=Lirias&tab=default_tab&lang=en_US&fromSitemap=1.

information held by dominant undertakings.⁷⁰ As the remedy of imposing access obligations is draconian, and infringes on the dominant undertaking's freedom to choose its trading partners, it has been used sparingly, and mostly within the network industries.

In order to oblige a dominant undertaking to grant access to its assets to competitors under the essential facilities doctrine, the requested asset needs to be indispensable in order to compete in a downstream or adjacent market, and not granting access must lead to market foreclosure. Only under these circumstances does a refusal to give access constitute an abuse. Due to indispensability being a determining factor in establishing an abuse, the question around substitutability of data becomes crucial.⁷¹ According to the report, substitutability will very much depend on the type of data. Volunteered data, for example, can probably be volunteered repeatedly (e.g. a user can give her name, e-mail address, or age to several service providers). It is thus often substitutable. When it comes to inferred data, the undertaking in question has put considerable effort into analysing raw data to generate unique insights. Even if not substitutable, inferred data should not be subject to an access obligation, as it would chill investment incentives considerably and would probably entail a sharing of business secrets. The type of data that is not easily substitutable and could be subject to access obligations is observed data, as well as volunteered data that is very costly or impossible to collect again.⁷² The report notes that this analysis is developed against the backdrop that there are no property rights in raw data as EU law currently stands.⁷³ The analysis might look different if the scope and limitations of rights in data by a data controller or a data producer were established. The following sections of this paper discuss this scenario.

Lastly, the report discusses difficulties with designing the access regime under competition rules. First, the requested data might be personal data, which would presuppose that the company wishing access has a consent from the data subjects involved, or any other legitimate basis for processing that data.⁷⁴ Second, giving access to data in real time and in a specific might

⁷⁰ CRÉMER, ET AL., *supra* note 56 at 98.

⁷¹ *Id.* at 101.

⁷² The authors give an individual's calendar data as an example of a type of volunteered data that can be difficult to volunteer again.

⁷³ *Id.* at 96.

⁷⁴ *Id.* at 102. The report gives the example of the UK and French competition authorities that consulted their respective data protection authorities regarding access to personal data and subsequently ordered data access in the energy sector on an opt-out basis. *Id.*

impose considerable costs on the dominant undertaking and its IT systems.⁷⁵ Giving access to standardized, historical data to a downstream competitor once would be a much lower burden than giving access to individual-level data in real time on a constant basis. Third, it might very resource-intensive for competition authorities to monitor compliance with access obligations imposed on dominant undertakings.

Ownership in Data

The survey of the EU Commission's policy documents makes clear that, currently, access and control over data is largely governed by contract. While this is what industry wants,⁷⁶ it seems to lead to sub-optimal levels of data sharing in the eyes of the Commission, and at times, it carries risks of data-lock in and might impair the contestability of markets. Consequently, this paper asks how the situation would change for competition and trade in the data economy if the EU legislature introduced ownership rights over data, in particular in the form of a data producer's right.⁷⁷ Before addressing this question in the following sections of this paper, it is first necessary to assess in how far it would be feasible to introduce such ownership rights in light of existing EU law.

If we think about ownership rights as something attached to property it is worth recalling that property usually entails a right to exclude, as well as rights to acquisition, use, and disposal.⁷⁸ In addition, property rights are *rights in rem* and thus enforceable *erga omnes*, i.e. against third parties. These characteristics of ownership and property rights also apply to IP rights, such as patents and copyright. In contrast to physical property, however, IP rights come with important limitations and exceptions, reducing exclusivity when it is necessary for overriding public interest reasons.

⁷⁵ *Id.*, at 100.

⁷⁶ This is what the Commission learned in its public consultation between the 2017 and 2018 Communications on the data economy, as discussed above.

⁷⁷ The question of ownership rights in data has been extensively discussed in the literature in general. For a particularly thorough overview, see JOSEF DREXL, DATA ACCESS AND CONTROL IN THE ERA OF CONNECTED DEVICES: STUDY ON BEHALF OF THE EUROPEAN CONSUMER ORGANIZATION BEUC (2018). The direct impact of a data producer's right on competition has been discussed to a lesser extent, and on international trade not at all, as far as the author knows.

⁷⁸ Adam Mossoff, *What is Property? Putting the Pieces Back Together*, 45 ARIZ. L. REV. 371-443 (2003).

Types of data for which IP protection exists

Under certain circumstances, it is already possible to protect data under existing IP rights.⁷⁹ Possible candidates for the protection of data would be patent law, copyright, the EU *sui generis* database right and trade secret law. Each of these will be discussed briefly.

The conditions of patentability, novelty, non-obviousness, and usefulness would usually exclude raw data from patent protection. Patent law, however, does offer protection to data sets under very specific circumstances. This could be, for example, data generated through a patent-protected process.⁸⁰ Clinical trial data would be another example of a type of data that can be protected under patent rules.⁸¹

Copyright, at least as conceived in the EU legal system, would not cover data as such,⁸² especially if generated by a machine.⁸³ At the same time, protection under the EU *sui generis* database right, would be possible for structured data sets into which a database maker has put a creative effort and investment into obtaining, verifying, and presenting the contents.⁸⁴ The database right protects against the “extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database”.⁸⁵ Considering, however, that the majority of big data is raw and unstructured means that a large majority of big data is excluded from the

⁷⁹ The availability of IP protection for data has been discussed extensively in the context of EU law, therefore I only briefly summarize the points made in the literature. See P. Bernt Hugenholtz, *Data Property: Unwelcome Guest in the House of IP*, in PETER DRAHOS ET. AL., KRITIKA: ESSAYS ON INTELLECTUAL PROPERTY, VOL. III (Edward Elgar 2018); DREXL, *supra* note 77 ;Daniel Gervais, *Exploring the Interfaces Between Big Data and Intellectual Property Law*, 10 JIPITEC 3 (2019).

⁸⁰ On this possibility see DREXL, *supra* note 77, at 87.

⁸¹ See TRIPS, art. 39(2) and Daniel Gervais, *Exploring the Interfaces Between Big Data and Intellectual Property Law*, 10 JIPITEC 3 (2019).

⁸² Note, however, that copyright protection is available for databases that are original, i.e. when databases “by reason of the selection or arrangement of their contents, constitute the author’s own intellectual creation”, see Article 3 (1) Directive 96/9 on the legal protection of databases.

⁸³ Drexel notes how the lack of human authorship would probably preclude the copyright in databases to arise. See DREXL, *supra* note 77.

⁸⁴ 7(1) Directive 96/9 on the legal protection of databases.

⁸⁵ *Id.*

scope of the European database right.⁸⁶

Lastly, data sets can be protected against misappropriation under trade secret law. While trade secret law can offer an alternative form of legal protection compared to IP rights, it is not a property regime, but a tort-based regime.⁸⁷ Under the EU Trade Secret Directive,⁸⁸ know-how or business information receives protection if it is kept “secret”, i.e. not “generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question”.⁸⁹ In addition it must have commercial value because of its secrecy, and the business claiming trade secret protection must have undertaken reasonable steps to keep it secret.⁹⁰ A trade secret can really be any kind of information, including raw data.⁹¹ While it can be difficult to estimate the threshold for commercial value that a trade secret needs to meet in order to qualify for protection,⁹² data sets that are kept secret could qualify as trade secrets. The limitations of trade secret law is that the holder of the trade secret has no exclusive entitlement to use the information. The holder is only protected against unlawful acts, unlawful acquisition, use, or disclosure.⁹³ Once a data set has lost its secrecy, the holder can no longer rely on trade secret protection to exclude others from using the data.

This section has shown that, to a limited extent, data can be protectable subject matter under existing IP law and related regimes. Nonetheless, for the most part, the type of unstructured data that mainly makes up big data would not be covered by these rights, except trade secrets. In addition, there exists

⁸⁶ Note, that the Database Directive was recently evaluated by the Commission and the Commission also found that “the *sui generis* right does not apply broadly to the data economy (machine generated data, IoT device, big data, AI, etc); it only covers databases that contain data obtained from external sources (for example industries like publishers, who seek out data in order to commercialize databases)”. See EU COMMISSION (2018). SECOND DATABASE DIRECTIVE FINAL EVALUATION REPORT from 25 April 2018, available at <https://ec.europa.eu/digital-single-market/en/news/staff-working-document-and-executive-summary-evaluation-directive-969ec-legal-protection>. No amendments to the EU database regime are currently planned.

⁸⁷ DREXL, *supra* note 77.

⁸⁸ EU law on trade secrets has been harmonized recently under Directive 2016/943 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure.

⁸⁹ Trade Secret Directive 2016/943, art. 2 (1) (a).

⁹⁰ *Id.* at art. 2 (1) (b) and (c)

⁹¹ DREXL, *supra* note 77 at 93.

⁹² *Id.* at 94.

⁹³ Trade Secret Directive 2016/943, art. 3(2) and 4.

a range of EU law instruments containing data access regimes that would make it difficult to claim any kind of exclusive right in data, as discussed in the next section.

Types of data for which ownership is excluded or would be severely limited

The summary of the policy discussion provided above has shown that there exists a considerable amount of access obligations to certain types of data thwarting attempts to create ownership rights over those types of data under EU law. For these types of data, it has been made impossible to categorically exclude others. They can be categorized into three main groups: (i) personal data, (ii) public-sector data, (iii) and certain sector-specific data.

For personal-data, the GDPR has created a portability right for the data subject. Under Article 20 GDPR, the “data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided”.⁹⁴ As a user of a streaming service, for example, this should allow me to receive my data about the preferences I have revealed, my play lists, etc. and have those transferred to another streaming service provider. The portability right can thus counter data lock-in to a certain extent, as it allows users to transfer their data from one provider to another provider. A business thus cannot exclude users from accessing data relating to them and their use of the service. In addition, this portability right can indirectly allow other competitors to access user data.⁹⁵

For public-sector data, the PSI II Directive discussed above provides for non-exclusive access obligations to data held by public authorities, public undertakings, and publicly-funded research.⁹⁶ Access needs to be provided at low or no cost.⁹⁷ For certain data sets falling into the category of “high-value datasets” access even needs to be provided in machine-readable format through

⁹⁴ Note that the portability right under Article 20 GDPR is limited to data processing based on the data subject.

⁹⁵ Obviously, the effectiveness of the portability right to foster access to data depends on other factors, such as network effects and in how far a user will have little incentives to switch to another platform given that many of its social or business contacts are on the platform previously used.

⁹⁶ PSI II Directive, art. 1(1).

⁹⁷ *Id.* at art. 6(1).

APIs.⁹⁸ In addition, the Directive provides that the re-use of public-sector data shall not be made subject to conditions, unless such conditions are objective, proportionate, non-discriminatory, justified on grounds of a public policy objective, and do not restrict competition.⁹⁹ Lastly, the Directive precludes public entities from entering into exclusive agreements with private parties to re-use public data. The only exception to this rule applies when exclusive access is necessary for the provision of a service in the public interest.¹⁰⁰ In this case, the grant of such exclusive access needs to be regularly reviewed, at least every three years. The PSI II Directive thus only allows for a very limited set of circumstances where a private entity might acquire exclusive control over public-sector data for a limited period of time.

When it comes to sector-specific data, a range of regulatory instruments provides access obligations to facilitate competition and innovation. Within the financial sector, in a quest to promote open banking and smaller payment services, users have the right to use third party payment or account services.¹⁰¹ This means that banks must give access to these third party services to the user's data. In the automotive sector, car manufacturers are obliged to give unrestricted and standardized access to vehicle repair and maintenance information, including monitoring and test data, to independent repairers.¹⁰² Another example are smart meters for gas and electricity. Under EU law, competitors in retail electricity and gas markets must be given access to metering and consumption data, data required for customer switching, demand response and other services.¹⁰³ Access obligations also exist for electricity network data to allow for the development of smart grids.¹⁰⁴ These are just some examples, more exist in the area of intelligent transport systems,¹⁰⁵

⁹⁸ *Id.* at art. 14.

⁹⁹ *Id.* at art. 8.

¹⁰⁰ *Id.* at art. 12.

¹⁰¹ Articles 66-67 Directive 2015/2366 on Payment Services in the Internal Market.

¹⁰² Article 6 Regulation 715/2007 as amended by Regulation 595/2009 on type approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6) and on access to vehicle repair and maintenance information.

¹⁰³ Directive 2019/944 on common rules for the internal market for electricity, Directive 2009/73/EC concerning common rules for the internal market in natural gas.

¹⁰⁴ Commission Regulation (EU) 2017/1485 of establishing a guideline on electricity transmission system operation.

¹⁰⁵ Directive 2010/40 on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport.

eCall in-vehicle systems,¹⁰⁶ as well as digital content.¹⁰⁷

Due to existing access obligations, the creation of a right over data would thus not be possible in the EU for the categories of data covered in this section. Such a right could only be created for non-public, non-personal data outside the scope of access obligations in sector-specific regulations. The most relevant area where such right could apply would be probably data produced or collected by smart machines and sensors. Another area could be data collected by online platforms about their business users.

The Data Producer's Right

In 2017, the Commission briefly floated the idea of introducing a data producer's right into EU law.¹⁰⁸ The intellectual idea for such a right can be traced back to 2012, when Herbert Zech provided new impetus to the discussion of protecting information as property.¹⁰⁹ He later used the term "data producer's right" in a 2016 article, where he argues that the creation of a data producer's right could enhance the functioning of data markets, and allow for more widespread and equitable access to data.¹¹⁰

According to Zech, a data producer's right could be designed in the following way: its subject matter would be raw data created by automated means, e.g. through tracking algorithms, sensors etc., or simple computing power.¹¹¹ The owner would be the data producer, i.e. the "economically responsible operator of equipment that generates the data".¹¹² Alternatively, a data producer could

¹⁰⁶ Regulation 2015/758 concerning type-approval requirements for the deployment of the eCall in vehicle system based on the 112 service. On this point, see also Bjorn Lundqvist, *Big Data, Open Data, Privacy Regulations, Intellectual Property and Competition Law in an Internet of Things World – The issue of Access*, Stockholm Faculty of Law Research Paper Series, No.1 (2016).

¹⁰⁷ Directive 2019/770 on certain aspects concerning contracts for the supply of digital content and digital services

¹⁰⁸ See *supra*.

¹⁰⁹ Zech's original habilitation on the topic in 2012 was in German and is published as HERBERT ZECH, INFORMATION ALS SCHUTZGEGENSTAND (Mohr Siebeck 2012). The German title can roughly be translated into 'Information as Property-protected Subject Matter'. Zech subsequently published English articles based on his habilitation, see Herbert Zech, *Information as Property*, 6 JIPITEC 192-197 (2015), and Herbert Zech, *A Legal Framework for a Data Economy in the European Digital Single Market*, 11(6) J. OF INTELL. PROP. L. & PRAC. 460-470 (2016).

¹¹⁰ *Id.* at 460-470.

¹¹¹ *Id.*

¹¹² *Id.*

also be the owner or long-term user of a device as proposed by the Commission in its 2017 Communication.¹¹³ The scope of exclusivity would be the activity of carrying out statistical analysis on that data. Independent reproduction through new measurements would not be covered by the scope of exclusivity. The right would be complemented with non-commercial use exceptions, and other public interest exceptions known from IP law. Zech provides a range of arguments in favour of creating such a right. First, it would further a clear allocation of property rights, which would eliminate uncertainties as to title and therefore improve the functioning of data markets.¹¹⁴ It would also lead to a fairer allocation of profits along value chains, which has been pointed out before as one of the problems brought about by the data economy.¹¹⁵ Lastly, it would facilitate user-access to data.

Zech's idea led to a flurry of critiques and pushbacks by other European IP scholars. Kerber argues that there is no economic justification for introducing such a right. The introduction of an exclusive right over information would usually presuppose an under-production of such information on the market, which would be remedied by the incentives derived from the prospect of an exclusive right.¹¹⁶ Data, however, is anything but under-produced.

Other scholars argue in addition that it would be impossible to make such a right operational. Drexl points out that such a right would significantly increase costs of doing business in the data economy, since market participants would have to engage in massive rights clearance before being able to analyse an aggregated data set from a large stock of data producers.¹¹⁷ In addition, it might be difficult to disentangle personal and non-personal data generated by a device, as these two data types are often collected jointly.¹¹⁸ Furthermore, a data producer's right would have to be paired with an access right, as the data

¹¹³ See *supra* and DREXL, *supra* note 77 at 140.

¹¹⁴ This argument essentially follows the Coase Theorem that finds that markets will produce efficient results if property rights are defined and allocated clearly, and transaction costs are low. The foundation of the Coase Theorem can be found in Ronald H. Coase, *The Problem of Social Cost*. 3 J. OF L. AND ECON. 1-44 (1960).

¹¹⁵ This point is also raised at UN level. See UN (2019). *Data Economy: Radical Transformation or Dystopia?* FRONTIER TECHNOLOGY QUARTERLY, available at https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/FTQ_1_Jan_2019.pdf.

¹¹⁶ This incentive theory is one of the basic economic theories justifying IP rights. See Wolfgang Kerber, *A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis*, GRUR INT. 989, 997(2016).

¹¹⁷ DREXL, *supra* note 77 at 137.

¹¹⁸ *Id.* at 138.

would not be in the hands of the data producer but would probably be stored centrally (e.g. on the servers of the device manufacturer).¹¹⁹ Hugenholtz adds to these critiques that it would be very difficult to establish the subject matter and scope of the right at each moment in time, as, e.g. real-time data is being created constantly.¹²⁰ In addition, it might be difficult to single out the producer of a data set, as multiple actors up and down in the supply chain might be contributing in the process of creating data.¹²¹

Questions would also arise as to derivative data (i.e. inferred data in the terminology used in the 2019 Report on Competition Policy in the Digital Era). Would a data producer also have an entitlement to the proceeds of derivative data, where most value would be generated?¹²² If so, significant chilling effects could occur, as the economic incentives for analysing data by the controller would be diminished if data producer's would have to be compensated. A data producer's right should thus not be extended to derivative data.

A solution to the problem presented by the critics of a data producer's right would be to understand the right not as an ownership right or a new form of IP. Rather, one could understand the data producer's right similar to the data subject's portability right for personal data under the GDPR, i.e. and access and re-use right.¹²³ In this case, the right to access and portability would cover non-personal data of a device that is owned or used in the long-term by the data producers. While there exists no right to portability of non-personal data to date in the EU legal system,¹²⁴ it could be introduced as part of the Data Act that the Commission proposes in its 2020 Data Strategy (which the legislative

¹¹⁹ *Id.* at 142. Note however, that this might be shifting as we move to a fringe internet, where data is stored at the user's end, too (see EU Commission 2020 European Data Strategy).

¹²⁰ P. Bernt Hugenholtz, *Data Property: Unwelcome Guest in the House of IP*, in PETER DRAHOS ET. AL., *KRITIKA: ESSAYS ON INTELLECTUAL PROPERTY*, VOL. III (Edward Elgar 2018).

¹²¹ *Id.* On this point see also OECD (2015). *Data-Driven Innovation: Big Data for Growth and Well-Being*, available at <https://www.oecd.org/sti/data-driven-innovation-9789264229358-en.htm>, 195-196.

¹²² DREXL, *supra* note 77 at 143-144.

¹²³ For a similar suggestion see *id.* at 154.

¹²⁴ CRÉMER, ET AL., *supra* note 56 at 188.

Article 6 of Regulation 2018/1807 on the Free Flow of Non-Personal Data contains an encouragement for the EU Commission to facilitate the development of self-regulatory codes of conduct for the porting of data, which, however, do not strictly speaking provide for an access regime. On this point, see also Bjorn Lundqvist, *Portability in Datasets under Intellectual Property, Competition Law, and Blockchain*, Faculty of Law, Stockholm University Research Paper No. 62 (2018), available at SSRN: <https://ssrn.com/abstract=3278580>.

proposal for the Data Governance Act does not address).

The Impact of Data Producer's Rights on Competition Law in the EU

The impact of data producer's on competition law would depend on whether it is understood as an IP-akin ownership right, or rather as a mere access and portability right. In either of the two cases it could lead to market power of a dominant firm translating less directly into a data advantage. In both cases, a dominant undertaking would not be able by default to control data exclusively via contract and technological protection measures. Such an effect from the data producer's right would be limited to the data being accumulated through a device or a service used long-term by a data producer, however, and would not extend to data advantages gained by analysing the data and producing inferred data. Any variation of the data producer's right would thus not affect data advantage and the resulting market power from superior analytics capabilities. This would be desirable from the perspective of maintaining incentives to innovate and invest in analytics capabilities.

Critics may argue that dominant players continue to hold more bargaining power than data producers and might simply ask data producers to assign their rights to them by contract, depriving the data producer's right from any actual effect on the market.¹²⁵ Two arguments could be held against the argument that nothing would change with a data producer's right. First, the simple fact that the data producer's right exists might improve the bargaining situation of data producers compared to the right not being enshrined in legislation.¹²⁶ Second, competition authorities would have a blueprint to assess the limits and boundaries of access rights by the law, which might make it easier to define and impose access obligations under competition rules when it comes to abuses of dominance, e.g. in the framework of the essential facilities doctrine,¹²⁷ or when determining whether a dominant undertaking used unfair business practices in its contractual arrangements covering data.¹²⁸

¹²⁵ DREXL, *supra* note 77 at 139.

¹²⁶ In addition, the data producer's right in its portability version could simply be made non-waivable. See on this point DREXL, *supra* note 77 at 156.

¹²⁷ CRÉMER, ET AL., *supra* note 56 at 99. (the Report notes that "Any legal definition of property rights is accompanied by a specification of their boundaries and limitations. Where no such property rights exist, the boundaries and limitations of the rights of a data controller are missing. Consequently, the responsibility to define such limitations is partly shifted to competition law, which should balance considerations of protection of investments and greater competition."). This could also facilitate access obligations remedy in merger cases where a merger would lead to high levels of market concentration and a higher risk of market foreclosure.

¹²⁸ Under Article 102 (a) TFEU, "directly or indirectly imposing unfair purchase or selling prices or other unfair trading conditions" constitutes a form of abuse of dominance.

A data producer's right would allow a business that helped produce non-personal data, e.g. through the use of a machine or a service provided by an upstream player or an online platform, to take their data and either use it to generate their own insights, develop their own products or services, or license it to a third party. Under these circumstances, the risk of data-lock in would be lower, and it would allow device and service users to multi-home or facilitate the purchase of complementary goods and services from independent providers. Especially if the data producer's right is understood as a portability right, it should thus improve the free flow of data, the working of markets, the distribution of the benefits from creating data, and the rate of innovation.

Some effects on competition would differ depending on whether the data producer's right were to be introduced as an IP-akin type of right compared to a mere right to access and portability. The most important difference would lie in the fact that an IP-akin type of a data producer's right would necessarily require a component of exclusivity. This would entail, for example, that the data producer needs to license to, e.g. a device manufacturer, the data producer's right in order to allow for the manufacturer to pool and analyse the data. The data producer's right in its IP-akin version would thus increase the possibility for data producers to exploit the data produced by their devices, but it would make it more difficult for the manufacturer to analyse its entire data set covering all devices or to allow third parties to do so. As a result, a data producer's right understood as ownership or property, while increasing individual-level access to data, might reduce the overall access to data on the market.¹²⁹ The access and portability right for a data producer would not imply exclusivity and would thus not carry the risk of re-establishing data lock-ins in a different form. From a competition perspective, the latter is thus to be preferred over the former.

Impact of a Data Producer's Right on EU Trade Relations

The entire paper has so far mainly referred to the EU data economy, as if it were a self-contained system. The EU data economy, however, is highly interconnected with the global data economy, including international trade flows. Which impact would a data producer's right have on the EU's trade relations?

When it comes to the global data economy, the Commission seems to suggest that the EU is trying to build a different system compared to the US and China.

¹²⁹ DREXL, *supra* note 77, at 143.

The Commission notes in its 2020 Data Strategy that the US has adopted a system where the private sector is left free to organize the data economy “with considerable concentration effects”.¹³⁰ When it comes to China, the Commission notes pervasive “government surveillance with strong control of Big Tech Companies over massive amounts of data without sufficient safeguards for individuals”. The Commission seems to suggest that the EU is striving to make the data economy flourish in a different, third way, without massive privacy invasions on the one hand, and on the other without allowing excessive concentration of markets and accumulation of private economic power when it comes to data-driven business.

When it comes to personal data and the GDPR applying beyond EU borders, the Commission has already imposed its regulatory standards on businesses wanting to offer goods and services that involve the handling of personal data in the EU. In addition, the EU has reached out to governments in other parts of the world to persuade them to adopt similar regulatory regimes for the processing of personal data in their own jurisdictions as it did with the GDPR.¹³¹ As the Commission notes in its 2020 Data Strategy, it has engaged in “digital diplomacy recognizing 13 countries as providing adequate level of protection for personal data”.¹³² This statement does not prove whether all of these jurisdictions also offer a portability right for non-personal data as under Article 20 GDPR, which would make for an interesting comparative study, but it could be possible that other jurisdictions have such portability rights in their regulatory schemes for personal data.

The creation of a data producer’s right in the EU, given the global reach of supply chains and interconnectedness of various businesses when offering e.g. smart objects, or platform services, would likely need some form of international backing in order to be effective. This could entail, for example, a portability provision similar to the GDPR in non-personal data with a scope of application extending outside the EU. It could also mean digital diplomacy and advocacy in bilateral and multilateral talks. The US and Japan Digital Trade Agreement signed in November 2019 could offer a blueprint, as it contains provision on the free flow of data, a prohibition of national data

¹³⁰ EU COMMISSION (2020). A EUROPEAN STRATEGY FOR DATA COM(2020) 66 final.

¹³¹ Anu Bradford refers to this as the “Brussels Effect”, and uses data protection as one case study. See ANU BRADFORD, *THE BRUSSELS EFFECT - HOW THE EUROPEAN UNION RULES THE WORLD* (Oxford University Press 2020).

¹³² EU COMMISSION (2020). A EUROPEAN STRATEGY FOR DATA COM (2020) 66 final.

location requirements, and provisions on the protection of personal data.¹³³ Lastly, the EU could try to bring the suggestion to the WTO to introduce such a right through a multilateral treaty on trade in data that could be brought under the auspices of the WTO.¹³⁴

Conclusion

How should good big data regulation look like? This paper could not give an overall answer to this very big question, but analysed smaller sub-questions. It looked at it from a perspective of competition law, data ownership, and data access and concretely asked to which extent the introduction of a data producer's right would be possible as EU law currently stands, and which impact it would have on competition law in the EU and on the EU's international trade relations.

The paper made clear that a data producer's right could only be introduced in relation to non-public, non-personal data outside the scope of access obligations in sector-specific regulations. Given that the version of the data producer's right in the form of an IP-akin property right as received a host of justified criticism, a lighter version of that right in the form of an access and portability right would be desirable. Such a right would be beneficial for competition, as it would reduce the risk of data-lock in, improve the free flow of data, and therefore the functioning of data markets. In order to fully unfold its positive effects, however, the data producer's right would need to be promoted also in international trade relations. The EU could use either bilateral or multilateral trade diplomacy to achieve this.

¹³³ See Articles 11, 12, and 15. The text of the Agreement is available at https://ustr.gov/sites/default/files/files/agreements/japan/Agreement_between_the_United_States_and_Japan_concerning_Digital_Trade.pdf

¹³⁴ For a taxonomy of measures restricting cross-border data flows and an initial assessment under WTO Law see Martina Ferracane, *Restrictions on Cross-Border Data Flows: A Taxonomy* (November 18, 2017). ECIPE Working Paper No. 1/2017, Available at SSRN: <https://ssrn.com/abstract=3089956> or <http://dx.doi.org/10.2139/ssrn.3089956> .