Designing Competitive Markets for Industrial Data

Between Propertisation and Access

by **Josef Drexl***

Abstract: As part of the project to establish a Digital Single Market, the European Commission has launched a 'Free Flow of Data' initiative. This initiative is meant to enhance the growth potential of the emerging data economy, which is characterised by the digitisation of production (smart factories) and the advent of digitised products such as smart—driverless—cars, or smart wearables that will be able to communicate with each other and the environment through the Internet of Things. Furthermore, the enormous amount of data generated and controlled by the industry could serve as a most valuable input for other new data-driven services and for applications in the public interest, such as the operation of smart cities, smart and resource-efficient farming, or measures to prevent the spread of infectious diseases. Obviously, this new data economy has to rely on the commercialisation of data. But what kind of regulation is needed in order to make the data economy work? Do we need new ownership rights in data? Or should regulation focus on access in order to make data as widely available as possible? The European Commission is currently trying to formulate answers to these questions. This article aims to assist the Commission by working on a pro-competitive framework for issues of both ownership and access. In so doing, this article undertakes two things: first, it analyses to what extent intellectual property laws already provide control over data and then discusses the need and justification for introducing new rules on data ownership. Second, it analyses whether EU competition law already provides remedies to promote access to data, and furthermore explores whether and under which conditions the introduction of new access regimes would be advisable. This article is to be considered as ongoing research. It does not yet take into account more recent developments in 2017.

Keywords: Data ownership; access to data; data sharing; data economy; data-driven economy; Internet of Things; data analytics; database rights; trade secrets protection; EU competition law; refusal to license; essential facilities; data portability

Recommended citation: Josef Drexl, Designing Competitive Markets for Industrial Data – Between Propertisation and Access, 8 (2017) JIPITEC 257 para 1.

A. Introduction

1 The advent of the data economy and the Internet of Things (IoT) is currently challenging regulators across the globe. Buzzwords such as 'big data' or 'data as the oil of the modern economy' are used everywhere, and questions like 'Who owns the data?' are not only asked by the media, but are also heard and taken up by decision-makers in the political arena.

2 In the EU, potential new regulation for the data economy, concerning both data ownership and access to data, is part of the Commission's current priority project to implement a Digital Single Market.[1] In May 2015, the Commission identified 16 key actions for the implementation of this Digital Single Market,[2] including the 'building of a data economy'. This 'action' is supposed to contribute to the third pillar of the Digital Single Market project, aiming at 'maximising the growth potential of the digital economy'.[3] More concretely, the Commission announced a 'Free Flow of Data' initiative for 2016, which would address in particular the restrictions on the free movement of data beyond the protection of personal data with the objective of enhancing the cross-border use of data in a world of big data and the Internet of Things. Yet the initiative also includes a mandate to look at the issue of ownership. The announcement reads as follows:

> *The Commission will propose in 2016 a European 'Free flow of data' initiative that tackles restrictions on the free movement of data for reasons other than the protection of personal data within the EU and unjustified restrictions on the location of data for storage or processing purposes. It will address the emerging issues of ownership, interoperability, usability and access to data in situations such as business-to-business, business to consumer, machine generated and machine-to-machine data. It will encourage access to public data to help drive innovation.*[4]

3 As regards ownership, the mandate does not indicate the direction in which later regulatory actions may ultimately go. In the light of the objective to promote access to data, one could expect the Commission to consider whether existing 'ownership' regimes are in need of additional exceptions and limitations to promote access. This would have been in line with the debate in other fora, such as OECD in particular, where a study of 2015 highlighted the need to promote access to big data in order to generate maximum benefits for society.[5] Rather than taking data ownership as the starting point of the regulation of the data economy, the OECD study recommends developing and improving 'data governance regimes' that 'overcome ... barriers to data access, sharing and operability'.[6]

4 As regards the EU, however, the debate quickly shifted direction. While the responsibility to work on the initiative was allocated to the Digital Value Chain unit of DG CONNECT, which is also responsible for the open data policy of the EU as regards public sector information in particular, it was the German Commissioner Günther Oettinger responsible for DG CONNECT who publicly contributed to the impression that the Commission would soon propose legislation

* Dr iur (Munich), LLM (UC Berkeley), Director of the Max Planck Institute for Innovation and Competition, Munich, Honorary Professor at the University of Munich.

This article complements the Position Statement of the Max Planck Institute: Josef Drexl, Reto M Hilty, Luc Desaunettes, Franziska Greiner, Daria Kim, Heiko Richter and Gintarė Surblytė, 'Data Ownership and Access to Data' (16 August 2016), available at: <http://www.ip.mpg.de/en/link/positionpaper-data-2016-08-16.html> (accessed 10 September 2016). The views expressed in this article are however only those of its author.

This article was first made available online as Research Paper No. 16-13 of the Max Planck Institute for Innovation and Competition Research Paper series on 8 November 2016. The text remains substantially unchanged. It does not take into account the debate following the EU Commission's Communication of 10 January 2017 on Building a European Data Economy, COM(2017) 9 final. On this Communication see the Position Statement of the Max Planck Institute: Josef Drexl, Reto M. Hilty, Jure Globocnik, Franziska Greiner, Daria Kim, Heiko Richter, Peter R. Slowinski, Gintare Surblyte, Axel Walz and Klaus Wiedemann, 'On the European Commission's Public Consultation on "Building a European Data Economy"' (26 April 2017) available at: <http://www.ip.mpg.de/fileadmin/ipmpg/content/stellungnahmen/MPI_Statement_Public_consultation_on_Building_the_EU_Data_Eco_28042017.pdf> (accessed 19 October 2017). On this, see also Josef Drexl, 'On the Future Legal Framework for the Digital Economy: A Competition-based Response to the "Ownership and Access" Debate', in Reiner Schulze and Dirk Staudenmayer (eds), *Trading Data in the Digital Economy: Legal Concepts and Tools* (Baden-Baden: Nomos, forthcoming) 222-43.

1 Implementation of the Digital Single Market is one of four 'priority projects' of the current European Commission under the aegis of President Jean-Claude Juncker. See Jean-Claude Juncker, 'My priorities', available at: <http://juncker.epp.eu/sites/default/files/attachments/nodes/en_01_main.pdf> (accessed 10 September 2016).

2 See Communication of the Commission of 6 May 2015 to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions—A Digital Single Market Strategy for Europe, COM(2015) 192 final. See also European Commission, 'A Digital Single Market for Europe: Commission sets out 16 initiatives to make it happen', Press Release of 6 May 2015, available at: <http://europa.eu/rapid/press-release_IP-15-4919_en.htm> (accessed 10 September 2016).

3 See Chapter 4.1 of the Commission Communication (*supra* n 2) at 14-15.

4 *Ibid*, at 15. (Emphasis added.)

5 OECD, 'Data-Driven Innovation: Big Data for Growth and Well-Being' (2015) 195-98, available at: <http://www.oecd.org/sti/data-driven-innovation-9789264229358-en.htm> (accessed 10 September 2016).

6 *Ibid*, at 195-99 (in particular at 198).

on a new 'data use right' (*Datennutzungsrecht*).[7]

5 The data economy and its regulation attract particular attention in Germany, where the industry is deeply involved in the development of new business models of the Internet of Things. In Germany, in 2011, the 'Industrie 4.0' initiative was launched as a joint initiative of the government, the private business sector and the public research sector to manage and promote a fourth industrial revolution characterised by the integration of manufacturing in modern information and telecommunications networks, including the Internet of Things.[8] This initiative not only aims at optimising the manufacturing process, whereby the product itself, in the various production phases, communicates with, and steers, the production process. It also targets the logistics sector, aiming to foster an 'Internet of Services' that builds on smart products as a basis for new kinds of services provided to consumers. This early initiative may also explain why, in Germany, legal regulation of the industrial dimension of the data-driven economy, namely, beyond the issues of protection of personal data in particular, attracted attention much earlier than in other parts of the EU, both from the academic community[9] and from the stakeholders' side. As regards the latter, the *Bundesverband der Deutschen Industrie* (BDI, German Industry Association) published a study on the legal ramifications of the data-driven economy that, *inter alia*, argued against the introduction of a new right of data ownership.[10] A report of the Bavarian Industry Association (*Vereinigung der Bayerischen Wirtschaft*) argued that ownership for single pieces of data and small datasets could lead to a scarcity of data and distort innovation through big data analytics.[11]

6 Indeed, scepticism about introducing a new intellectual property right expressed by the industry that is expected to rely on this right for protecting its own investments is not something that experts in intellectual property law would necessarily expect. However, the same scepticism was voiced by the representatives of the 'Industry 4.0' sector who were invited to a hearing of DG CONNECT on the 'legal regime fit for an efficient and fair access to and usage and exchange of data' in Luxembourg on 17 March 2016.[12] The hearing concentrated on the legal protection of the investment in data collection capabilities and the exploitation of the value represented by that data. The hearing was not least held for the purpose of learning more about the legal instruments that are used and needed to

7 See, for instance, 'Oettinger: Versicherungen brauchen mehr digitale Produkte', *Der Standard* (25 November 2015), available at: <http://derstandard.at/2000026414259/Oettinger-Versicherungen-brauchen-mehr-digitale-Produkte> (accessed 20 May 2016) (reporting on a talk by the Commissioner at a conference of the German insurance industry association in November 2015 where the Commissioner called upon the insurance industry to take part in the discussion on such a right). See also the association's website: 'Versicherungstag 2015: Es geht mehr denn je um den Kunden' (25 November 2015), available at: <http://www.gdv.de/2015/11/versicherungstag-2015-chancen-der-digitalen-welt/> (accessed 10 September 2016). The author of this paper personally attended another talk given by the Commissioner at a conference of the Forschungsinstitut für Wirtschaftsverfassung und Wettbewerb (FIW) in Innsbruck on 25 February 2016, where the Commissioner made similar statements. See 'Rede (Kommissar Oettinger) auf dem 49. FIW-Symposion (2016) in Innsbruck zur Digitalisierung' (25 February 2016), available at: <http://www.fiw-online.de/de/aktuelles/aktuelles/rede-kommissar-oettinger-auf-dem-49.-fiw-symposion-2016-in-innsbruck-zur-digitalisierung> (accessed 10 September 2016) (reporting on the Commissioner asking who owns the data that are produced by modern cars in a world of the Internet of Things). In a more recent speech at the occasion of a Commission conference on the 'Free Flow of Data' initiative, however, the Commissioner did not repeat this claim for a data usage right. See Günther Oettinger, Speech at the Conference 'Building European Data Economy' (17 October 2016), available at: <https://ec.europa.eu/commission/2014-2019/oettinger/announcements/speech-conference-building-european-data-economy_en> (accessed 30 October 2016).

8 See the public announcement of the initiative made on the occasion of the 2011 Hanover trade fair: Henning Kagermann, Wolf-Dieter Lukas and Wolfgang Wahlster, 'Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution' (1 April 2011), available at: <http://www.vdi-nachrichten.com/Technik-Gesellschaft/Industrie-40-Mit-Internet-Dinge-Weg-4-industriellen-Revolution> (accessed 10 September 2016).

9 More in favour of such a right Herbert Zech, 'Daten als Wirtschaftsgut—Überlegungen zu einem "Recht des Datenerzeugers"' (2015) *Computer und Recht* 137; most recently see Herbert Zech, 'A legal framework for a data economy in the European Digital Single Market: rights to use data' (2016) 11 *J Int Prop L & Prac* 460; Herbert Zech, 'Data as tradeable commodity' in Alberto de Franceschi (ed.), *European Contract Law and the Digital Single Market* (Insentia: 2016 forthcoming) 51; against such a right Michael Dorner, 'Big Data und "Dateneigentum", Grundfragen des modernen Daten- und Informationshandels' (2014) *Computer und Recht* 617. See also Alberto De Franceschi and Michael Lehmann, 'Data as Tradable Commodity and New Measures for their Protection' (2015) 51 *Italian LJ* 51 (seemingly supporting the recognition of a 'data usage right').

10 Konrad Żdanowiecki, 'Recht an Daten' in Peter Bräutigam and Thomas Klindt (eds), *Digitalisierte Wirtschaft/Industrie 4.0*, Study of Noerr LLP for BDI (November 2015) 18-28, available at: <http://bdi.eu/media/themenfelder/digitalisierung/downloads/20151117_Digitalisierte_Wirtschaft_Industrie_40_Gutachten_der_Noerr_LLP.pdf> (accessed 10 September 2016).

11 Zukunftsrat der Bayerischen Wirtschaft, 'Zukunft digital—Big Data: Analyse und Handlungsempfehlungen (July 2016) at 99, available at: <https://www.vbw-zukunftsrat.de/pdf/big_data/vbw_zukunftsrat_handlungsempfehlungen_langfassung_v15_rz_web.pdf> (accessed 10 September 2016).

12 The author of this paper took part in this 'Round Table' as a representative from the academic community. The results of the event are documented in a synthesis report not publicly available of the Unit for Data Value Chain (available from the author).

implement new business models based on big data. Unanimously, the industry participants stressed that they were able to implement their business models involving data-sharing by relying on contract law. 'Ownership' was even considered a concept that does not fit the needs of the data economy; introduction of a new right was seen as a form of government intervention that needs to be avoided. At the same time the need to promote access, with a potential role of competition law, was discussed. Ultimately, the Digital Value Chain Unit's representative indicated that the Commission would come forward with policy conclusions in the form of a Communication, which was published in January 2017.[13]

7 There seem to be two obvious, yet related, reasons why the industry rejects the introduction of new property rights for data: first, many firms are producers of data and have to rely on access to data of other players at the same time. Hence, it is not clear to them whether the introduction of new rights would provide them with more benefits than drawbacks. Second, the criteria on who would qualify as the owner of the new right are not at all clear. Many stakeholders, in one way or another, contribute to the same data-based business model and may have very diverse kinds of interests. Therefore, allocation of data ownership is indeed a major issue.[14] This is also an issue of considerable complexity because of the particularities of the specific sectors. The interests of stakeholders regarding the data collected by the sensors of a car, in which public authorities also have an interest, so as to protect the environment or to increase driving safety, are likely to be different than those in the case of health-related data derived from blood tests of patients for which a patented diagnostic tool is used, which, taken together with similar data from other labs, may help authorities around the globe to fight the spread of infectious diseases. The difficult question to whom the new data ownership should be allocated led the BDI study to conclude that the legislature should refrain from creating such a right from the outset.[15] In such a situation it should not come as a surprise that firms, which cannot foresee, and do not have any legitimate expectation, that they will be recognised as owners of data rights, will be hesitant to support any additional legislation. If it was accepted that there should be ownership of everybody to whom specific data can be allocated, the result would be multiple ownership of the same data with considerable negative effects on access to that data.[16]

8 This article aims to produce additional insights on how the data economy should be regulated as regards data collected by the industry. Ideally, it will also assist the European Commission in its task of designing its regulatory approach to promoting the data economy in the interest of society. For that purpose, the article looks at the issues of both data ownership and access to data.

9 As a starting point, this article argues that the question 'Who owns the data?' is fundamentally misguided. This is so for two reasons: first, it skips the prior question of whether there is a need to recognise any ownership. There is no natural law that says that data as an asset, although it may have economic value, has to be owned by anybody. Rather, recognition of any new right should, as is the case in intellectual property in general, be considered a form of government regulation of the market, which is in need of a particular justification. In terms of data ownership, which enables its owner to commercialise data, this justification needs to be an economic one.[17]

10 The second reason is that identifying the owner does not resolve all issues of ownership. In the field of intellectual property law, the legislature has to decide upon a series of issues: first and foremost, the subject-matter of protection has to be determined. Hence, the law would have to clarify what is meant by 'data' in the context of 'data ownership'. And then there is the issue of 'how' ownership should be protected. In other words, the legislature has to decide on the scope of protection—namely, what kind of interests and uses are protected— whether there are certain exceptions and limitations that

13 European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of The Regions Building A European Data Economy*, COM (2017) 9 final, Bruxelles, 10.1.2017. As mentioned, this article does not yet discuss this Communication. For further references see at n * above.

14 The OCED (*supra* n 5) at 196, lists ten different kinds of stakeholders. It thereby relied on literature—David Loshin, 'Knowledge integrity: Data ownership' (2002) (no longer available on the Internet)—that predates the big data debate and, in particular, does not yet take account of big data analyses and big data brokerage.

15 Żdanowiecki (*supra* n 10) at 28.

16 This could be considered a situation of a 'tragedy of the anti-commons' in which too many property rights in the same asset lead to inefficient underuse of that asset. See Michael A. Heller, 'Tragedy of the Anti-Commons: Property in the Transition from Marx to Markets' (1998) 111 *Harv L Rev* 621.

17 This distinguishes 'data ownership' from the protection of personal data. It is to be noted that data protection rules in the EU only protect natural persons. See 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 on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), [2016] OJ L 119/1. Corporate entities may also have an interest in keeping back information that has the potential of harming their 'corporate reputation'. However, this can be seen as part of their commercial interests. In this context, trade secrets rules may provide some protection. On this, see at C.II. below.

take into account conflicting interests and, finally, which remedies will be made available to the right-holder. In making such decisions on the framing of the new right's regime, the economic arguments that justify the recognition of a new right as such have to play a key role.

11 In addition, any new legislation on data ownership should take into account the public interest in maintaining competition in the market. Additional rights regarding data as an asset may enhance market power deriving from the control of data. As in other fields of intellectual property law, the guidepost should be that both property rights and competition pursue the goal of enhancing innovation.[18] If the data ownership right is supposed to create incentives to invest in new data-based business models by controlling the use of data, and if competition is designed to maintain competitive pressure on the right-holder to maintain its incentives to invest, the best approach will be to take the competition dimension into account as a core consideration for the design of the property rules. This approach has the advantage of reducing the need for later reliance on competition law as a countervailing legal regime. Accordingly, the interest in maintaining access to data in the interest of society would have to be one of the criteria that guide any future legislation on data ownership.

12 In the following, the article will first take a look at the phenomenon of the emerging data economy and how value is generated in that economy (section B. below). Then, it will explore to what extent there is already control over data, in the form of either factual control or legal control based on specific protection regimes (section C. below). Against this backdrop, it will be possible to discuss whether and to what extent there is an economic justification for additional protection (section D. below). Furthermore, the article will explore the different issues concerning the design of an additional protection regime (section E. below). Yet the analysis is not limited to the question of whether additional ownership rights are needed. Rather, in part F. this article will analyse and discuss legal regimes, including competition law and more targeted forms of legislation, to enhance access to data in order to promote a pro-competitive data economy.

18 See Communication from the Commission—Guidelines on the application of Article 101 of the Treaty on the Functioning of the European Union to technology transfer agreements, [2014] OJ C 89/3, para 7.

B. The phenomenon of the data economy

13 For the purpose of this article, a number of particular features of the data economy need to be understood properly in order to answer the policy issues surrounding data ownership. This includes a description of the use of data as an asset in the data economy and the economic and societal benefits of that economy (part B.I. below), the phenomena of 'data' and 'big data' in this context (part B.II. below), specific features of how value is generated in this economy (part B.III. below), and finally the interests of specific stakeholders that need to be taken into account in designing any future legislative action (part B.IV. below).

14 All of these issues are closely linked to new business models that are currently evolving in very diverse sectors of the data economy. This means that the following analysis has to do with very dynamic phenomena of high complexity and variety. Anybody who engages in this topic has to understand what is actually going on in the market concerning the underlying business models; also, generalisations need to be considered with caution. This is already an important lesson for the legislature. Any rule that is adopted against the backdrop of one case scenario also has to fit other scenarios to which it may apply. In addition, property legislation in particular should not only respond to the needs of today's economy, but also the needs of tomorrow. This argues against precipitate legislative action, despite the enormous speed of the development of the data economy, at least as regards the recognition of new property rights without a clear understanding of the business models that will be affected now and in the future. Such new rights have the potential of increasing market power, creating barriers to access to important data and, ultimately, curbing rather than fostering the data economy.[19]

19 An example of such premature legislation was the introduction of the *sui generis* database right by the EU legislature in 1996. See Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, [1996] OJ L 77/20. Ten years after its adoption, the Commission had to admit that there was no evidence that the Directive had indeed produced the expected positive economic effects as regards the information market in the EU. The Commission even considered a withdrawal of that protection, without, however, recommending it. See DG Internal Market and Services Working Paper—First evaluation of Directive 96/9/EC on the legal protection of databases (12 December 2005), available at: <http://ec.europa.eu/internal_market/copyright/docs/databases/evaluation_report_en.pdf> (accessed 10 September 2016).

I. Data as a most important asset of the data economy and the societal benefits deriving from it

15 In the data economy, data have become the key asset for conducting business. This explains why data are often called the 'oil' of the new economy.

16 Beyond the use of this buzzword, it is more important to understand why and how data are used. Different forms of use relate to different stages of the development of the Internet. At its first stage, the Internet was used as a tool for providing information. This was the time when politics started to realise that an 'information society' with new services was emerging that was in need of new legislation.[20] At this first stage of development the Internet emerged as an information and selling platform (web 1.0).

17 At the second stage, new business models developed that provided consumers with other kinds of services, yet still related to information, without charging them a price. These services, such as search engines or social platforms that connect people with people (web 2.0), were often exclusively financed by advertising. Whereas, at the first stage, information was largely limited to information as an object of the service; at the second stage, personal data became a most important input for new kinds of business models that were information-related. The advertising value of a service or platform increases with its attractiveness for private users who, in turn, provide its operator with personal data as the key input for such business models.

18 In the Internet of Things, physical objects get connected with each other and with the environment. This brings about another major boost of the data that are collected and an extension of the data that enter into big data collections and business models. At this stage of Internet development, any data that is collected by somebody for a particular purpose can become a most important asset for other economic players or public entities for very different purposes. For instance, smart cars nowadays collect data for steering driverless cars and for providing better and timely—even predictive—maintenance services. But cars may also register the driving habits of the driver, in which the insurance companies are interested, the geographical location of the car at a given moment can inform providers of geographic data, such as Google Maps, about a change of the direction of a one-way road, and inform the public authorities about the volume of use and traffic conditions of roads at a given time. The social benefits of data will even increase with the inclusion of the data in larger datasets that bring together data from different sources, such as from different car manufacturers to get a more comprehensive picture of the concrete traffic conditions in a particular geographic area. The innovative character of this kind of use of data consists in linking large datasets in order to answer many different questions based on mere correlations between different kinds of data (often called 'data mining') in the interest of individual businesses or the public.

19 In this big data world, it also seems that the role of the state is beginning to change. At an earlier stage of the development of the Internet, states started to realise that it is becoming increasingly important to grant private businesses access to publicly held data (so-called 'public sector information', PSI) for commercial re-use in order to promote new commercial information services.[21] Conversely, the modern private data economy is increasingly producing data from which big data analytics in particular can extract new knowledge that can optimise public decision-making—whether it is about increasing traffic security based on data collected by cars, protecting the environment, for instance, by relying on information that is collected by machines used in the agricultural sector, or revolutionising health care around the world by collecting and analysing the clinical, genetic, environmental, and behavioural data from myriad sources.[22] In other words, the public sector is a major contributor, as well as a beneficiary of the data economy and big data analyses.[23]

20 In sum, in the development of the 'data economy' a shift of focus can be observed. Whereas the business models of major Internet platform operators are built on the use of personal data and, accordingly, may give rise to particular concerns about effectively protecting the use of personal data, the data economy will no longer be limited

20 In the EU, see in particular Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce'), [2000] OJ L 178/1; Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, [2001] OJ L 167/10.

21 See Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information, [2003] OJ L 345/90, as revised by Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013, [2013] OJ L 175/1; consolidated version available at: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02003L0098-20130717&from=EN> (accessed 10 September 2016).

22 On the benefits for health care, see in particular the study by OECD (*supra* n 5) at 331-78.

23 The OECD argues that the governments should 'lead by example' in promoting data-driven innovation by granting access to public-sector information. See OECD (*supra* n 5) at 404-48.

to the use of personal data for advertising and marketing purposes. There are two more important innovation-driven features of the data economy that can be witnessed. On the one hand, in the era of the Internet of Things, data collection by sensors will allow consumers to be provided with innovative smart products and services that will increasingly replace traditional products. On the other hand, the data collected in this industry will be of particular utility to private actors in very different business sectors and to public entities. Hence, data collected by smart products will become an important input, both for other businesses and for the government.

II. What do we mean by data and big data?

21 Asking the question of who owns the data suffers from the terminological weakness of what is meant by the term 'data'. There are two aspects to the problem. First, more precision is needed in defining the individual data. The second aspect relates to the aggregation of larger datasets and their protection.

22 The first issue relates to the question of the potential object of protection of data. Take the following example: a smart car of manufacturer A, through the sensors attached to its dampers, locates a pothole. This information is not yet noticed by any natural person; however, it is stored in the form of digital data on a server of manufacturer A. If the law recognised ownership of A in this data, the question arises whether ownership relates to the pure digital dataset in the form of bits and bytes, or to the 'information' the digital dataset contains. This makes a major difference from a competition-oriented perspective. The pothole can of course be registered by the smart cars of different manufacturers (A and B) that follow each other. Hence, the 'information' in which the public road authority is interested could be extracted from two different (competing) datasets.

23 This example shows that the concept of data is in need of additional precision. When we use the term 'digital data', we typically refer to 'machine-readable encoded information'.[24] However, the interest in 'protecting data' relates to the information encoded in these bits and bytes. As regards this information, in turn, a distinction can be made in terms of semiotics between the different levels of information.[25] For data protection, the distinction between the syntactic and the semantic level is key. The syntactic level regards the representation of information in particular signs, for instance as a text, a photograph or a video. In contrast, the semantic level relates to the meaning. Take the example of a camera at a public square that produces a video. The syntactic information is the video as such, which can be stored on different carriers. In contrast, the meaning that can be extracted from that video, for instance, how many people or vehicles cross the square on a single day, is placed on the semantic level. These distinctions can be further illustrated by the example of a novel printed as a book. The book is the physical carrier of the information. The syntactic information consists in the text printed in a sequence of letters and words. The semantic information is the story told by the novel. If somebody does not speak the language in which the novel is written, to this person the information will only be accessible on the syntactic level.

24 Hence, whenever the law protects 'data', it has to make clear what it really protects. There is no general argument against protecting semantic information. Indeed, trade secrets protection and private data protection relates to the semantic level of information.[26] The know-how of a firm consists in technical knowledge; it does not matter whether this knowledge arises from a drawing, a text or a combination of both, or whether this knowledge is stored in a digital format or not. Similarly, individuals are protected against unauthorised processing of information relating to them, whether this information is contained in a text, photographs, or audiovisual recordings. In contrast, in the abovementioned example on the potholes in the street, it would be better to avoid protecting the semantic information the sensors of a car collect. Hence, the question of whether the law should protect the semantic or the syntactic information, or even only the integrity of the digital file, will depend on the circumstances. This analysis would seem to argue for context-specific regulation. Even a general regime on the protection of industrial data would thus appear problematic since, in some instances, protecting semantic information such as in the case of trade secrets seems the right approach, while protection of data collected through sensors in the public sphere should probably not be extended to the meaning these data are able to convey. To

24 Definition used by Herbert Zech, 'Data as tradeable commodity' in Alberto de Franceschi (ed.), *European Contract Law and the Digital Single Market* (Insentia: 2016 forthcoming) 51, at 53.

25 On this distinction see also Maximilian Becker, 'Rechte an Industrial Data und die DSM-Strategie' (2016/1) GRUR Newsletter 7, available at: <https://www.grur.org/fileadmin/daten_bilder/newsletter/2016-01_GRUR_Newsletter.pdf> (accessed 10 September 2016); Andreas Wiebe, 'Protection of industrial data—a new property right for the digital economy' (2016) *Gewerblicher Rechtsschutz und Urheberrecht Internationaler Teil* (*GRUR Int*) 877, at 881; Zech (*supra* n 24) at 53-54.

26 Art 4(1) General Data Protection Regulation (*supra* n 17) defines 'personal data' as 'any data relating to an identified or an identifiable person'.

draw the line between the two approaches is not an easy task. Constitutional rights can argue in favour of protecting semantic information, such as in the case of personal data. Yet in other instances constitutional rights and competition policy will argue against ownership in semantic information, if such protection has the potential of undermining the free flow of information.[27]

25 The second problem arises from the fact that firms do not only hold individual pieces of data. Data are collected and then included in larger datasets. This raises the issue of whether there should be protection of each and every data information or whether there should be protection of the whole dataset in its particular composition.

26 This second issue directs the attention to the features of big data, the technical features of big data analytics and, ultimately, big data business models. At the outset, it should be stressed that big data analyses are only one application where data held by one person is used by another person in the data economy. The purpose of big data analyses is to optimise decision-making. The decision-maker can be any person or entity, usually a firm or a public entity. The following three features are key to the technical understanding of big data: volume, velocity and variety (the so-called '3 Vs').[28] 'Volume' relates to the exploding volume of data that is produced by different sources, including the Internet of Things and social media. Big data is defined by the fact that the volume of data to be analysed transcends the current capacity of storage and processing systems. 'Velocity' relates to the dynamic nature of big data. Indeed, big data constantly changes as new data is produced. To keep up with the speed of this process is key in big data analytics because the users of the results of such analyses will usually have to rely on real-time analyses for decision-making in a constantly changing world. 'Variety' relates to a wide range of different kinds and formats of data. Data may originate from very different sources, such as machine sensors, websites or social platforms; it may be structured or unstructured; and it may consist in texts, pictures, audio or video. While it would be important to combine different kinds of data in big data analyses, the large variety of data constitutes a major technological challenge to big data analytics.[29]

27 These technical features also need to be taken into account when it comes to the policy decision of whether additional data ownership rights should be created. The general claim to be made is that data ownership should not create obstacles to big data analyses, because it is through these analyses that new insights and social benefits will be generated. The issue of volume indicates the difficulty of storing all data that needs to enter into an analysis on one server. This means that big data analyses may have to take place in a decentralised manner. Either the 'code has to be brought to the data' or individual datasets need to be screened first for the critical data, which is then transferred for the analysis.[30] In both cases, it is clear that the big data analyst is in need of access to different data sources and that the different data sources cannot *ex ante* be considered as substitutes for each other. Creating new data rights at the upstream level of holding such datasets could therefore considerably obstruct big data analyses.

28 Velocity may be an even more important feature to be taken into account for the regulation of ownership. Velocity indicates that 'data' should not generally be considered as a 'commodity' that

27 In this context, the *Magill* competition law case of the European Court of Justice (now Court of Justice of the EU, CJEU) should be recalled. Since British and Irish copyright law recognised copyright protection for the mere listings of TV programs, TV stations were able to monopolise the downstream market for printed TV programs and prevent the emergence of comprehensive TV guides combining the programs of different TV stations. The case gave rise to the EU case-law on refusal to license. For more detail see at F.II.1. and F.II.2. below. Copyright protection blocked access to the 'information' contained in the TV listings and, thereby, gave rise to dominance of TV stations in the upstream information market and allowed the TV stations to eliminate competition in the downstream market. See Judgment in *RTE and ITV v Commission* ('*Magill*'), C-241/91 P and C-242/91 P, ECLI:EU:C:1995:98, [1995] ECR I-743. For this case, it can be argued that copyright went too far in the first place by blocking access to information. On this case see also at F.II. below.

28 See, for instance, Amir Gandomi and Murtaza Haider, 'Beyond the hype: Big data concepts, methods and analytics' (2015) 35 *Int'l J. Inf. Manag.* 137, 138; Stephen Kaisler, Frank Armour, J. Alberto Espinosa and William Money, 'Big Data: Issues and Challenges Moving Forward', (2013) *46th Hawaii International Conference on System Sciences* 995, available at: <http://www.computer.org/csdl/proceedings/hicss/2013/4892/00/4892a995.pdf> (accessed 10 September 2016); Daniel O'Leary, 'Artificial Intelligence and Big Data' (2013) *IEEE Intelligence Systems* 96, available at: <http://people.westminstercollege.edu/faculty/ggagne/fall2014/301/chapters/chapter1/mex2013020096.pdf> (accessed 10 September 2016); Paul Zikopoulos and Chris Eaton, Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming (2011). The first author to have hinted at these three features seems to be Doug Laney, '3-D Data Management: Controlling data volume velocity and variety' (2001), available at: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (accessed 10 September 2016). From a competition law perspective see Daniel L Rubinfeld and Michal S Gal, 'Access Barriers to Big Data' (16 August 2016) at 8-9, available at: <http://ssrn.com/abstract=2830586> (accessed 10 September 2016).

29 On the technique and process of data analytics see Gandomi and Haider (*supra* n 28) at 140-143.

30 These two solutions are identified by Kaisler et al. (*supra* n 28) at 997.

can be traded like other commodities. Rather, the modern data economy typically has to rely on real-time information. Hence, a concept of ownership in data, similar to copyright in a work, which would invariably be protected for a fixed period of time, would not serve the needs of such data services and big data analytics. Big data analyses that are confronted with dynamic processes and have to serve a purpose in a dynamic environment, such as steering the traffic management system of a smart city, will have to rely on permanent access to real-time data sources. Ownership in individual data will hardly be able to constitute the backbone of such a service.

29 Velocity is closely linked to another 'V' that is increasingly mentioned as an additional feature of big data and which is key from a legal perspective, namely, 'veracity'.[31] Data needs to be reliable to serve the purposes of a data economy. Where real-time data are needed, but not delivered, the service also misses the requirement of veracity. From a legal perspective, veracity indicates that the supply of data should also come with particular responsibility.

30 In this regard, it is worth noting that the EU is currently moving in the direction of fixing uniform standards of 'quality' of 'digital content' that need to be respected if digital content is supplied under a contract with a consumer.[32] The Proposal for a Directive on the supply of digital content defines 'digital content' as '*data* which is produced and supplied in digital form, for example video, audio, applications, digital games and any other software'.[33] The Directive would have the effect of creating a harmonised regime of contractual liability for both physical goods, which are also often sold over the Internet, and data. This, however, does not automatically lead to the recognition of ownership in the underlying data.[34] Whether there is contractual liability if digital content does not meet the quality that is to be expected under the contract and whether the supplier transfers ownership in the framework of such a contract are two separate legal issues. Most importantly, ownership implies a third-party effect while the proposed Directive only creates rights and obligations between the parties to the sales contract.[35]

31 In addition, also as regards big data analyses, the difference between the syntactic and semantic level of data is to be taken into account. Big data analytics consists in reading large datasets to discover 'new' meaning—in the sense of (semantic) information—that has so far not been observed. Big data analytics acts like a person who is able to read the data in a different way by identifying correlations between different data—again in the sense of information—to draw conclusions from those correlations. Hence, the information that big data analyses produce is already hidden in the pre-existing datasets. However, it is big data analytics that allows us to discover this semantic information. This explains how problematic it would be to recognise protection of all semantic information contained in the pre-existing datasets for those who control access to these sets. It is indeed the contribution of the data analyst that leads to the discovery of that information and, hence, any right in this information should be vested in the data analyst[36] rather than the holder of the datasets that are analysed.

III. From value chains to value networks

32 For considering whether new property rights in data are to be recognised from a functional perspective, it is crucial to understand who generates economic value and, as a follow-on question, whether this contribution depends on the recognition of a property right. In this regard, it is important to understand that in the data economy, value is generated differently than in the traditional economy.

33 In the traditional economy, the still dominant paradigm relates to vertical value chains. Manufacturers purchase input for the production of goods in upstream markets and then sell them through distribution chains—often including wholesalers and distributors—to consumers. At each level of the production and distribution chain, some economic value is added.

31 An example is big data analytics in the healthcare sector; see Wullianallur Raghupathi and Viju Raghupathi, 'Big data analytics in healthcare: Promise and potential' (2014) 2(3) *Health Information Science & Systems* 1, at 2, available at: <https://hissjournal.biomedcentral.com/track/pdf/10.1186/2047-2501-2-3?site=hissjournal.biomedcentral.com> (accessed 10 September 2016).

32 Article 6 of the Proposal of Commission of 9 December 2015 for a Directive of the European Parliament and of the Council on certain aspects concerning the supply of digital content, COM(2015) 634 final.

33 *Ibid*, Art 2(1)(a).

34 See, however, De Franceschi and Lehmann (*supra* n 9) at 59-60 and 71 (relying on the corresponding rule contained in the previous draft for a Common European Sales Law and attributing a property dimension to this proposal).

35 As regards the recognition of ownership in the download of a computer program by the CJEU in the Judgment in *UsedSoft*, C-128/11, ECLI:EU:C:2012:407, paras 45-52, see at C.V. below. See also De Franceschi and Lehmann (*supra* n 9) at 60-63 (relying on this decision in their yet cautious support of data ownership).

36 Such information can constitute trade secrets. On trade secrets protection see at C.II. below.

34 In contrast, in a world of smart goods and the Internet of Things, economic value is increased in very complex and dynamic value networks, which can be disruptive for traditional value chains,[37] through collaboration of the different participants in the network. This paradigm shift from value chains to dynamic value networks is identified as a core feature of the current digital transformation of the industry.

35 Four sub-factors are relevant for this shift:[38] (1) *Improving decisions based on data:* sensor-generated industrial data and analysis of big data help firms optimise their decisions. For instance, predictive maintenance becomes possible. (2) *Full automation:* Automation through digital technology, including robotics, revolutionises production and the use of products (e.g. driverless cars). Automation increases the speed of production and decreases the likelihood of defects. (3) *Connectivity:* Objects and machines within the factory and beyond get connected over the Internet and allow supply and production to be steered from the perspective of the need of the customer, which results in quicker production and distribution while saving resources. (4) *Increasing role of Internet intermediaries:* The intermediaries from the Internet sector who have the best access to and knowledge of the needs of consumers and of controlling the data interfaces between different markets gain a competitive advantage in the industrial sector where smart products are produced. This explains why Google and other firms are today trying to expand their activities into the industrial sector. Google, or Alphabet as Google's parent company, may now already have considerable competitive power for entering the market for smart, driverless cars based on its control of geographic data, and may provide most efficient transport services to passengers who, in the future, will no longer buy their own cars but become passengers of Google transport services. At the same time, by expanding their activities to the production and operation of smart products, these Internet intermediaries will gain control over new sources of data.

36 Hence, whereas the digital transformation of the industry decreases existing entry barriers and may even force industrial incumbents out of the market, control over data enables firms originating in the Internet sector, such as Google, to enter into and gain considerable market power in a large variety of different markets for the production and operation of smart products. Recognition of data ownership may therefore have the unwanted effect of strengthening the market power of these firms even more, while, from a competition perspective, it would be wiser to promote access to data that is needed by other market players to operate in such markets.

IV. The interests of different stakeholders

37 The preceding analysis already provides some important insights into the interests of different stakeholders. This analysis underlines the observation in the introduction (at A. above) that industrial players who have already started to invest in the Internet of Things are reluctant to advocate data ownership.

38 The major technological challenges of the Internet of Things relate to big data analytics. This is the area where most investment is needed for tackling the technological obstacles to handling rapidly growing dynamic datasets and solving the problem of analysing a large variety of different kinds of data. However, such innovation is more likely to be fostered through copyright protection for the software solutions employed in the framework of big data analyses rather than through ownership in the data analysed.[39]

39 Moreover, it is to be acknowledged that the non-economic interests of natural persons in the use of their personal data deserve to be safeguarded, also in the data economy. While personal data protection needs to be taken into account, it does not argue as such against the recognition of an economic ownership right of a firm that collects data about the use of a smart product by a natural person. Both rights can coexist. This has the important consequence that rules on the protection of personal data can prevent a data owner from commercialising that data. The industrial holder of personal data can also respect data protection rules by making the data collected from individual natural persons available to third persons in an aggregated and anonymised form in larger datasets. To the extent that big data analytics manages to reproduce personal data, data

37 This has recently been highlighted by a study conducted by Roland Berger Strategy Consultants on behalf of Bundesverband der Deutschen Industrie (BDI). See Roland Berger Strategy Consultants and BDI, 'Analysen zur Studie "Die digitale Transformation der Industrie"' (February 2015) 4-8, available at: <http://bdi.eu/media/user_upload/Digitale_Transformation.pdf> (accessed 10 September 2016).

38 *Ibid*, at 8.

39 Another kind of protection would consist in patent protection for algorithms. However, this is rejected by Josef Drexl, Reto M. Hilty, Luc Desaunettes, Franziska Greiner, Daria Kim, Heiko Richter and Gintarė Surblytė, 'Data Ownership and Access to Data—Position Statement of the Max Planck Institute for Innovation and Competition of 16 August 2016 on the Current European Debate', paras 12-17, available at: <http://ssrn.com/abstract=2833165> (accessed 12 September 2016).

protection rules may apply again as regards the re-use of that data.

40 As regards personal data, it is important to note that the fact that a natural person is and will often be the source of specific data does not automatically argue in favour of allocating data ownership as an economic right to commercially exploit that data to this person. Protection of personal data is neither vested in the natural person for economic purposes, nor is it an absolute right.[40] Personal data protection does not allocate economic value.[41] Hence, there is room to grant economic rights of exploitation of data originating from natural persons to other persons or firms.

41 The same applies as regards the property of the purchaser of a smart product. The property in the car as a physical object does not automatically extend to the commercial exploitation of the data that are produced by the sensors of that car. The question of whether data ownership should be recognised, and for whom and with which scope of protection, should only be decided against the backdrop of economic welfare considerations.

C. Existing protection regimes as a basis for 'data ownership'

42 Already at the end of the preceding part, it was clarified that at least two rights that are recognised by law do not provide a sufficient basis for data ownership; namely, personal data protection and real property in a smart product that produces the relevant data. However, there are other legal regimes that could provide protection in favour of the firm that controls data. Most obvious candidates are database rights and trade secrets protection. Beyond this, in certain circumstances, the question may arise whether patent protection extends to data that is generated through a patented process. Moreover, one could also contemplate unfair competition rules and the like, as well as a generalisation of property in tangibles as a civil law concept. In sum, none of these regimes provides a convincing or comprehensive basis for data ownership. In contrast, it will be shown that factual control over data can enable the data holder to commercialise that data without additional legal protection by relying on contract law.

40 See Recital 4 of the General Data Protection Regulation (*supra* n 17). See also Pamela Samuelson, 'Privacy as Intellectual Property', (2000) 52 *Stanford L Rev* 1125.

41 Zech (*supra* n 24) at 60.

I. Database protection

43 At first glance, database rights present a most obvious property regime for controlling access to data.[42] However, this kind of protection has particular limitations that explain why it will often fail to provide protection to data for the new business models of the data industry.[43]

44 The EU legal regime for database protection provides for a two-tier system: Copyright protection is granted to creative databases;[44] *sui generis* protection is granted to databases based on 'substantial investment'.[45]

45 The availability of copyright protection can be excluded from the outset. Article 3(1) of the Database Directive clarifies that the character of a creative work defined as 'the author's own intellectual creation' has to relate either to the selection or to the arrangement of the database's contents. According to the CJEU this originality requirement is satisfied if 'through the selection or arrangement of the data which it contains, its author expresses his creative ability in an original manner by making free and creative choices ... and thus stamps his "personal touch"'.[46] Already this definition explains that the individual data as such will not be copyright protected. This is also explicitly confirmed by Article 3(2) of the Directive, which states that copyright protection for databases will not extend to the contents as such. Hence, even if data were included in a copyrightable database, such copyright protection would not extend to that data.

46 *Sui generis* database protection may at first glance provide a better basis for protecting data generated in a world of the Internet of Things.[47] However, this form of protection also has its limitations. They arise from both the subject-matter of protection and the scope of protection.

47 As regards the subject-matter of protection, a 'database' is uniformly defined as a 'collection of

42 Arts 7-10 Database Directive (*supra* n 19).

43 See also Wiebe (*supra* n 25).

44 Art 3 Database Directive.

45 Art 7(1) Database Directive. Note that both forms of protection may also coincide. A given database may be both creative and based on substantial investment.

46 Judgment in *Football Dataco v Yahoo! UK*, C-604/10, ECLI:EU:C:2012:115, para 38 (adopting the general originality concept of EU copyright law as developed by the Court for other categories of works to databases).

47 It is even argued that the sui generis database right will often protect big data databases; see Giulio Corragio, 'Big data and IoT–a great match with troubles...' (19 June 2015), available at: <http://www.medialaws.eu/big-data-and-iot-a-great-match-with-troubles/> (accessed 10 September 2016).

independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means'.[48] Protection will also be granted if the arrangement and storage is accomplished by 'electronic, electromagnetic or electro-optical processes'.[49] Hence, collections of digital data can usually be considered as databases in the sense of the Directive.[50] However, a *sui generis* database right only subsists if 'there has been qualitatively and/or quantitatively a *substantial investment* in either the *obtaining, verification or presentation of the contents*'.[51] The CJEU has interpreted these requirements in a very restrictive way. It clarified that the investment has 'to refer to the resources used to seek out existing independent materials and collect them in the database, and not to the resources used for the creation as such of independent material.'[52] The CJEU explained this with the objective of the Directive to create incentives for the making of databases and not for the creation of the data that goes into the database.[53] Hence, a distinction is to be made between the 'creation' of the materials contained in the database and the 'obtaining' of these materials.[54] This leads to the conclusion that the creation of smart products with sensors that collect data should not be considered for the assessment of whether the investment in the database was 'substantial'.[55] The same applies to big data analyses. These may well require substantial investment. However, such analyses only lead to the creation of new data in the form of knowledge, which may then be included in databases. For the protection of these databases, the investment in the big data analyses is not to be taken into account.

48 As regards the scope of protection, it is important to note that the *sui generis* database right only protects the database as a collection of data and not the individual data. The Directive thereby aims to keep the (semantic) information that can be derived from the data in the public domain.[56] Extraction and re-utilisation of individual data only fall within the scope of protection of the database if these data form a 'substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database'.[57] The concepts of 'extraction' and 're-utilisation' further restrict the scope of protection. In particular, big data analyses, whereby the 'code comes to the data' in order to generate new information, will not lead to any 'extraction' since there will be no 'permanent or temporary transfer of all or a substantial part of the contents of a database to another medium'.[58]

49 In sum, it is quite obvious that the Database Directive is based on a database technology that no longer corresponds to the use of data in an era of 'Industry 4.0' or the Internet of Things. In particular, by protecting a collection of materials for a given period of time (15 years as of the completion of the database),[59] the concept of a database is much too static to adequately respond to the features of constantly changing datasets and real-time data services.

50 This latter point may raise the question of whether the Database Directive is in need of a reform. However, the fact that the Directive does not respond to the needs of the modern data industry in a technologically appropriate manner cannot by itself justify reforming the Directive by introducing a right of data ownership. Rather, such reform is in need of an economic justification, which is part of the analysis further below (section D. below).

II. Trade secrets protection

51 Trade secrets protection is another protection regime that inevitably comes to mind as regards the protection of data.

52 The EU has recently adopted a directive for harmonising the national rules on trade secrets protection.[60] As regards the modern data industry, this Directive may already be considered as technologically out-dated, since at the time of the preparation of the Commission Proposal, the implications of the new data economy were not yet fully perceived or understood.[61] As a consequence,

48 Art 1(2) Database Directive (*supra* n 19).

49 Database Directive, Recital 13.

50 Zech (*supra* n 24) at 70.

51 Art 7(1) Database Directive (*supra* n 19) (emphasis added). This means at the outset that there may be databases fulfilling the definition of a 'database' in the sense of the Directive that, however, are not protected since they meet the requirements neither for copyright-protected databases, nor for *sui generis* databases. Confirmed by the CJEU in its Judgment in *Ryanair v PR Aviation*, C-30/14, ECLI:EU:C:2015:10, paras 35-40.

52 Judgement in *British Horseracing Board*, C-203/02, ECLI:EU:C:2004:128, [2004] ECR I-2195, para 31.

53 *Ibid*.

54 *Ibid*, para 32.

55 See also Żdanowiecki (*supra* n 10) at 21.

56 See Zech (*supra* n 24) at 71.

57 Art 7(1) Database Directive (*supra* n 19).

58 Article 7(2)(a) Database Directive.

59 Article 10(3) Database Directive only takes changes to contents of databases into account to the extent that such changes amount to a new substantial investment, which leads to a revival of protection for 15 years.

60 Directive (EU) 2016/943 of the European Parliament and the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use or disclosure, [2016] OJ L 157/1.

61 See Proposal of the Commission of 28 November 2013 for a Directive of the European Parliament and of the

the text of the Directive is rather unclear as to what extent, for instance, data produced by smart products benefit from trade secrets protection.

53 In comparison to database protection, trade secrets protection has the obvious advantage of protecting the specific information. However, there are other shortcomings:

54 Most importantly, trade secrets protection relies on rather narrowly defined requirements for the subject-matter of protection. According to Article 2(1) of the Directive, the know-how or business information (1) needs to be 'secret' in the sense that it is not 'generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question'; (2) the information must have 'commercial value' because of its secrecy; and (3) it has to be subject to 'reasonable steps ..., by the person lawfully in control of the information, to keep it secret'. None of these three requirements can be easily applied in the context of data produced by sensors attached to smart products. First, while the secrecy could be confirmed for data that is produced by the machines inside a factory, data collected by smart cars on freely accessible roads could be collected by the cars of many manufacturers and, hence, will not fulfil this requirement.[62] Second, while data may nowadays have great commercial value, it is quite questionable whether it will always be possible to establish a causal link between the secrecy of the information and its commercial value. In the context of big data analyses, an individual piece of information may appear quite trivial, but particular value may arise from correlations with other data.[63] Third, it is very unclear which steps will be required of the person in control to keep the information secret.[64] Fourth, where data is generated in a network of different entities connected through a value network, it will be particularly difficult to allocate protection to a single person controlling the secret.[65]

55 Yet another question is whether the subject-matter of protection needs to be interpreted narrowly in the light of the objectives of the Directive. The Directive pursues the goal of promoting the competitiveness and innovative strength of businesses through protecting secret information.[66] However, data are nowadays largely produced as a by-product of smart machines and goods, whereas these data can be commercialised in completely different markets and for completely different purposes (not least in the public interest). Here, data is largely used by the data holder as an asset for generating additional income. In addition, protection of the data as trade secrets will not always promote innovation through the holder of that data. Rather, the challenge will often consist in promoting access to that data for other firms and public entities that may generate additional knowledge from that data through big data analyses. This argues for making a distinction between information that serves the core business of the holder of data, such as personal data held by Internet platform operators, as the backbone of the underlying business model, as well as data generated through machine sensors that are designed to be immediately used for the production process on the one hand, and other data, which are rather a by-product of the firm's core business, on the other hand.

56 Finally, it should be noted that trade secrets protection is much narrower in scope than an exclusive data use right. It does not protect against any use of the data, but requires 'unlawful' conduct which, to summarise the different provisions of the Directive, can be regarded as contrary to honest commercial practices.[67] Hence, the Trade Secrets Directive only establishes a system of liability for specific tortious conduct and not a property rights system.[68] However, such further limited protection can be considered as better suited to serve the purposes of the data economy, by focussing on the particular way in which a third party has specifically acquired access to the data instead of granting exclusive protection against the use of data. Such exclusive property protection would easily conflict with the fundamental right of freedom of information.

III. Patent law

57 In limited sets of cases one could even consider protection based on patent law. The reason for this is that the scope of process patents also extends to 'products' that are obtained through that process. For instance, in the European Union, Article 25(c) of the—yet not effective—Agreement on the Unified

Council on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, COM(2013) 813 final. See also Wiebe (*supra* n 25) at 880 (pointing out that the drafters of the Directive did not have big data in mind).

62 In this context, it is important to note that independent discovery of the same information will not lead to unlawful acquisition of the information. See Article 1(3)(a) Trade Secrets Directive.

63 See Zech (*supra* n 24) at 63 (therefore criticising Recital 8 Trade Secrets Directive according to which trivial information should not be protected).

64 On the difficulties to keep information secret in a network environment, see Wiebe (*supra* n 25) at 880.

65 See also Wiebe (*supra* n 25) at 880.

66 See Trade Secrets Directive, Recital 1.

67 See, in particular, Art 4(2)(b) Trade Secrets Directive.

68 See also Drexl et al (*supra* n 39) at paras 18-20.

Patent Court stipulates that a process patent also provides the right to prevent a third party from 'offering, placing on the market, using, or importing or storing for those purposes a product obtained directly by a process which is the subject-matter of the patent'.

58 The question in this regard is whether 'data' can be a 'product' that is obtained by using a process patent.[69] This question would become relevant for instance where data is produced in a factory in applying a patented production method or, maybe more relevant, in the context of a process patent applied in medical diagnostics. In the latter case, the patent owner would also 'own' the 'result' of the diagnosis.

59 However, at the outset, such protection would only become relevant where the patent is used without the consent of the right-holder. Only if the patented process is used without a licence does the patent holder have a right to prohibit the commercialisation of the product as the offspring of the process.

60 The reason why the legislature extends the protection of process patents to the commercialisation of products is that process patents are much weaker than product patents. The owner of a product patent enjoys full protection against price competition from imitators in the product market. In contrast, the holder of a process patent runs the risk of having to compete with firms that offer essentially the same product manufactured with an alternative process. Extending protection to the products that are produced with the process assimilates process patents to product patents regarding the economic incentives arising from the patents. It also addresses the problem that third parties could otherwise legally serve the market with products produced abroad by applying the process if the process patent is only protected in the importing country.

61 However, already as a matter of principle, it does not seem appropriate to extend patent protection to information as the product of a process patent. Moreover, German courts seem to deny protection for information that is derived from a process patent. An interesting decision in this regard is the one by the District Court of Düsseldorf in the *Hunde-Gentest* case.[70] In this case, the process patent for a genetic test for dogs was protected in Germany, but not in Slovakia. The defendant, who previously applied the test in Germany, moved the testing to Slovakia to avoid a patent infringement. Therefore, the Court was only requested to decide whether the plaintiff can rely on a process patent to prevent the defendant from communicating the test results to Germany. The Court denied such protection, arguing that the test results as pure information cannot be considered the product of the process. The Court noted that, since information is directly accessible for humans without any further technical process, information as such lacks technicity and therefore cannot be patented. Yet the Court refrained from arguing that the 'product' of a process needs to be patentable by itself in order to be protected within the scope of the process patent.[71] Rather, the Court showed great sensitivity for the free flow of information. It rejected protection so as to avoid using patent law as a kind of trade secrets protection. In particular, the Court stressed that patent law should not support a claim to ban communication of the test result to anybody in Germany, which, in the last resort, would even include denying a person who knows about the test result entrance to German territory.

IV. Unfair competition law and similar protection regimes

62 In many jurisdictions, unfair competition laws and similar protection regimes, such as the tort of misappropriation in common law countries, may provide subsidiary means of protection against free-riding where other protection mechanisms are not available.

63 However, whether such a role should be attributed to these general principles or laws as regards the holding of data, is again a policy issue which should only be answered in the affirmative if there is sufficient economic justification for protection against free-riding (see section D. below). Free-riding as such should not be considered a violation of the law unless it undermines incentives for investment in the production of the asset that is copied.

69 On the similar issue whether patent protection for a computer-based process for producing aesthetic creations extend to these creations see Jean-Marc Delthorn, 'Councours de droits sur les œuvres numériques—Le cas des creations issues de procédés brevetés', (2016) 60 *Propriétés intellectuelles* 285.

70 *Landgericht Düsseldorf* of 16 February 2010, Case 4b 0 247/09—Hunde-Gentest, available at: <https://www3.hhu.de/duesseldorfer-archiv/?p=813> (accessed 10 September 2016). See also Oberlandesgericht München (Higher District Court Munich) of 22 October 2015, Case 6 U 4891/14, (2015) *Beck-RS* 18783.

71 This is also the view of the EPO. See EPO, Decision of the Enlarged Board of Appeals, G 1/98, Transgenic plant/NOVARTIS-II, [2000] OCJ EPO 111, at 138. The Enlarged Board of Appeals confirmed the availability of process patents, including protection of the products deriving from the process according to Art 64(2) EPC, even in a case where the product would be a plant, which is excluded from patentability under Art 53(b) EPC.

V. 'Digitisation' of the civil law concept of property?

64 Civil law countries are not unlikely to discuss nowadays whether the concept of property found in the national Civil Code, which is usually limited to the ownership of tangible items and land, should be opened, namely, in a move to 'digitise the Civil Code', to also include data. For instance, in 2016, the *Deutsche Juristentag*,[72] which is the most important private discussion forum for legal reform in Germany, bringing together law professionals from all different sectors, considered whether German civil law is in need of a 'digital up-date'.[73]

65 Yet, to equate data with tangible objects as a subject-matter of property is a rather risky undertaking. The risk is that, as an expression of general enthusiasm and striving for modernisation, the legislature or courts will not give sufficient consideration to the different economic characteristics that distinguish markets for non-tangible objects from those for tangible objects.

66 Hence, the question of whether civil law is in need of being 'updated', should be considered carefully and within the specific context of protection. To transfer the principles of contractual liability developed for the sale of tangible goods to defects of digital goods, is one thing;[74] to recognise a property right for holders of data with exclusionary effects on third parties is another thing. In Germany, the debate is mostly triggered by certain limitations of tort law. Under Section 823(1) German Civil Code, there is only a claim for damages if somebody injures the 'life, body, health, freedom, property or another right' of someone else.[75] Courts have continuously extended the range of 'other rights', to include, for instance, the general personality right, but they have also limited those rights to 'absolute rights'. This is why it is now also discussed whether courts should recognise 'data ownership' as another absolute right to protect the integrity of datasets against injuries committed by third parties. For instance, the need for such protection is quite evident when computer viruses delete large and valuable datasets, while the physical carrier and its functions remain intact. The downside of this is that recognition of such a right in the framework of Section 823(1) Civil Code would also provide for injunctive relief to prevent injury. For that purpose, German courts rely on an analogy to Section 1004 Civil Code, the basis for injunctive relief in case of unlawful interference with property.

67 Injunctive relief raises the more important question regarding the extent to which the scope of protection of such data ownership is to be assimilated to property in tangible objects. Property in tangibles basically provides two sub-rights, a right of integrity and a right to exclude others from any use.[76] The debate on data ownership is inspired by the lack of protection as regards the integrity of data, whereas the recognition of a right to exclude other persons from *any* use of the data would amount to a very powerful intellectual property right that would have the potential of undermining the free flow of information.[77] Also, from an economic standpoint, a right to exclude others from the use of data is less needed than in the case of tangibles. Data are not rivalrous; hence, someone else's use of the same data does not prevent the 'owner' from using these data. Accordingly, from an economic perspective, it is easier to justify protection of the integrity of data than to provide full protection, including injunctive relief, as regards the use of data.

68 This debate on extending the property concept to digital data was more recently also inspired by the *UsedSoft* decision of the CJEU.[78] In this case, the Court explicitly recognised 'ownership' of the person legally downloading a computer program from the Internet. However, this holding was limited to the application of the exhaustion rule in the Computer Programs Directive.[79] Exhaustion of the distribution right under copyright law requires a first 'sale' of a copy of the work through the right-holder or with her consent. The CJEU defined a 'sale' as 'an agreement by which a person, in return for payment, transfers to another person his rights of *ownership*

72 The Deutsche Juristentag convened in Essen on 13-16 September 2016.

73 The debates at the Deutsche Juristentag revolve around *Gutachten* (expert reports), which are usually prepared by law professors. The 'digital update' of the German Civil Code is assessed in the *Gutachten* by Florian Faust, *Digitale Wirtschaft—Analoges Recht—Braucht das BGB ein Update? Gutachten A zum 71. Deutschen Juristentag* (Munich: C.H. Beck, 2016), also available at: <http://static1.1.sqspcdn.com/static/f/1376130/26847040/1455040340113/Faust+Digitale+Wirtschaft+Analoges+Recht+Gutachten+fur+den+71.+DJT.PDF?token=73St8IVwwV4tYnJQSVMQJmH3F8c%3D> (accessed 10 September 2016).

74 See the Commission Proposal for a Directive (*supra* n 32).

75 English translation of the *Bürgerliches Gesetzbuch* available at: <http://www.gesetze-im-internet.de/englisch_bgb/> (accessed 10 September 2016).

76 As regards the right to exclude under German law, see Sec 903 Civil Code. On the distinction between the three different rights of property regarding data ownership, including (1) possessing data—with the possibility to exclude access—, (2) using data, and (3) destroying data (right of integrity), see Zech (*supra* n 24) 56-57.

77 See also Wiebe (*supra* n 25) at 882. (considering whether recognition of data ownership would lead to a paradigm shift in protecting information).

78 Judgment in *UsedSoft* (supra n 35).

79 See Art 4(2) Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs, [2009] OJ L 111/16.

in an item of tangible or intangible property'.[80] By recognising ownership in the digital copy of the program, which is provided to a customer on a permanent basis,[81] the Court managed to transfer the concept of exhaustion to the digital field. Hence, in *UsedSoft*, the CJEU did not recognise any general concept of data ownership.[82] Rather, the Court only relied on ownership in a digital download to limit the exclusivity of the copyright as another property right. We can learn from this judgment that limited recognition of property rights can also have a liberalising effect and thereby promote the free movement of data in the digital economy. However, such recognition should not be generalised by arguing in favour of allocating ownership involving third-party effects wherever persons are legally in permanent control over the use of any data. This may well have the opposite effect of hampering the free flow of data and information in the data economy.

VI. Factual exclusivity and contract law

69 Despite the uncertainties and shortcomings of the different protection regimes, the players of the data economy do not seem to suffer from the lack of recognition of general data ownership. The reason is that markets can also develop with relatively little legal exclusivity where access can effectively be controlled by technical means.[83] Factual exclusivity has the potential of forcing parties into negotiations and can trigger transactions in very similar ways as in the case of intellectual property.

70 Such data contracts based on the factual holding of data are therefore meant to grant access to these data.[84] However, this does not exclude agreement on certain limitations of the use of data. Accordingly, contract law may exercise even stronger restrictions on the use of data than a new ownership agreement that could provide for mandatory exceptions and limitations.[85]

71 A very prominent example of an area where markets for immaterial exploitation emerge with very little legal exclusivity is the marketing of sports rights. There are only few jurisdictions which provide special intellectual property rights for the audiovisual exploitation of sporting events.[86] Other jurisdictions manage to provide the same conditions for markets for sports rights with comparable value streams without such legislation. The reason for this is that the organisers of such events can control access to the premises of the sporting events and thereby charge a price from the broadcaster that is allowed to produce the broadcast.[87] Of course, there is a risk that third parties will use the broadcasts without authorisation. However, it suffices in this regard that the broadcasting corporation that was granted access to the event is protected by its investment by copyright, or at least its original related right, in the broadcast.

72 As regards the data economy, this example of the sports rights may explain that, even where misappropriation by third parties is a concern, there is no need to recognise ownership of the data holder as long as the investor in access to the data—such as the big data analyst—disposes of an intellectual property right that prevents third-party use, such as the copyright in the software tools for analysing big data. The data holder itself will regularly be able to exclude others from access through technical means, including technical protection measures. Rules of criminal law that make unauthorised access to data a crime, such as data or computer espionage, can further strengthen factual exclusivity without recognition of ownership in the sense of private law.

D. Potential justifications for recognising data ownership

73 Against the backdrop of the uncertainties and shortcomings of existing protection regimes, we now turn to the question of whether there is an economic justification for the recognition of data ownership. In this regard, the analysis can rely on insights from intellectual property scholarship.

80 *UsedSoft* (*supra* n 35) at para 42.

81 *Ibid*, at para 45.

82 This is also confirmed by authors who rely on this judgment to argue in favour of a concept of general data ownership. See De Franceschi and Lehmann (*supra* n 9) at 60-63.

83 See also Żdanowiecki (*supra* n 10) at 25.

84 See Zech (*supra* n 24) at 59.

85 On the question whether promoting access may hence justify introduction of a data ownership see at D.V. below.

86 The most prominent example is French law. Arts L333-1 through L333-5 Sports Code (*Code du sport*) vest the sports associations with an exclusive right of audiovisual exploitation. Original French text of the *Code du sport* available at: <https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000006071318> (accessed 10 September 2016).

87 See also Thomas Margoni, 'The Protection of Sports Events in the EU: Property, Intellectual Property, Unfair Competition and Special Forms of Protection' (2016) 47 *IIC* 386 (arguing that, in principle, the combination of the exclusivity of the sports venue and contract law is capable of making markets for sports rights work).

I. Incentives for generating and collecting data

74 The standard argument in favour of recognising intellectual property rights is based on a utilitarian incentive theory. Intellectual property is designed to promote innovation. However, the subject-matter of protection of these rights, such as inventions and works of creativity, is characterised by the features of public goods. Without the recognition of legal exclusivity, everybody else would be able to free-ride by copying and, consequently, nobody would be willing to invest in the production of such public goods.[88]

75 As demonstrated further above, the generation and collection of data allows for very new and innovative business models that lead to large gains in allocative efficiency in manufacturing and maintenance, as well as far-reaching social benefits based on big data analyses. Hence, there is a case for also fostering incentives for generating and collecting the underlying data. However, it is less clear whether, for that purpose, data ownership is required. In this regard, the incentives of different players need to be analysed.

76 There is always some human act that can be found at the very beginning of the generation of data and the commercial exploitation of these data. A manufacturer may decide to employ machines and robots that are equipped with sensors to control and steer the production process. The owner of a smart car decides where to go with this car and where the car will register data about the density of traffic or the physical conditions of the road. A patient provides the blood for a blood test, the result of which may go into datasets that are subsequently analysed. In all of these cases, the relevant person would and should certainly know about the generation of the digital data, and may even have to give her consent based on the rules on data protection. However, additional ownership in the data is not necessary as an incentive to generate such data. Hence, in principle, it is possible to conclude that there is no need to vest the person at the beginning of the value chain with exclusive rights to exploit that data as a means to create incentives for the generation of that data.

77 The same holds true for the next step of exploitation. The data produced by a smart car will be transferred to the manufacturer of that car. The car manufacturer will be sufficiently motivated to generate data that will guarantee smooth operation and maintenance of the car. Generation of that data is very much part of the firm's business model. Furthermore, the potential of follow-on markets creates sufficient incentives for collecting the data, whether database rights are available or not, even in cases where the main business model does not require the data to be stored on a permanent basis.

78 Nor are additional incentives needed as regards the business model of Internet platform operators (e.g., search engines, social media etc.), for which the collection of personal data is the very core of the success of the underlying business model. Yet the fact that firms nowadays know that, in an emerging data economy, any data may become interesting and that they may be able to commercialise that data based on factual exclusivity, it cannot be argued that there is suboptimal generation and collection of digital data. In general, data are not a scarce resource.[89] The sheer volume and variety of big data constitute the basis but also the particular challenge of big data analytics.

79 Hence, there is not sufficient evidence of the need of data ownership as justified by the incentive theory concerning the generation and collection of data. However, there could be a need for more incentives to invest in tools for technologically challenging big data analyses. Within the value stream of exploiting data, data analyses generate major social value by producing new knowledge and thereby optimise decision-making in many fields. However, although the evolving business models of big data analyses may still be in need of further research, it seems that data ownership will not be the appropriate mechanism for protecting the interest of big data analysts. Access to data held by others should be more of a concern to big data analysts than acquiring ownership in data. It is more important for big data analysts that the data they have access to respond to the challenges of velocity and veracity than having claims against third parties for unauthorised use of the data they produce. Since, in many instances, real-time data is key, data analysts do not have to be so much afraid of competitors' free-riding. What counts more is getting access to the various datasets from which they can gather new knowledge. As regards the other side of the market, namely, the firms and public entities to which big data analysts provide new knowledge for optimising decision-making, data ownership will not be needed either. Such relationships will often be based on contracts for services through which customers are supplied with accurate knowledge at a given point in time. From a competition perspective, the core question is whether data analysts need to rely on data ownership in competition with other data analysts.

88 On the public goods theory for intellectual property, see, in general, William M. Landes and Richard A Posner, *The Economic Structure of Intellectual Property Law* (Cambridge, MA and London, UK: The Belknap Press of Harvard University Press, 2003) 12-16.

89 See also Becker (*supra* n 25) at 7.

This question has to be answered in the negative. Data analysts will not gain a competitive edge by 'owning data' at the expense of their competitors. Rather, they will prevail in competition if they manage to have better access to the various sources of big data, for which they will not rely on ownership but contractual business relationships with the holders of such datasets, on the one hand, and the effectiveness and accuracy of their big data analyses, on the other hand. As regards the latter, it is more important that big data analysts control the technology for big data analysis. For this, they will rely on copyright protection in the software infrastructure and possibly technical know-how rather than data ownership.[90] The same holds true for firms that deliver—typically software-based—tools for big data analysis of other firms.

80 At the last stage, the customers to whom information is delivered based on big data analyses are not in need of data ownership either. To the extent that these data are kept secret and the data analysts are under a contractual obligation to keep that information secret, this information may enjoy trade secrets protection. Public entities as customers of big data analysis services will be less likely to have an interest in keeping the result of big data analysis secret. In the framework of emerging laws on open data, public institutions may even be under an obligation to provide access to the data both to the public and, pursuant to public-sector-information (PSI) laws, for commercial re-use by private actors.

II. Incentives for the commercialisation of data

81 Another and more modern justification for property rules is the goal of creating incentives for the commercialisation of the subject-matter of protection. In the context of patent law, this is often called the 'prospect theory'—in contrast to the traditional incentive theory, whereby the latter is designed to reward those who invest in the generation of the subject-matter for that investment.[91]

82 In general, innovation does not end with the generation of the subject-matter of protection and the acquisition of the IP right. Innovation will only serve society if it reaches the market. And quite often more investment will be needed for the commercialisation of the subject-matter of protection than for its generation.

83 A good example of this can be observed in the pharmaceutical sector. The major investment that goes into the development of drugs relates to the financing of the lengthy and risky clinical trials, which typically take place after the filing of patents. Indeed, in order to protect investment in the clinical trials against free-riding by others, the pharmaceutical company is in need of patent protection prior to making that investment. In most cases, the patent holder will also be the firm that conducts the clinical trials and brings the product to the market. However, the patent holder may also decide to license the patent to another company that, based on that licence and with the prospect of having a secured market later on, will make the investment in developing the drug.

84 Similarly, investment in the commercialisation of copyrighted works is not typically effectuated by the creator, but by the representatives of the copyright industries, such as publishers and producers. Only in countries that follow a work-made-for-hire doctrine will the latter be considered initial copyright owners, whereas in other countries they can rely on exclusive copyright licences or, at best, related (neighbouring) rights.

85 These examples show that the original right does not necessarily have to be vested in the person who makes the investment in the commercialisation. The licensing system, based on contract law and exclusive licences, can provide for the same incentives. Granting the original right at the stage of the creation of the content, however, may produce additional distributional effects. The copyright protected in favour of the creator may generate additional income for the creator, at least if there are additional rules in place that guarantee fair remuneration.

86 As regards the data economy, however, no case for recognising data ownership can be identified based on the goal of producing additional incentives for the commercialisation of the data. The major argument is that the holders of data do not have to be afraid that competitors will free-ride on investment in the commercialisation of their data. Likewise, there is not any particular risk that the data will be copied by competitors for the purpose of substituting the data holder's offer, nor does the grant of access to the data to others, such as big data analysts, involve particular investment by the data holders.

87 Nor are the big data analysts unable to recoup their investment in the commercialisation of their data

90 Against a justification of patent protection for the algorithm, see Josef Drexl et al. (*supra* n 39), paras 12-17.

91 The foundations of the prospect theory were laid by Edmund Kitch, 'The Nature and the Function of the Patent System' (1977) 20 *J L & Econ* 265. On a more modern market-related patent theory that departs from the classical reward theory, see also Daniel F Spulber, 'How Patents Provide the Foundation of the Market for Inventions' (2015) 11 *J Comp L & Econ* 271.

without data ownership. They are much more likely to rely on the control of their software solutions to protect their innovation under competitive conditions.

88 The situation is likely to be different as regards data brokers. Data brokers can play an important role in the enabling of big data analyses in particular.[92] Data brokers may also act as aggregators of datasets. Property rights have the potential of stabilising their activities. However, these brokers can also rely on factual exclusivity regarding the control of datasets that are transferred to them. Concerning situations where real-time data is key, data brokers are less likely to act as intermediaries that buy and resell identifiable datasets. They are more likely to act as agents that bring together providers of large and dynamic datasets with customers that are interested in services that build on big data analyses. Such brokers will enable direct transactions between data providers, on the one hand, and big data analysts and their customers, on the other hand. To do this, they are not in need of property rights in the data.

III. Data ownership as a means to stabilise transactions

89 Property rights can also stabilise and, thereby, facilitate transactions. Conversely, this is an effect which cannot be provided in the framework of trade secrets protection. Transactions on trade secrets suffer from major instability. Every sharing of trade secrets increases the risk that the information will ultimately become publicly available with no possibility for the holder of the trade secret to act against the re-use of that information.[93] Accordingly, recognition of data ownership is advanced as a means to facilitate trading with data as a commodity. The argument is that, even where there is factual exclusivity, without ownership there are no direct remedies against unauthorised use by third persons once the data has been disclosed.[94]

90 Yet considering the risk that business models will be undermined by unwanted free-riding in an environment in which the availability of real-time data is key, this argument of stabilising transactions will hardly ever be very convincing.

IV. Legal certainty

91 Another argument relates to legal certainty. Clear attribution of ownership can enhance legal certainty by informing the stakeholders about their rights and obligations.

92 This, however, is not very convincing as regards data ownership either. On the one hand, new property rights will always give rise to additional conflicts and litigation. At the same time, allocation of property rights may not be so clear at all. As regards data ownership that is recognised independently of factual control over data in an environment where individual data may constantly be integrated and arranged in different datasets, data ownership is more likely to reduce transparency and increase the risk of unintentional infringement of rights.

V. Ownership as a means to enhance access

93 A final potential justification for data ownership may look counterintuitive at first glance, but in particular deserves closer attention.

94 As has already been explained above in the context of the discussion of the *UsedSoft* decision of the CJEU,[95] property rights regimes can also be used as a means to enhance the free flow of data. In this decision, a limitation of copyright protection regarding digital downloads was used as a means to promote free circulation of digital copies of computer programs.

95 This example shows that general recognition of property rights can also make sense where factual exclusivity is already particularly strong. Adoption of a fully-fledged rights regime can include far-reaching mandatory exceptions and limitations that cannot be set aside by contractual restrictions.[96] For

92 See Federal Trade Commission, 'Data Brokers—A Call for Transparency and Accountability' (2014), available at: <https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf> (accessed 12 September 2016). The business models of data brokers were however heavily criticised in the US in particular, where those brokers have contributed to the spread of personal data and provided uncontrolled access of the government to personal data. See Chris J. Hoofnagle, 'How ChoicePoint and Other Commercial Data Brokers Collect and Package Your Data for Enforcement' (2004) 29 *NC J Int'l L & Com Reg* 595.

93 According to Art 3(3) of the new EU Trade Secrets Directive the 'use' of trade secrets is only unlawful under rather restrictive conditions, namely, when the user has acquired the information unlawfully or is in breach of a confidentiality agreement or any other agreement on how to use the information. Once the trade secret has become known to third persons, these persons can lawfully use the information.

94 See, in particular, Zech (*supra* n 24) at 60.

95 At C.V. above.

96 See also Becker (*supra* n 25) at 9 (assuming that the industry may even refuse to claim new legislation on data ownership since such legislation could provide more access than they

instance, such exceptions and limitations can also be found in the French legislation on the exclusive right of sports associations as regards the audiovisual exploitation of sporting events.[97] Hence, such ownership systems could provide better guarantees for access than reliance on general contract law based on the unrestricted principle of freedom of contract.

96 However, this approach is not without alternatives. Access can also be guaranteed by special legislation on access that takes precedence over contractual restrictions. As regards the commercial exploitation of sporting events, such access rules can be included in the general media law. Current EU law also enhances access to information held by public bodies. Thereby, the European rules on public sector information do not have to recognise ownership of public bodies in the information they hold in order to regulate the principles that apply to the licensing of the commercial re-use of such information.[98]

97 An interesting case is also presented by the current proposal of the Commission to introduce an unwaivable exception to copyright protection for carrying out text and data mining for the purpose of scientific research.[99] This proposal seems to prove the case that exceptions promoting access to data can easily be drafted within existing ownership systems. However, separate access legislation on data mining could also be drafted by building on the model of the proposal with application beyond copyright and with regard to other interests whenever the data holder has granted access to somebody in the framework of a contractual agreement. To do this there is no need to recognise data ownership up front.

98 An additional argument against adopting ownership as a means to enhance access arises from challenges regarding the form of regulation of such exceptions or limitations. There are two approaches, both of which are problematic. The first approach consists in a general clause similar to the fair use exception of US copyright law.[100] This approach has the advantage of general applicability but the disadvantage of lack of precision. It would hence cause legal uncertainty, give rise to legal disputes and potentially favour the interests of those parties that have less of a problem to finance litigation. As regards data ownership in particular, this approach has the additional drawback that it would have to be formulated in an extremely general way in order to be adaptable to the very different sectors of the data economy. Hence, it is very doubtful whether such a 'fair use' clause would really be able to enhance access in practice.

99 The second approach would consist in formulating a precisely defined exhaustive catalogue of exceptions and limitations that takes care of specific countervailing interests. However, this would require the legislature to fully anticipate the interests of a large number of potential stakeholders in highly diverse sectors of a data economy that is only just emerging.[101] There is a clear risk that legislation on exceptions and limitations would largely be postponed to the future, while the legislature would immediately adopt a strong rights system that goes beyond the restrictions data holders can implement under contract law. In sum, this approach would entail the risk of largely hampering the free flow of information without sufficient remedies for addressing problems of access.

100 In addition, balancing conflicting interests is more difficult for the legislature, where the question of who should be the owner remains a most difficult issue.[102] Whomever the legislature singles out as the right-holder, this will produce an additional negative impact on the interests of other stakeholders and may intensify a conflict of interests. In contrast, by choosing the alternative approach of balancing factual control over data by access-only legislation, the legislature will react to the conflict as it arises from the specific context of the market without intervention.

101 In sum, it seems more advisable to prefer an approach of progressive adoption of access regimes as part of sector-specific regulation. Such an approach could still develop principles and guidelines that emerge over time and ultimately rely on general models of regulation.[103]

102 It can be thus concluded that no reasons can be identified that would argue in favour of introducing data ownership in favour of any of the stakeholders.[104]

E. Problems related to the design of the rules on data ownership

103 Since there is no clear case for introducing legislation on data ownership, the question of how to design such legislation is not even relevant. Yet,

currently are willing to provide under contract law).

97 See Arts L333-6 through L333-9 *Code du sport*.

98 See PSI Directive (*supra* n 21).

99 Art 3 Commission Proposal of 14 September 2016 for a Directive of the European Parliament and of the Council copyright in the Digital Single Market, COM(2016) 593 final.

100 See Sec 107 US Copyright Act.

101 On the many and very context-dependent stakeholders in the data economy see at B.IV. above.

102 See at E.I. below.

103 On this see at F.IV. below.

104 Also against adopting legislation on data ownership, Wiebe (*supra* n 25) at 884.

some challenges regarding such legislation should nevertheless be addressed since, in the current debate, it seems that these challenges are not sufficiently discussed[105] and, consequently, largely underestimated when the idea of data ownership is advanced.[106] There are many reasons why the design of such protections is enormously complex. Several dimensions of this problem can be identified:

I. Complexity of the legal issues

104 For any intellectual property rights system, a decision has to be made on what subject-matter is to be protected, on who should own the right, and on the scope of protection, including the exceptions and limitations. As to the latter aspect, a decision is to be made regarding the terms of protection.

105 As regards the subject-matter of protection, it has already been mentioned that the law has to decide whether data should only be protected on the syntactic or also on the semantic level. The latter should rather be avoided because of the risk of obstructing the free flow of information.[107] However, the question still remains whether data can be protected as 'raw data' on the syntactic level. This is questioned because data is in need of specification on the semantic level in order to qualify as subject-matter of protection beyond the encoding in the form of bits and bytes.[108] If, however, protection was granted on the semantic level, the very practical problem is to identify whether information is 'new'.[109] Another issue is whether data ownership should relate to individual data or datasets in their entirety. The latter would follow the example of the Database Directive with all its shortcomings, namely, that it fails to protect the individual data. Yet, if each and every individual piece of data were protected, data ownership of individual persons in a world of big data would disappear like drops of rain in the sea. Such a system would present major challenges in terms of its governance and of the enforcement of myriad individual rights, not to mention the challenges for users in the context of rights-clearing.

106 As regards potential owners, it has been shown in this analysis that in a complex world of networks where a considerable number of different players collaborate in generating value, not least by contributing their data, the allocation of data ownership is particularly difficult.[110] Furthermore, if everybody contributing to the generation of data in a value network is vested with ownership, this allocation could easily run the risk of creating too many property rights, which would block efficient exploitation of big data in particular.[111] The proposition to vest consumers with the ownership of their personal data in order to enhance trading with that data as a commodity[112] does not explain why allocating the economic value to consumers can be justified from an economic perspective.[113]

107 Moreover, the definition of the scope of protection also remains a difficult task. It is not clear at all in which situations there is a particular risk that the need for investment will be distorted by the free-riding of third parties. The proposal to limit protection to the copying of encoded information, while allowing for the re-generation of the same data,[114] would only confirm that data should not be protected on the semantic level of information.

108 The definition of the subject-matter of protection, the identification of the owner of the right and the scope of protection will be most relevant for finally identifying the need for exceptions and limitations. In the light of the large number of stakeholders, it would be particularly difficult to clearly identify the conflicting interests and to design rules for balancing these interests.

109 The interaction between all of these issues reaches an enormous level of complexity, which argues in favour of preferring legislation on access regimes to the implementation of a fully-fledged new ownership system.[115]

105 See, however, the discussion of a data producer right by Zech (*supra* n 24) at 74-78.

106 This is also true for EU Commissioner Oettinger. His idea of a 'data use right' does not explain what this right should protect, who should be the owner and how far protection should go.

107 See also Zech (*supra* n 24) at 74 (delineating his data producer right only on the syntactic level). For a review of different proposals see Wiebe (*supra* n 25) at 882.

108 Wiebe (*supra* n 25) at 883.

109 See Wiebe (supra n 25) at 882, highlighting that this requires a showing that the same information has not been stored before in form of 0s and 1s. In addition, it ought to be remembered that the same information can be represented differently on the syntactic level, for instance, in a different language or a different form (eg, a video and not a text).

110 This is considered a main counterargument against devising a property right in data according to Wiebe (*supra* n 25) at 883.

111 See also Wiebe (*supra* n 25) at 883 (against co-ownership because of the conflicting interests).

112 See, in particular, Zech (*supra* n 24) at 60.

113 This is also conceded in principle by Zech (*supra* n 24) at 69.

114 In this sense Wiebe (*supra* n 25) at 882.

115 See also discussion on adopting an ownership regime as a vehicle for promoting access through exceptions and limitations at D.V. above.

II. The one-size-fits-all issue

110 In addition, legislation on data ownership would have to respond adequately to very diverse circumstances in which data is generated and used in the future. The data economy and the use of smart products are predicted to enter all different fields of modern life. However, data collection as regards the operation of smart cars is very different from the processing of data in the healthcare sector. Whether it is possible *ex ante* to conceive uniform rules on the subject-matter of protection, the person owning the rights and the uses that will be covered by the right, while the peculiarities of different sectors are delegated to exceptions and limitations, remains rather doubtful.[116]

III. The dynamic character of the data economy

111 Several times it has been underlined in this analysis that the data economy and big data in particular, is not about stable datasets but about the 'moving target' of highly dynamic data. 'Velocity' and 'veracity' are a fundamental concern in this economy.

112 This however questions the very appropriateness of a property approach to regulating that economy. IP systems are largely based on the paradigm of protecting intangible assets, such as technologies in particular, that play a role as input in the production of physical goods. Such a paradigm does not seem to fit a world in which customers have to rely on real-time and accurate information as an input. This contradiction becomes most obvious if one addresses the issue of the terms of protection. In an environment where it is key to capture the moment and where being late leads to wrong decisions, asking the question of how long data should be protected will simply miss the needs of this economy. Rather, the starting point of any legislation should be a clear analysis of the emerging new business models and the question of what kind of protection firms need in order to make their business models successful in competition with other firms and in the overarching interest of society.

113 As a matter of principle, contract law seems to provide the better regime for such protection. It allows the parties to specifically design the rights and obligations as needed for making new business models work. Contract law provides the parties with the possibility to experiment with different arrangements over time and with the flexibility to adapt to different circumstances in very different sectors of the data economy.

116 Similar doubts are expressed by Wiebe (*supra* n 25) at 884.

F. Regulating access to data

114 However, contract law cannot be expected to make the data-driven economy work without frictions. Contract law will only work in instances where the holder of data has an economic interest in sharing the data with others and where the bargaining power of the contracting parties is equally strong. Hence, the question arises whether government and legislative action is needed to promote access.

115 From the outset, it has to be clear that a refusal to grant access by itself is not sufficient to justify intervention. In line with the rationale of trade secrets protection, such refusal should not be considered illegitimate where exclusive control over data provides firms with a competitive edge over others and, thereby, creates the necessary incentive to invest in data-based business models. This also means that the leading firms of the data economy such as Google and Facebook should not blindly be forced to share their user data, the most valuable asset they have to conduct their business.

116 Striking the balance between access to and legitimate control of data is hence a most difficult task. The field of law that first comes to mind to tackle the issue is competition law. In this regard, a more thorough analysis of competition law is needed in order to assess competition law's potential to provide a workable access regime. For this purpose and as a preliminary clarification, it is important to place competition law as a tool for enforcing access to data in the context of the current competition policy debate on big data (section F.I. below). This will be followed by an analysis of the potential application of rules of EU competition law to refusals to grant access to data (section F.II. below). This analysis will help in discussing additional actions that could enhance access to data (sections F.III. and F.IV. below).

I. The current competition law debate on big data

117 The debate and literature on how and whether competition policy should react to the advent of big data has exploded within a remarkably short period of time.[117] The discussion is mostly driven by the enormous success and expansion of firms in the digital economy such as Google or Facebook, whose business models are largely built on the control of user data. There is in fact growing awareness that control over big data should play a more prominent

117 Among the major and most recent contributions from competition law scholars are Rubinfeld and Gal (*supra* n 28); Maurice E Stucke and Allen P Grunes, *Big Data and Competition Policy* (Oxford: Oxford University Press, 2016).

role in assessing market power and dominance, not least in the framework of mergers.[118] The EU merger cases of *Google/DoubleClick*[119] and *Facebook/WhatsApp*[120] are among the first cases where control over user data in terms of 'data concentration'[121] was taken into account for assessing the effects of mergers on the online advertising market.[122] Yet in both cases the Commission held that the emerging data concentration was not sufficient to significantly impede competition in this market.[123] The growing role of data in the digital economy has also convinced competition law enforcers to further develop their policies as regards the impact of control over data on competition.[124]

118 Yet this discussion on how competition law should react to the challenges of the data economy and big data is based on a particular perspective. First, control over data is considered to be a potential competition problem. This corresponds to the general role of competition law to ban anti-competitive conduct. Second, the focus is very much on market structure, market power and dominance,[125] as well as on market entry barriers arising from the control of big data.[126] This is explained by the fact that anti-competitive effect, especially in unilateral conduct cases, depends on the ability to behave independently of the competition.

119 Within the framework of the current 'Free Flow of Data' initiative of the Commission, however, the role attributed to government is a more proactive one of industrial policy. The question is not only how to protect the free market economy against anti-competitive conduct of firms. Rather, the question is what can be done in order to promote the digital economy.

120 In this regard, competition law has certain advantages but also shortcomings. On the positive side, competition law is in principle applicable to all sectors of the economy that are currently undergoing a digital transformation. Competition law can work as a platform on which legislatures can build to formulate more targeted, sector-specific rules whenever competition law does not provide sufficient remedies. In addition, competition policy and law can also prevent the legislature from excessive intervention. In instances where there is no identifiable harm to competition, policy makers will have to look for an alternative justification for adopting access rules.

121 On the negative side, competition law is likely to be too limited to provide sufficient remedies. As regards its substantive criteria, competition law only reacts to one particular kind of market failure. Intervention is only justified where there is identifiable harm to competition. While the outer boundaries of what can be considered such harm is not at all clear, there are kinds of market failures that cannot specifically be addressed by competition law. For instance, in a world of big data analytics involving techniques of data mining by searching datasets for correlations, negotiations about access to data may simply fail because of information asymmetries regarding the value of the data.[127] From an institutional

118 See, for instance, Inge Graef, 'Market definition and market power in data: the case of online platforms' (2015) 38 *World Competition* 473.

119 Commission Decision of 11 March 2008, Case No COMP/M.4731—*Google/DoubleClick*, available at: <http://ec.europa.eu/competition/mergers/cases/decisions/m4731_20080311_20682_en.pdf> (accessed 10 September 2016).

120 Commission Decision of 3 October 2014, Case No COMP/M.7217—*Facebook/WhatsApp*, paras 164-67 and 181-91, available at: <http://ec.europa.eu/competition/mergers/cases/decisions/m7217_20141003_20310_3962132_EN.pdf> (accessed 10 September 2016).

121 *Facebook/WhatsApp* (*supra* n 120) para 164.

122 From the perspective of the data economy, the Commission Decision of 4 September 2002, Case No COMP/M.6314—*Telefónica UK/Vodafone UK/Everything Everwhere/JV*, available at: <http://ec.europa.eu/competition/mergers/cases/decisions/m6314_20120904_20682_2898627_EN.pdf> (accessed 10 September 2016) may even be more interesting. In this case, the Commission assessed the impact of the joint venture for the introduction of an electronic payment system ('mobile wallet') on the market for data analyses.

123 In *Facebook/WhatsApp*, the Commission specifically looked at WhatsApp as a potential source of user data for better targeting Facebook's advertising activities. It finally concluded that even if Facebook implemented such a policy post-merger, it would only control a small share of user data on the Internet as a resource for online advertising. See *Facebook/WhatsApp* (*supra* n 120) paras 180-89.

124 See, in particular, the joint policy paper by of French and German competition authority: Autorité de la concurrence and Bundeskartellamt, 'Competition Law and Data' (10 May 2016), available at: <http://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.pdf?__blob=publicationFile&v=2> (accessed 10 September 2016).

125 In their joint policy paper on data, the French and German competition authorities devoted the whole second half to the role of data for assessing market power. See Autorité de la concurrence and Bundeskartellamt (*supra* n 124) at 25-52.

126 See, in particular, the thorough analysis of potential barriers to entry caused by big data by Rubinfeld and Gal (*supra* n 28).

127 This is known as the 'information paradox'. Contractual negotiations on data as a commodity can easily fail because the purchaser, not knowing which information can be extracted from the data, will not be able to assess the value of the data. If, however, the data is made accessible to the prospective purchaser for solving the information problem, the purchaser will no longer be willing to pay for access. The 'information paradox' was first framed by Arrow in the context of patent law. See Kenneth J Arrow, 'Economic Welfare and the Allocation of Resources for Invention' in: National Bureau of Economic Research (ed), *The Rate and Direction of Inventive Activity* (1962) 609. But it is also to be noted that markets can provide solutions to

perspective, competition law enforcers are able to ban identifiable anti-competitive conduct, but they are not well equipped for regulating markets *ex ante* by imposing positive rules of conduct in the form of behavioural remedies that require on-going monitoring.

122 Hence, already based on these general observations, it is very likely that actions will be needed that go beyond competition law. But competition law should be placed at the beginning of the following analysis (section F.II. below). Competition law thinking as a market-compliant approach will however also prove important for devising additional pro-competitive regimes that promote access to data (sections F.III. and F.IV. below).

II. Addressing refusals to grant access to data under EU competition law

123 EU competition law has not yet developed specific case-law on access to data in the data economy that is only now about to emerge. However, as the following analysis will show, the practice on refusals to deal and, more concretely, refusals to license can produce some indications on how to assess future data-related cases. At the outset, it should be noted that it is not important whether data to which access is requested is protected by intellectual property (IP) rights or not.[128] Even in cases in which neither IP protection nor trade secrets protection is available, but the holder of data can rely on factual exclusivity provided particularly by technological protection measures, a refusal to grant access can be captured as a refusal to deal under competition law. For the assessment of such cases, under Article 102 TFEU, the question is whether the holder of data is market dominant and whether the refusal to grant access to data constitutes an abuse. These issues will be addressed in the framework of the following review of the existing case-law.

124 The three major cases that established the foundations for assessing refusals to license, namely, *Magill*,[129] *IMS Health*[130] and *Microsoft*,[131] are all, in one way or another, 'information-related'. Beyond these three cases, the following analysis will also take into account the more recent *Huawei* case, which dealt with a refusal to license a standard-essential patent (SEP).[132]

1. The requirement of dominance

125 For cases regarding access to data in the context of the currently emerging data economy, *Magill* and *Microsoft* are most suitable precedents. In both cases, the holder of information that was indispensable for entering a downstream market refused to grant access to that information. In *Magill*, the TV stations broadcasting in the Republic of Ireland and Northern Ireland refused to grant a copyright licence for their TV listings and thereby excluded a publisher from the market who intended to offer comprehensive TV guides to consumers. *Microsoft* is perhaps an even better precedent for refusals to grant access to data because, in this case, the interoperability information for the Windows operating system as such was not freely available to the competitors in the market for work group server operating systems.[133] Yet *Magill* laid the foundations for dealing with the issue of information-based dominance. The Court convincingly stated that, due to copyright protection, the TV stations were the only source of the relevant information and that, therefore, the three TV stations had to be considered as *de facto* monopolists with regard to the information contained in their respective TV listings.[134] The situation in *Microsoft* was very similar. However, market dominance did not arise from an IP right, but from the fact that Windows, based on network effects, had emerged as a *de facto* standard in the market for operating systems, which made the

the information paradox. For instance, data analysts can be appointed as trustees to do tests on the utility of datasets for the purposes of a prospective customer to assess the value of the dataset, without providing direct access to the information contained in the datasets to this customer.

128 In the *Microsoft* case, which was on access to the interoperability information contained in the Windows program, both the Commission and the General Court (GC, former Court of First Instance) left open whether this information was IP-protected or not and applied the test developed by the Court of Justice of the EU (CJEU) for refusals to license an IP right. See Judgment in *Microsoft v. Commission*, T-201/04, ECLI:EU:T:2007:289, [2007] ECR II-3601.

129 *Magill* (*supra* n 27).

130 Judgment in *IMS Health*, C-218/01, ECLI:EU:C:2004:257, [2004] ECR I-5039.

131 *Microsoft* (*supra* n 128).

132 Judgment in *Huawei*, Case C-170/13, ECLI:EU:C:2015:477.

133 Art 6 Computer Programs Directive (supra n 79) allows for decompilation of an existing computer program where this is necessary to obtain interoperability information for the purpose of establishing interoperability for an independently created computer program. However, this exception and limitation is insufficient in a modern software environment, where the interoperability information can constantly be changed by updates. Hence, competition law may still be needed to order the dominant holder of a computer program to provide access to the interoperability information. Recital 17 of the Computer Programs Directive explicitly safeguards the applicability of EU competition law in such instances.

134 *Magill* (*supra* n 27) para 47.

interoperability information an indispensable input for offering interoperable programs that would run on Windows.

126 The two cases demonstrate that it is easiest to show dominance in data-related cases where the petitioner seeks access to concrete information that is indispensable for doing business in a market.

127 More typical for the data-driven economy are however cases in which somebody, such as a big data analyst, seeks access to large datasets for purposes of data mining. In the light of its utility, namely, to rely on statistical correlations among different pieces of information contained in larger sets of aggregated data for generating new knowledge, such datasets have to be considered a kind of resource which is distinct from concrete semantic information such as in the case of *Magill*. Yet the test of *Magill*, as an expression of general competition law principles, can be adapted to meet the challenges of cases that deal with access to large datasets to enable big data analyses. The test in both cases is whether the respective dataset can be considered the 'only source' of the resource.

128 This leads to the issue of substitutability of datasets. The fact that data are non-rivalrous and, therefore, individual data could be found in various datasets seems to count against dominance. Whether datasets are substitutable or not will depend on the concrete circumstances, including the very nature of the information contained in the data. If, for instance, a supplier of parts wants to have access to the data collected by the end manufacturer after the sale of the final product to control the quality of its parts, the end producer's datasets will indeed be the only source of that data. However, if a city is in need of information about the qualities of streets which is collected by smart cars, different car manufacturers may be able to provide access to that information through their datasets. The reason is that the latter kind of information is freely available in the public in the first place, and, hence, can be duplicated in the datasets of any other data collector. Publicly accessible information is by nature non-rivalrous[135] and can therefore be registered by anybody in a digital format.

129 Yet assessing dominance in a world of big datasets by using the concept of substitutability remains a most difficult task, since even the petitioner for access, such as a big data analyst, will often only have a vague understanding about the kind of data contained in the dataset and about which data will produce the most valuable new information based on observable correlations.

130 However, larger collections of data will generally guarantee a higher level of accuracy of the new information, since such information derived from correlations of data within such datasets is based on statistical likelihood. Hence, just as in the case of multisided platform markets, the collection of datasets for the purpose of enabling big data analysis may exercise particular network effects and enhance market power of the firm that controls access to the larger dataset.[136] The same may occur in the case of data-sharing platforms. An example of such a platform is provided by the joint venture of the three German car manufacturers, Daimler, BMW and Audi, that acquired Nokia's digital map HERE as an important element of their systems for autonomous driving. For instance, such digital platforms could be used for collecting and exchanging real-time information about the weather conditions of roads. The quality and reliability of such an information-sharing platform would obviously increase with the number of cars contributing information to this system. Accordingly, the three car manufacturers should have a strong self-interest in convincing other car manufacturers to join the system.[137] At the same time, this may tip the market and give rise to market dominance of the joint venture.

2. The four requirements for abuse according to Magill and IMS Health

131 The two cases of *Magill* and *IMS Health* have established the European test for assessing whether a refusal to license constitutes an abuse. In *IMS Health* this test was phrased as one with three cumulative conditions, which, however, contained the additional underlying condition that the resource to which access is sought be indispensable for conducting a business.[138] In *Microsoft*, the General Court rephrased this test in a better and more structured way.[139] According to the Court, the following four conditions for a refusal to license need to be fulfilled in order to present 'exceptional circumstances' for considering the refusal an abuse:

135 The character of non-rivalry of data is also highlighted by Autorité de la concurrence and Bundeskartellamt (*supra* n 124) 36-37.

136 See also Rubinfeld and Gal (*supra* n 28) at 42.

137 Indeed, when the *Bundeskartellamt*, the German competition agency, cleared the acquisition under German merger control law, it specifically considered that other car manufacturers would not be excluded from participating in the system. See Bundeskartellamt, 'BMW, Daimler and Audi can acquire Nokia's HERE mapping service' (6 October 2015), available at: <http://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2015/06_10_2015_HERE.html?nn=3591568> (accessed 10 September 2016).

138 *IMS Health* (*supra* n 130) para 38.

139 *Microsoft* (*supra* n 128) para 332.

(1) The refusal relates to a product or service that is indispensable to the exercise of a particular business in a related (secondary) market;

(2) The refusal excludes effective competition in that related market;

(3) The refusal prevents the emergence of a new product for which there is consumer demand;

(4) The refusal is not objectively justified.

132 In applying these conditions to refusals to grant access to data and larger datasets in particular, several issues arise:

133 First, as regards the indispensability requirement, a problem arises when data relate to information that it is publicly accessible but can only be found in a digital format in the datasets of one undertaking. Since registration and digitisation makes the information retrievable and treatable, including for purposes of big data analysis, the digital data should be considered a product with added value that differs from the original, publicly accessible information. Accordingly, the holder of the digital data in such a situation can indeed be considered a monopolist and, hence, a potential addressee of Article 102 TFEU. However, this does not automatically mean that the data is also 'indispensable' in the *Magill/IMS Health* sense, since anybody else including the petitioner could also register the same information in a digital format.

134 For understanding the concept of indispensability, the judgment in *Bronner* is most relevant; although the case did not deal with access to data but access to a nationwide home delivery scheme for newspapers. According to the CJEU in this case, access to a resource of a competitor cannot be considered indispensable if there are no 'technical, legal or even economic obstacles capable of making it impossible, or even unreasonably difficult' to duplicate the resource.[140] Thereby, the Court showed reluctance to accept the argument of lack of economic viability too easily. The Court stressed that it is not enough to show that duplication of the resource would not be economically viable against the benchmark of the petitioner's scope of business in the secondary market.[141] Rather, the question is whether it is economically viable to create the resource 'for production on a scale comparable to that of the undertaking which controls the existing product or service'.[142]

135 This seems to indicate an objective standard for indispensability that does not depend on the size of the petitioner's business and that imposes on the petitioner the burden to make the same investment as the one made by the dominant undertaking. Regarding cases on refusal to grant access to data, this may well mean that indispensability cannot be argued where the information as such is freely accessible and it is only a matter of registering the data in a digital form. On the other hand, it would be easier to argue indispensability where data is generated through business models that are characterised by strong network effects such as search engines and Internet platforms like the HERE data-sharing system described above. The possibility to duplicate similarly large and valuable datasets is excluded by the economic characteristics of these markets.[143]

136 Second, the requirement of excluding competition in a secondary market qualifies the European rule on refusal to licence as one, which is based on a leveraging and exclusion theory. This presupposes that the dominant firm is also active as a competitor in the secondary market. This, however, will frequently not be the case when firms refuse access to data. It is a typical feature of the new data economy that data is collected for one purpose, such as enabling predictive maintenance services, but turns out to be interesting for very different purposes pursued by other firms of a very different sector and even the government. In such instances, the European rule on refusals to license and refusals to deal, as developed in the abovementioned case-law, would not apply.

137 More recently, in the *Huawei* judgment, the CJEU clearly indicated that the 'cumulative' *Magill/IMS Health* conditions are not the only 'exceptional circumstances' to make a refusal to license an abuse. the CJEU accepted that exceptional circumstances are also present in the case of a refusal to license an SEP if (1) the standard was fixed by a standardisation body[144] in return for which (2) the patent holder has irrevocably committed to license on fair, reasonable and non-discriminatory (FRAND) terms.[145] Since the Court did not reiterate the condition of exclusion of competition in a secondary market as part of the description of these exceptional circumstances, the question may be asked whether a refusal to license or a refusal to deal can also be considered abusive if the dominant firm is not vertically integrated. However, the *Huawei* decision itself presents many uncertainties in this regard, because the Court in its reasoning still indicates that harm to competition

140 Judgment in *Bronner*, C-7/97, ECLI:EU:C:1998:569, [1998] ECR I-7791, para 44.

141 *Ibid*, para 45.

142 As rephrased in *IMS Health* (*supra* n 130) para 28, with reference to *Bronner* (*supra* n 140) para 46.

143 This problem of 'access to data', though not in the context of the indispensability test, is also addressed by Autorité de la concurrence and Bundeskartellamt (*supra* n 124) at 38.

144 *Huawei* (*supra* n. 132) para. 49.

145 *Ibid*, para 51.

is conceived as harm through exclusion of a competitor in a downstream market. In particular, the Court reasoned that 'the fact that the patent has obtained SEP status means that its proprietor can prevent products manufactured by competitors from appearing or remaining on the market and, thereby, reserve to itself the manufacture of the products in question'.[146] From this, one could conclude that exclusion of competitors in a secondary market also remains a requirement in SEP cases. This would indeed be correct if one accepted the conservative approach to competition law, according to which competition law can only promote innovation indirectly, namely, only in cases in which there is identifiable harm to competition through exclusion.[147] In contrast, the Commission also argued a violation of Article 102 TFEU in the *Rambus* case against a non-vertically integrated SEP holder who tried to extract excessive royalty rates from the implementers in a case of patent ambush.[148]

138 This debate, however, may not be very relevant for cases on access to data. The reasons for this are twofold. First, those cases do not involve SEPs related to standards adopted by a standardisation body. Hence, the alternative 'exceptional circumstances' accepted in *Huawei* will not apply. Second, the alternative, dealing with refusals to grant access to data by non-vertically integrated data holders as a pure case of exploitative abuse in the form of excessive pricing under Article 102 lit. a) TFEU, would turn competition law enforcers into general price regulators. Fulfilling such a role would particularly be difficult in cases on access to data in which the parties also encounter major information problems as regards the economic value of data contained in large datasets. Accordingly, it is very unlikely that a claim of abuse of market dominance will be successful in a case where access to data is sought and the holder of those data is not a competitor of the petitioner in the secondary market in which the petitioner wants to use those data. This would exclude reliance on competition law in two very important sets of cases. The first case concerns big data analysts who seek access to data for purposes of data mining. The holders of such data will typically not be active as competitors in the market of providing new information generated through big data analyses. The second case regards cases where the government seeks access to data in the public interest. In such cases, a secondary market is missing in the first place, since the government will not make use of that data as an undertaking in the sense of EU competition law.

139 Third, the question is whether the requirement of the prevention of a new product (so-called 'new product' rule) also applies to cases of a refusal to grant access to data. According to the General Court in *Microsoft*, this is an additional requirement that only applies to cases involving the refusal to license an intellectual property right, but not to general refusal-to-deal cases.[149] As demonstrated further above,[150] it is very unlikely that data are already protected by intellectual property rights. The judgment in *Magill*, where access to the relevant information was controlled by a copyright, can only be explained by the very low standards of copyrightability under the British and Irish copyright case-law of that time, which most likely can no longer be maintained against the backdrop of more recent copyright decisions of the CJEU.[151] To the extent that there is trade secrets protection, the question is still left unanswered by the European Courts whether the test on refusals to license an IP right would also apply.[152] Yet if the European legislature decided to create a new intellectual property right in data, this may well make it more difficult to control access to data based on European competition law since, then, there should be less doubt as to whether the additional requirement of the prevention of a new product applies.

146 *Ibid*, para 52.

147 This is indeed the approach advocated by Pablo Ibáñez Colomo, 'Restrictions on Innovation in EU Competition Law' = LSE Law, Society and Economy Working Papers 22/2015 (2015), available at: <http://ssrn.com/abstract=2699395> (accessed 14 May 2016).

148 Commitments Decision of the Commission 9 December 2009, Case COMP/38.636—*Rambus*, available at: <http://ec.europa.eu/competition/antitrust/cases/dec_docs/38636/38636_1203_1.pdf> (accessed 10 September 2016). The Commission's approach is supported by Josef Drexl, 'Innovation as a Parameter of Innovation and its Implication for Competition Law Application', Paper presented at the 11th ASCOLA conference (30 June 2016) (forthcoming) (in favour of protecting dynamic innovation competition beyond cases involving exclusion).

149 *Microsoft* (supra n 128) para 334.

150 At C. above.

151 The CJEU requires that there be scope for the author to make 'free and creative choices', by way of which the author 'stamps the work created with his personal touch'. See Judgment in *Football Association Premier League and Others*, C403/08 and C429/08, ECLI:EU:C:2011:631, [2011] ECR I9083, para. 98; Judgment in *Painer*, ECLI:EU:C:2011:798, [2011] ECR I12533, paras 89 and 92; Judgment in *Football Dataco v Yahoo! UK* (*supra* n 46) para 38.

152 In 2005, under the impression of the *Microsoft* case, the Commission argued that applying the standard developed for refusals to license an IP right 'may not be appropriate' in cases on a refusal to grant access to interoperability information that is protected as a trade secret. See Commission, 'DG Competition discussion paper on the application of Article 82 of the Treaty to exclusionary abuses' (December 2005), available at: <http://ec.europa.eu/competition/antitrust/art82/discpaper2005.pdf> (accessed 10 September 2016). For arguments in favour of such a distinction see Gintarė Surblytė, *The Refusal to Disclose Trade Secrets as an Abuse of Market Dominance—Microsoft and Beyond* (Berne: Staempfli, 2011) 173-210.

140 More specifically, in the context of the data-driven economy, many complex issues would arise in applying the new-product rule. From the outset, it is to be remembered that this rule presupposes that both the data holder and the petitioner for access are competitors in the same secondary market. Only under this condition does the question make sense whether the petitioner for access would offer a 'new' product as compared to the product of the dominant firm. In cases on access to data, the product offered by the entity that seeks access to data can be enormously diverse. If it is about use of the data by big data analysts, the new product will consist of new knowledge or information, which may then be offered in a secondary information market. How to apply the concept of a 'new product' in relation to different information is rather unclear. To argue that the information produced by the petitioner differs from that produced by the data holder may seem convincing at first glance. However, this is less clear in the light of the rationale of the new-product rule, which is based on a balancing of the interest in protecting competition with the interest in protecting the intellectual property right. Accordingly, the new-product rule was devised to guarantee that the IP right, which is meant to promote innovation, can only be restricted if the petitioner for the licence is also an innovator.[153] However, whether the generation of (any) new information can be considered innovation, remains rather doubtful. Of course, data may also be used to offer diverse goods and services in secondary markets. Access to data may especially lead to the improvement of goods and services. Yet it is not settled whether any improvement of a product can be considered a 'new' product. In *Microsoft*, the General Court seemed to argue this way by pointing out that, according to Art 102 lit. b) TFEU, there is not only an abuse when the dominant undertaking limits production, but also in the case of a limitation of 'technical development' to the prejudice of consumers.[154] It is to be noted that the new-product rule would also exclude application of competition law to public entities that seek access to data in the public interest where these entities do not engage in any economic activity in the sense of the concept of an undertaking under EU competition law.

141 Fourth, as regards potential justifications, it is still very unclear whether and what kind of efficiencies can be considered in the framework of an efficiency defence in cases of a refusal to grant access to data.[155]

153 See *IMS Health* (*supra* n 130) paras 48-49.

154 See *Microsoft* (*supra* n 128) para 648.

155 See only Stucke and Grunes (*supra* n 117) ch 19 (at 302-12) on 'data-driven efficiency claims' (however with a particular focus on the efficiency defence in merger control law).

142 In sum, the analysis of the case-law on refusals to licence under EU competition law produces a number of limitations und uncertainties.The requirement to show market dominance based on control over larger datasets presents particular challenges for assessing whether different datasets can be considered as substitutes. The case-law so far can only be applied with certainty to vertically integrated data holders, while, in many instances, the petitioners for access and the data holder will not be competitors in any markets. The case-law will not provide any remedy when government bodies seek access to data in the public interest. The rule on exploitative abuse (Article 102 lit. a) TFEU) will hardly fill the gap since it would require competition law enforcers to act as price regulators where it is extremely difficult for the parties themselves to assess the value of data. Hence, this analysis highlights the shortcomings and uncertainties of the current state of competition law to provide adequate remedies against refusals to grant access to data in the data-driven economy.

3. Access to indispensable tools for data treatment

143 The analysis so far has concentrated on access where data or whole datasets are an indispensable input. However, the European case-law on refusals to license has more to offer.

144 In *IMS Health*, the CJEU used the *Magill* judgment as a template for assessing a case that nevertheless presented very distinct features.[156] The reason for doing this was that an intellectual property right, namely, a copyright protecting a database, was at stake and this made *IMS Health* a refusal-to-license case similar to *Magill*.

145 As a precedent for cases relating to the data-driven economy, it should however be noted that the subject-matter of copyright protection in *IMS Health* was characterised by a particular functionality. The so-called 1860-brick structure, representing a map of Germany subdivided into 1860 geographical sectors, was used as a tool for collecting and treating data on the sale of drugs. IMS Health was dominant in the service market for the collection of sales data to assist the pharmaceutical companies in designing their marketing activities. A smaller competitor encountered major problems entering the market with its own 'structure' since the pharmaceutical companies refused to work with a different structure. The reason for this was that IMS Health's brick structure had emerged as a *de facto* standard in the industry, which led the smaller competitor to simply use the 1860-brick structure; this competitor was

156 *IMS Health* (*supra* n 130).

then sued by IMS Health for copyright infringement in Germany. In this context, the question arose whether the defendant could rely on a competition-law defence.

146 By only looking at the fact that the brick structure was protected by copyright law, the CJEU missed the point that the case was indeed one on *de facto* standardisation regarding the tools used in the relevant service market for collecting data. Therefore, the distinction between two related markets, the upstream licensing market and the downstream product market, as well as the application of the leveraging theory based on an extension of market dominance from the upstream to the downstream market, appears rather formalistic.

147 As regards cases on access to data, *IMS Health* produces the particular insight that the tools for treating data have a tendency to emerge as *de facto* standards since they allow data to be communicated between the different market participants involved at the different levels of the value chain of treating and analysing data. Use of the same tools in the industry will produce positive network effects. On the downside, *de facto* standardisation will create access problems regarding the use of these tools. These tools will regularly be software-based and hence protected by copyright law. Market participants that are not allowed to use these tools will encounter difficulties to enter the market for the treatment of such data.

148 The *IMS Health* judgment would directly apply to such cases. From a competition policy perspective, the CJEU should have given more weight to the fact that the IP right controlled access to a standard with a foreclosure effect on competitors. This places cases such as *IMS Health* in between *Magill* and *Huawei*.[157] The question in such cases is whether the new-product requirement makes sense in the first place.[158] Also in *Huawei*, the CJEU did not require the prevention of a new product for considering the refusal an abuse.

149 Of course, the better option would be to promote standard-setting through standard-setting bodies and licensing of such standards regarding the tools for data treatment on FRAND terms. To the extent that such standards will emerge, the *Huawei* judgment would become directly relevant.

157 *Huawei* (*supra* n. 132).

158 This has already been questioned by Josef Drexl, 'Intellectual Property and Antitrust Law. IMS Health and Trinko—Antitrust Placebo for Consumers Instead of Sound Economics in Refusal-to-Deal Cases' (2004) 35 *IIC* 788.

4. Learning from the judgment in Huawei

150 Indeed, the judgment in *Huawei* can also provide inspiration for dealing with cases on access to data. As regards SEPs, the problem is that patent holders enter into a FRAND commitment *vis-à-vis* the standard-setting organisation (SSO) when the patents are notified as standard-essential, but later no agreement can be reached between the patent owner and the standard implementer on the concrete royalty rate. Such disputes are prone to being affected by strategic behaviour by either party of the licensing negotiations. Since rights-clearing is enormously difficult in the telecommunications industry, which is characterised by several thousands of declared SEPs held by multiple right-holders, to require users to wait with implementation until they have cleared all rights would considerably delay implementation of the standard in the industry. At the same time, the FRAND declaration creates a legitimate expectation that the licence will be granted. However, once the user has started to implement the standard by producing standard-compliant goods, the SEP holder may try to extract excessive royalty rates by challenging the implementers with claims for injunctive relief (so-called 'patent hold-up'). Conversely, if injunctive relief is not granted at all, implementers can be tempted to reject any licence offer as non-FRAND-compliant so as to avoid any payment (so-called 'patent hold-out'). In order to strike a balance of interest, in *Huawei*, the CJEU devised a framework for negotiations that includes duties of both parties,[159] and this may help the parties reach an agreement without having to call upon the courts or arbitration tribunals to make a decision on the appropriate royalty rate.

151 In a data-related access dispute, one of the major difficulties may be that the parties are not easily able to agree on price. Hence, devising a negotiation framework for the parties similar to *Huawei* may assist the parties to reach an agreement. Such schemes could be implemented through private institutions—by way of private ordering—or through state regulation. This leads the analysis to the design of additional legislative measures to promote access.

III. Access regimes for existing contractual relations

152 As regards access regulation, a distinction can be made as to whether the parties already entertain a contractual relationship or not. Problems of access to data may also arise within existing contracts. The typical justification for legislative

159 *Huawei* (*supra* n. 132) paras 60-68.

intervention in contractual relations beyond the realm of competition law is unequal distribution of bargaining power.

153 Unequal bargaining power is addressed by different parts of the law. In particular, the EU has adopted such rules on consumer contract law in the form of the Directive on Unfair Contract Terms.[160] The Directive's scope of application is broad enough to also control standard contract terms on the treatment of data. However, there are also particular shortcomings. First, the Directive's general clause on unfair terms does not provide any guidance on how to assess clauses that relate to the collection and use of data. The indicative list of unfair contract terms in the Annex to the Directive does not respond at all to the modern challenges of a data economy. Second, since the application of the Directive is limited to consumer contracts, it fails to create a European legal framework for addressing the regulation of access to data in B2B cases.

154 However, as regards both B2C and B2B relations, there are alternative ways to address cases of unequal distribution of bargaining power.

155 As regards consumers, there is a considerable overlap of consumer law with data protection law. The rule on data portability in Article 20 of the General Data Protection Regulation[161] can rather be considered as one of consumer protection than of data protection. While the relevant data covered by Article 20 is personal data as protected by the Regulation in general, the purpose of the data portability provision is not to protect the individual's moral interests. Rather, the rule is designed as an access rule that will enable the individual to switch to other suppliers where access to the data is crucial for competition to work.[162] The German *Monopolkommission*, which, as a commission of competition experts, fulfils an advisory role to the German government, supported the right to data portability by stressing that it has the potential to help the individual overcome a lock-in effect[163] and to react to the problem that businesses, without ownership regulation in place, often claim control over personal data as part of their contractual arrangements.[164]

156 This rule was inspired by the situation of platforms, including social platforms that rely on user data. Yet it will prove particularly effective in the context of new data-driven business models built on the collection of data. For instance, car insurance companies have already begun to lower premiums of customers who accept digital registration of their driving habits.[165] The possibility to switch to another insurance company will be considerably enhanced by the possibility to use such data to prove that the customer is indeed a careful driver.

157 Since this rule on data portability constitutes a most suitable form of pro-competitive regulation, there is no reason why the right to data portability should be limited to personal data.[166] The lock-in effect is not necessarily restricted to such data.[167] Beyond consumer contracts, a lock-in problem can also arise with regard to industrial data where suppliers want to take data with them concerning the quality and longevity of their parts after the termination of the supply contract with the manufacturer of the final product. Hence, data portability rules should also be considered for industrial relations.

158 Yet use of access to data as regards the relationship between suppliers and an end producer could also be addressed as part of specific competition

160 Council Directive 93/13/EEC of 5 April 1993 on unfair contract terms in consumer contracts, [1993] OJ L 95/29.

161 General Data Protection Regulation (*supra* n 17).

162 The pro-competitive character of this provision was specifically highlighted and praised prior to the adoption of the Regulation by the German Monopolkommisson (Monopolies Commission) in its Special Report of 2015. See Monopolkommission, 'Competition Policy: The Challenge of Digital Markets', Special Report No. 68 (2015) paras S15, S37 and S105, available at: <http://www.monopolkommission.de/images/PDF/SG/s68_fulltext_eng.pdf> (accessed 10 September 2016).

163 *Ibid*, at para S105.

164 *Ibid*, at para S106.

165 On this see, for instance, Adam Tanner, 'Data Monitoring Saves Some People Money On Car Insurance, But Some Will Pay More' (2 September 2013), available at: <http://www.forbes.com/sites/adamtanner/2013/08/14/data-monitoring-saves-some-people-money-on-car-insurance-but-some-will-pay-more/#7bc2c423264a> (accessed 10 September 2016).

166 The French Parliament has just adopted a provision on data portability that builds on Art 20 General Data Protection Regulation (*supra* n 16) in Art L 224-42 of the *Code de la consommation* (Consumer Act) through the so-called *Loi Lemaire* (*Loi pour une République numérique*; Law for a digital Republic). The law was adopted by the *Assemblée nationale* on 20 July 2016 and finally approved by the French Senate on 28 September 2016; available at: <https://www.senat.fr/leg/tas15-131.html> (accessed 30 September 2016). See comments on Art 12 in the English Explanatory Memorandum, available at: <https://www.republique-numerique.fr/pages/digital-republic-bill-rationale> (accessed 10 September 2016).

167 Indeed, the new French portability rule is not limited to personal data. The new Art L 244-43-3 of the Code de la Consommation (Consumer Code), as amended by the *Loi pour une République numérique*, seems to apply to any data provided by a consumer. However, the rule is also more restricted than the General Data Protection Regulation in that it only applies where data are provided to an online service communication service provider (*fournisseur d'un service de communication au public en ligne*). This rule seems to apply to social platforms in particular, but not necessarily to a car insurance company, as in the example mentioned above.

law regulation. Regulation of supply traditionally forms part of the Block Exemption Regulation in the Motor Vehicle Sector.[168] In times of the advent of autonomous driving, a modernised regulation could also address the treatment of data on the functioning of the vehicle between the supplier of parts and the manufacturer of the vehicle. There is a particular risk that the latter, by relying on superior purchaser power, will implement contract terms on data treatment concerning parts that disadvantage the supplier. The question will be how to implement such rules within the framework of the Regulation. While the Regulation will continue to build on the market-share approach as a basis for the block exemption, restrictions regarding the access of data to the disadvantage of the supplier, including a restriction on data portability, could be included in the black list of hard-core restrictions. However, for formulating such a rule, precision is needed in order to clearly delimit the non-exempted clauses from those that can be exempted. In particular, one could imagine a rule that a supply contract cannot be exempted if it does not include a rule on free-of-charge data-sharing with the supplier concerning the functioning of the parts delivered by the supplier. Such a rule is justified by the fact that both parties belong to the same network that contributes to the generation of economic value.[169]

159 Of course, the issue of access to data by a supplier of parts is not specific to the motor vehicle industry. Hence, the Commission should consider creating a generally applicable access regime in favour of suppliers in the framework of its block exemption regulations.

160 Finally, the legislature is free to draft targeted rules that would ban contractual restrictions on the use of data under particular circumstances. The already mentioned Commission's proposal for an un-waivable copyright exception for text and data mining for purposes of scientific research provides such an example, which could be extended beyond the realm of copyright and applied for other purposes.[170] In this regard, Article 3(1) Commission Proposal for a Directive on Copyright in the Digital Single Market requires that a research organisation wanting to conduct text or data mining have legal access—typically based on a copyright licence—to the relevant subject-matter.

168 Commission Regulation (EU) No. 461/2010 of 27 May 2010 on the application of Article 101(3) of the Treaty on the Functioning of the European Union to categories of vertical agreements and concerted Practices in the motor vehicle sector, [2010] OJ 129/52.

169 On the new paradigm of 'value networks' see at B.III. above.

170 See at n 99 above.

IV. Access regimes outside of existing contractual relations

161 Regimes for access to data outside of existing contractual relations are more difficult to devise. In this field, a more cautious approach is needed in order to avoid excessive intervention in the market economy. In addition, the particularities of very different sectors where data is currently starting to play a major role in generating economic value from the outset seems to argue against a regime of general applicability. On the other hand, designing regimes for access to data is not an unprecedented exercise for legislatures. Existing models can be considered and discussed for cautious generalisations and potential transfer to other sets of cases.

162 In any event, devising access regimes outside of existing contractual relations depends on using certain criteria to balance the interests involved between exclusivity and access. Such criteria can be discussed as the kind of information contained in data, the identity of the data holder and the business model through which it generates data and, finally, the person or entity seeking access and the kind of use this petitioner is intending.

1. Kinds of information

163 As regards the kind of information contained in data, a first distinction could be made between information access to which is in the public interest—such as information that helps to fight infectious diseases—and other information in which there is only a commercial interest. Such a distinction, however, is very difficult to make, since information that seems commercial at first glance may still help the state to make decisions in the public interest. Hence, as regards 'public interest data', it is better to address this issue further below in the framework of the discussion of who is seeking access to data and for which purpose the data will be used.

164 Yet there are examples where access to information is promoted by specific legislative means based on the nature of the information. This is the case in particular as regards scientific information contained in publications. Access to such information is often controlled by academic publishers who seek an exclusive licence also with regard to the digital exploitation of the publications. In contrast, governments increasingly promote open-access publications. The tools used in this regard can be very diverse.[171] One approach consists of setting

171 As regards the European open access policy see Commission Recommendation of 17 July 2012 on access to and

financial incentives. In instances where the scientific information is the result of government-funded research, a commitment to open-access publication of the recipient can be made a requirement for the grant decision.[172]

165 Furthermore, open-access regimes can also be promoted through copyright law. In Germany, the legislature recently adopted a so-called 'secondary publication right', which vests the author with an un-waivable right to make the work available online after an embargo period of 12 months if the publication is the result of research activity that is at least 50 per cent publicly funded and provided that the second publication does not serve any commercial purpose.[173] The French legislature has just introduced similar legislation as part of its '*Loi Lemaire*' (*Loi pour une République numérique*).[174]

166 Such a secondary publication right is characterised by making use of the interest—namely, in reputation—of one stakeholder, namely, the author, to promote open access against the interests of another stakeholder, namely, the publisher. In doing so it indirectly benefits users, who get unrestricted benefits. Hence, this model has the advantage of promoting open access much more effectively than by requiring each and every user to claim access. This model could be transferred to other sets of cases where there is conflict of interest between two parties contributing to the information and where one party in contrast to the other is interested in open access. One such case regards libraries and other cultural heritage institutions that cooperate with private businesses such as Google in the digitisation of their public domain materials and works. While the private partner would usually be interested in exclusive exploitation, the cultural heritage institution will typically prefer open access.[175]

167 It is also to be noted that particular access features of the secondary publication right are also shared by the data portability rule of Article 20 Basic Data Protection Regulation (see section F.III. above). Moreover, in the latter case, two persons contributing to the collection and generation of digital data have opposing views on access of third parties to the data. In both cases, the law strengthens the rights of the person in favour of access, which will indirectly benefit third parties. From this perspective, these rules can be qualified as enacting partial, pro-access property rights. The legislature refrains from creating an exclusive ownership right relating to personal data under the Basic Data Protection Regulation that would allow the owner to prevent third parties from using those data,[176] but still promotes access of third parties based on the rights of the person from which the data originate. The un-waivable right is limited to the right to make the data available to third parties. In this context, also the recognition of copyright exhaustion for downloads of computer programs by the CJEU in the *UsedSoft* case comes to mind.[177] In this case, 'access' in form of tradability of the programs was enhanced by recognising ownership in the digital of the program downloaded by the licensee.

2. The data holder and its business model

168 Another distinction can be made concerning who holds the data and what business models they use. Access can be promoted by legal regimes that focus on particular groups of data holders.

169 Legislatures can in particular promote access to data where data is held by public institutions as part of an open-data policy. At the EU level, the Public Sector Information (PSI) Directive of 2003, in its revised version of 2013,[178] provides an evolving approach for the EU to overcome resistance among public bodies in the Member States to make data more accessible to the private sector.

170 As part of the *Loi pour une République numérique*, the French legislature has just taken further steps to make data more broadly available by going beyond public institutions. The Law adopts the concept of 'data in the general interest' to expand the open-data policy to private entities such as public service concession holders or entities that receive state

preservation of scientific information, C(2012) 4890 final.

172 This is also the policy applied by the EU within its Horizon 2020 research funding programme. See European Commission, 'H2020 Programme—Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020', Version 3.1 (25 August 2016), available at: <http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf> (accessed 10 September 2016).

173 Sec 48(3) German Copyright Act (entry into effect on 1 January 2014).

174 See Art 17 *Loi pour une République numérique* (*supra* n. 166); see also comments on Art 9 in the English Explanatory Memorandum (*supra* n. 166). The French provision however provides for an embargo period of 24 months, instead of 12 months, for publications in the human and social sciences.

175 In this context, see also Art 11(2a) of the PSI Directive (*supra* n 21). As regards public-private partnerships of cultural institutions with private entities for the digitisation of cultural resources, this provision limits the grant of an exclusive license for the re-use of the digitised version to 10 years.

176 Similarly, the un-waivable secondary publication right does not prevent the author from granting an exclusive licence covering the publication right to the publisher.

177 *UsedSoft* (*supra* n 35).

178 PSI Directive (*supra* n 21).

subsidies.[179] In the first case, the concession holder is under an obligation to provide all data collected in the framework of the concession to the public authority in a digital format. In the second case, the recipient of the subsidy is under an obligation to provide all essential data as stipulated by the grant agreement in a digital, reusable and exploitable format to the authority.

171 In all of these instances, the state appears either as the source, or as an intermediary for making data available to the public. However, the more difficult question is whether such access rights can also be devised with regard to fully independent private data holders. In this instance, for any access regime, a fundamental distinction could be made according to the features of the business model the data holder applies. In the first case, the creation of a dataset is only a by-product, and the commercialisation of the data in downstream data markets is not part of the main business of that entity. This is the case, for example, where a car manufacturer collects geographic data through the cars' sensors for the purpose of predictive maintenance, but other firms or the state would be interested in getting access to that data. In such cases, the private entity may anyhow be willing to grant access in order to generate additional income, but the parties may still be unable to agree on access due to information problems. Intervention in the form of access regimes that provide for a framework of negotiations, mediation and arbitration will not reduce in any way the data holder's incentives to generate the data.

172 The situation is however very different in the second case, where the collection of the data constitutes a key element of the business model in competition with other firms. Examples are in particular the business models of search engines or social platforms, such as Facebook, which build on the control of user data to compete more effectively in the market for online advertising. Access regimes should not facilitate access of weaker competitors to data where control over such data constitutes the most valuable asset for competition.

173 The same argument applies to the tools for collecting and processing information, in particular as regards big data analytics, since these tools are of crucial importance for the commercial success of big data analysts. However, where such tools become the standard for collecting and processing information, as explained above,[180] access regimes may be justifiable also from the perspective of sound competition policy.

179 Arts 10 and 11 *Loi pour une République numérique* (*supra* n 166).

180 At F.II.3 above.

3. The person seeking access and the intended use of the data

174 In particular, access to data is justifiable where public entities seek access for the fulfilment of tasks in the public interest. In the light of the large benefits deriving from big data analytics, which could help optimise public policies and decisions of the state in many regards, this sub-category for which access regimes could be implemented seems most important.[181] Such regimes could be implemented at the different levels of government through sector-specific regulation. Sector-specific regulation appears as the road to take, since the security interests of the state will most likely need different rules than the prevention of infectious diseases, the protection of the environment or the functioning of smart cities or traffic control systems.

175 As explained above,[182] this is a field in which the competition rules on refusal to deal will hardly be able to promote access.

176 Going yet a step further, access based on public interest does not have to be limited to public entities as petitioners of access. An example of an access regime in the public interest providing for access to data in favour of even competitors is provided by the REACH Regulation.[183]

177 This Regulation has the objective of ensuring 'a high level of protection of human health and the environment, including the promotion of alternative methods for assessment of hazards of substances, as well as the free circulation of substances on the internal market (...)'.[184] To enable the assessment of these hazards, the Regulation's registration provisions require manufacturers and importers to generate data on the substances they manufacture or import. To meet these obligations the manufacturers and importers have to submit a dossier that contains the relevant information to the European Chemicals Agency (ECHA). Registered substances are allowed to circulate within the internal market.[185]

181 See in this context in particular the study of OECD (*supra* n 5).

182 At F.II.2 above.

183 Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC, [2007] OJ L 304/1; consolidated version available at: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02006R1907-20150601&from=EN> (accessed 10 September 2016).

184 Art 1(1) REACH Regulation.

185 Recital 19 REACH Regulation.

178 Such assessment of hazards may also require the manufacturers or importers to conduct new tests.[186] Tests may include animal testing.[187] But the REACH Regulation tries to avoid testing with vertebrate animals by recourse to alternative test methods wherever possible.[188] As part of the regulatory framework for preparing and submitting a registration, Articles 27 and 30 REACH Regulation implement a scheme for information sharing that pursues the particular objective of avoiding animal testing.[189] More concretely, the potential registrant is under an obligation to request a sharing of information from previous registrants as holders of studies, whether these studies include tests with vertebrate animals or not. Thereby, the Regulation also takes into account the interest of the previous registrant in fair compensation for the testing it has already undertaken.[190] For that latter purpose, the owner of the existing study has to determine the costs of sharing the information in a 'fair, transparent and non-discriminatory way'.[191] Under this scheme, the parties are expected to enter into an information-sharing agreement.[192] In case such an agreement cannot be reached, the REACH Regulation provides for default rules. The potential registrant can inform the ECHA about the failure to reach an agreement.[193] Then, within one month, the ECHA gives the potential registrant permission to refer to the information requested in its dossier, provided that it has paid the previous registrant a share of the cost incurred. At the same time, the Regulation confirms the right of the previous registrant to claim a proportionate share of the cost. This amounts to an equal share of the cost if the previous registrant makes the full study report available to the potential registrant. This right of equal cost sharing is enforceable before the national courts.[194]

179 In sum, the REACH Regulation builds on particular features that could be used as guidance for similar legislation in other fields. First, a duty to share information is formulated against the backdrop of a particular public interest in avoiding the duplication of the generation of information. In this context, it is important to remember that, in contrast, the rules on refusal to deal under EU competition law following the CJEU's *Bronner* judgment do not exempt the petitioner from making the same investment as the holder of the essential facility.[195] Hence, the REACH Regulation facilitates access to information beyond the remedies available under competition law. Second, the subject-matter of access consists in identifiable information similar to the competition law cases in *Magill* or *Microsoft*. However, it is to be discussed whether this model could also be applied to cases where somebody seeks access to large datasets for the purpose of undertaking big data analyses or engaging in data mining. It seems that, to the extent that there is a particular public interest in obtaining access, such broader access regimes are also justifiable. Third, the REACH Regulation relies on a framework of contractual negotiations. It thereby favours a pro-market solution over direct government intervention. The detailed rules of the REACH Regulation are very context-specific; but the negotiation framework could be adapted to other sector-specific circumstances. Fourth, the data-sharing agreement also requires agreement on the price or compensation to be paid for the sharing of information. The REACH Regulation thereby relies on concepts that resemble the FRAND concept as used in particular by standard-setting organisations in their IP policies concerning SEPs.[196] However, the REACH Regulation is more concrete about the base for calculating the compensation, relying on the cost for undertaking the relevant study.[197] Fifth, a negotiation-based access regime will only work where the law offers a default rule that enables the public interest to prevail and that provides sufficient legal certainty for the parties when they assess whether it makes sense to depart from that rule. This default rule also has to include procedures of judicial enforcement through state courts or arbitration tribunals in case no agreement can be reached.[198]

186 Recital 26 REACH Regulation.

187 Such testing has to be conducted in conformity with Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes, [1986] OJ L 358/1.

188 Recital 47 REACH Regulation.

189 See also Recital 49 REACH Regulation.

190 Recital 50 and 51 REACH Regulation.

191 Arts 27(3) and 30(1)(2) REACH Regulation.

192 More concrete rules on the standards of negotiations are contained in the Commission Implementing Regulation (EU) 2016/9 of 5 January 2016 on joint submission of data and data-sharing in accordance with Regulation (EC) 1907/2006 of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), [2016] OJ L 3/41.

193 Art 27(5) REACH Regulation.

194 Arts 27(6) and 30(3) REACH Regulation.

195 See at F.II.2. above.

196 FRAND licensing is considered as a general solution to overcome barriers to entry by Rubinfeld and Gal (*supra* n 28) at 37.

197 In contrast, R&D costs are not an appropriate standard for calculating the value of a patent. There is agreement to the extent that the royalty base should relate to price of the product in which the technology is implemented. However, there is disagreement as to whether the royalty should be calculated as a percentage of the often very complex end product, or as a percentage of the smallest salable unit.

198 Note that the default rule is very weak in the case of SEPs for which the patent holder has committed to FRAND licensing. The problem here is that the default rule is not based on statutory rules but private ordering through the IP polices

180 The question may still be whether and to what extent an access regime like the one contained in the REACH Regulation could also be implemented for cases in which there is no additional public interest. Indeed, such an access regime would make sense if it is devised as a non-mandatory procedural framework for negotiations on access to information. For designing such a general framework, it would be wise to assess the effectiveness of models such as the REACH Regulation or the most recent experience with the negotiation framework devised by the CJEU in *Huawei* for the case of SEPs. Such schemes could especially be applicable for cases in which the holder of information publicly commits to grant access to data on FRAND terms. It is yet another question whether such a scheme should be implemented by the EU or national legislatures, or by way of private ordering, in particular through industry associations that provide for commercial arbitration. The European Union could cooperate with the latter institutions to promote such non-mandatory arbitration on access disputes.

G. Conclusion

181 This article shows that existing EU rules, as regards both protection of data holders and access to data based on EU competition law, are applicable in principle to the data economy. However, in particular the rules of the Database Directive, the brand-new Trade Secrets Directive, and EU competition law, present considerable uncertainties as regards their application to the data economy. These uncertainties cannot be expected to be clarified quickly by the European Courts.

182 Yet, although the Trade Secrets Directive was not drafted to meet the needs of the data economy, trade secrets protection can provide a sound approach to protecting firms in the data economy to some extent. Rather than recognising exclusive control over any use of protected information, as would be typical for intellectual property regimes, EU trade secrets law implements a tort law approach that bans specific conduct related to the acquisition, dissemination and use of trade secrets that can be considered as unfair. It is thereby better suited to balance the interest in protection and in free flow of information than the property approach.

183 While a clarification of the scope of trade secrets protection regarding data as it is collected and used in the data economy would certainly be welcomed, the analysis shows that there is no case for creating a new system of data ownership. Apart from the fact that the key issues to be addressed—namely, regarding the subject-matter of protection, the identity of the data owner, and the scope of protection—are of enormous complexity, the analysis does not produce any evidence for a need or an economic justification for such legislation. In principle, in the data economy, no incentives are needed for generating and commercialising data. Data holders are able to charge a price for making data available to third parties based on factual control over data, supported by technical protection measures.

184 Hence, the question remains as to whether there is a need for legislation on access. In principle, the legislature could also promote access through un-waivable exceptions and limitation as part of a comprehensive legislation of data ownership. However, this article favours stand-alone access regimes. This latter approach better suits the dynamic development of the data economy, which most likely will only gradually inform the legislature about impediments to access while business models develop. In contrast, immediate adoption of an integrated ownership system would result in general recognition of exclusive control, whereas unfounded trust in adequate operation of a fair-use provision or postponing legislation on targeted exceptions and limitations would fail to address the additional limitations on the free flow of information generated by new data ownership.

185 In principle, access can also be sought under EU competition law. However, this law shows considerable shortcomings as regards the data economy: first, the requirement of market dominance in Article 102 TFEU considerably limits the scope of application of this rule and requires an often burdensome assessment. Second, it is quite uncertain to what extent Article 102 TFEU can be applied in cases in which, as will be frequently be the case, the data holder is not competing with potential customers in downstream data-related markets. Of course, Article 102 TFEU can also be relied upon to remedy excessive pricing. However, competition law enforcers can hardly be expected to act as price regulators in the data economy, which is characterised by information problems and huge uncertainties regarding the value of data. This puts the state as a frequent end user of data services in a particularly uncomfortable situation. Where the state has to rely on access to privately held data and big data analyses to optimise its decisions for fulfilling tasks in the public interest, it does not act as an undertaking in the sense of competition law and,

of standard-setting organisations. To bring more precision to the concept of what FRAND actually means may raise competition concerns in the sense of an anti-competitive price agreement. Hence, the default rule is ultimately in need of judicial interpretation of the FRAND concept by courts. Hence, FRAND licensing of SEPs does not provide a perfect model for regimes to enhance access to data.

hence, the rules on refusals to deal based on theories of exclusion and leveraging of market dominance by vertically integrated firms will not apply from the outset.

186 Yet the state, including the legislature, could promote access to data in a pro-active and pro-competitive way. Where different stake-holders contribute to the generation of data and information and only some of these contributors are interested in promoting access, the legislature can decide to particularly vest these persons with rights to enforce access against the interests of the other stakeholders. Examples of this are the secondary publication right of authors of scientific publications and data portability rights. The latter can enhance competition where factual control of other parties creates a lock-in effect. Block exemption regulations can take care of conflicts over access to data between suppliers and end producers. The state can promote access as part of its funding policy and even when granting subsidies. More importantly, there is a case for implementing sector-specific access regimes in the public interest. While it is hard to conceive a general legal framework for access of the state to data in the public interest, progressive sector-specific legislation in diverse fields of law, including environmental law, public health law, medicinal law or road traffic law, can develop models for access regimes over time.

187 Public-interest considerations can also play a role where private parties seek access to information. European competition law sets a rather high threshold for a duty of a dominant firm to share an essential resource by requiring the person seeking access to make at least the same investment in duplicating the resource that was made by the holder of the facility. There is a case for access regimes below this threshold where additional public interests, such as in the case of producing data through animal testing or clinical trials with human beings, or the interest in promoting scientific research, argues against duplication of already available data.

188 A main barrier of access is uncertainty about the information contained in large datasets, the new information that can be drawn from existing data through data mining and big data analytics and, hence, the value of data and the appropriate price to be paid for access. The so-called information paradox makes it particularly difficult to agree on the price of access to information in contractual negotiations. Access regimes should address this issue by favouring a consensus-based approach to regulating prices. Where pubic interest or competition law justifies access, a cost-based approach to assessing the royalty rates seems most appropriate.

189 As regards access negotiations between private parties, the Commission could support schemes of private ordering that enable private initiatives to pool data of multiple data holders.[199] The Commission could also cooperate with institutions that have experience with arbitration to build up schemes for mediating negotiations on data licensing.

190 The functioning of the data economy will also depend on the interoperability of digital formats and the tools of data collecting and processing.[200] The relevant tools have to rely on interoperability and, hence, the markets for such tools will typically be characterised by network effects. In this regard, the Commission can cooperate and support industry initiatives for standardisation of these tools, whereby those initiatives should also develop disciplines that promote access to the standardised tools. Accordingly, these needs of the data economy should also be taken into account as part of the Commission's competition policy regarding standardisation agreements. The Guidelines on Horizontal Co-operation Agreements already recognise the principle that standard-setting organisations should require participants to commit to license their IP rights in the standard on FRAND terms in order to make the standard broadly accessible.[201] This approach is superior to *de facto* standardisation, not only because it will enhance quick and general data sharing based on interoperability of data across borders and across sectors,[202] but also in the light of the fact that EU competition law has so far not developed appropriate disciplines through its case-law on refusals to license regarding the access problems arising from *de facto* standards.

199 This also has a competition law connotation, as demonstrated by the rules on information sharing in the Communication from the Commission—Guidelines on the application of Article 101 of the Treaty on the Functioning of the European Union to horizontal co-operation agreements, [2011] OJ C 11/1, paras 55-110.

200 See the standardisation issues regarding data and big data analysis mentioned in Communication from the Commission—ICT Standardisation Priorities for the Digital Single Market (19 April 2016), COM(2016) 176 final, p. 9.

201 Horizontal Cooperation Guidelines (*supra* n 199) para 285.

202 Commission Communication on ICT Standardisation Priorities (*supra* n 200) at 9.