

MAGKS



**Joint Discussion Paper
Series in Economics**

by the Universities of
**Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen**

ISSN 1867-3678

No. 35-2017

Wolfgang Kerber

**Rights on Data: The EU Communication ‘Building a
European Data Economy’
From an Economic Perspective**

This paper can be downloaded from
<http://www.uni-marburg.de/fb02/makro/forschung/magkspapers>

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

Rights on Data: The EU Communication ‘Building a European Data Economy’ From an Economic Perspective

Wolfgang Kerber

Abstract

In its Communication "Building a European data economy" the EU Commission discusses the introduction of a new exclusive property right on data ("data producer right") for non-personal (or anonymised) machine-generated data, and mandatory access rights to privately held data for achieving more access, transfer and reuse of data, esp. in the context of "Internet of Things" applications. This article analyzes the problem of "rights on data" from an economic perspective (incentive problem, data markets, bargaining power problems, access problems in multi-stakeholder situations) and the reasonings and proposals in the Communication from an economic perspective. Important results are that a "data producer right" cannot be recommended but that access rights to data can be part of specifically tailored data governance solutions in certain sectors.

Key words: Big Data, machine-generated data, data ownership, data access, data markets, internet of things

JEL classification: K24, L86, O34

I. Introduction

In January 2017 the EU Commission published a Communication ‘Building a European data economy’¹ and an accompanying Staff Working Document (‘Data Economy SWD’)². This is part of a broader policy approach dealing with the general question how the legal and regulatory framework of markets in the EU have to be adapted to the new conditions of the digital economy, in which data are seen as the new critical resource for innovation and welfare.³ Although the Communication deals with several issues (as free flow of data within the EU, liabil-

¹ European Commission, ‘Building a European Data Economy’ COM(2017) 9 final (hereinafter ‘Data Economy Communication’).

² European Commission, ‘Staff Working Document to the Communication “Building a European data economy”’ SWD(2017) 2 final (hereinafter ‘Data Economy SWD’).

³ See the Communications from the European Commission, ‘Towards a thriving data-driven economy’ COM(2014) 442 final and ‘A Digital Single Market Strategy for Europe’ COM(2015) 192 final.

ity, portability, interoperability and standards), the main issue is ‘data access and transfer’, because the Commission sees the problem that many of the produced data in the digital economy are not shared and (re)used enough. In that respect, the Communication focuses not on personal data that are subject to European data protection law but on non-personal machine-generated data (especially raw data, e.g. through sensors), which will play an increasing role in smart manufacturing and the Internet of Things (IoT). Since property rights do not protect many of these data, this is directly linked to the recent academic discussion about proposals for the creation of a new sui-generis intellectual property rights (IPRs) on (non-personal) machine-generated (industrial) data.⁴ Starting with the diagnosis that most of these data are only used in-house and not shared with and reused by other firms, the Communication has suggested a number of policy proposals, which encompass, inter alia, the introduction of an exclusive ‘data producer right’ but also mandatory access rights to privately held data.⁵ In that respect, the Commission has triggered a new discussion about ‘rights on data’.

This article has the objective to analyze core questions, reasonings, and policy proposals in the Communication from an economic perspective.⁶ A particular focus will be on the question of the necessity of defining new rights on data, especially in regard to the suggested data producer right as an exclusive property right but also in regard to access rights to data. To some extent this contribution builds upon a previous article published prior to the Communication,⁷ in which I analyzed from an economic perspective whether we need a new IPR on non-personal data in the digital economy. In this former article I came to the conclusion that based upon our current theoretical and empirical knowledge there is no need for introducing a new IPR for non-personal data, and that such a new exclusive right might even have serious nega-

⁴ See for this discussion; Dorner, ‘Big Data und “Dateneigentum”, Grundfragen des modernen Daten- und Informationshandels’ (2014) CR 617; Drexl, ‘Designing competitive markets for industrial data: Between propertisation and access’ (2016) Max Planck Institute for Innovation and Competition Research Paper No. 16-13; Kerber, ‘A new (intellectual) property right for non-personal data? An economic analysis’ (2016) GRURInt 989; Wiebe, ‘Protection of industrial data – a new property right for the digital economy?’ (2016) GRUR Int. 877; Zech, *Information als Schutzgegenstand* (Mohr Siebeck 2012) 412–440; Zech, ‘Daten als Wirtschaftsgut – Überlegungen zu einem “Recht des Datenerzeugers”’ (2015) CR 137; Zech, ‘Data as a tradeable commodity’ in De Franceschi (ed.), *European Contract Law and the Digital Single Market* (Intersentia 2016) 51; Zech, ‘A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data’ (2016) JIPLP 460–470.

⁵ A public consultation on this Communication ended on 26 April 2017. The results of this consultation were not available at the date of the finishing of this article.

⁶ For assessments of the Communication from a legal perspective, see, e.g., Drexl, ‘Neue Regeln für die Europäische Datenwirtschaft? Ein Plädoyer für einen wettbewerbspolitischen Ansatz’ (2017) NZKart 339 (Part 1) and 415 (Part 2); Wiebe, ‘Von Datenrechten zu Datenzugang – Ein rechtlicher Rahmen für die europäische Datenwirtschaft’ (2017) CR 87; Zech, ‘Building a European data economy – The proposal of the European Commission for a data producer’s right’ (2017) ZGE (forthcoming); additionally a number of documents as part of the consultation process of the Commission can be found on the Internet.

⁷ Kerber (n 4). Other important contributions from an economic perspective are OECD, ‘Data-driven innovation: Big data for growth and well-being’ (OECD Publishing 2015) available under <http://www.oecd.org/sti/data-driven-innovation-9789264229358-en.htm> last accessed 25 August 2017; OECD, ‘Maximizing the economic value of data: Understanding the benefits and challenges of enhanced data access’ DSTI/CDEP(2016)4; Duch-Brown/Martens/Mueller-Langer, ‘The economics of owner, access and trade in digital data’ (2017) JRC Digital Economy Working Paper 2017-01.

tive effects on legal certainty, competition and innovation in the digital economy. However, the more general issue of ‘rights on data’ is extremely complex and so far not well-understood. This article intends to extend the analysis from an economic perspective in two ways: (1) It will analyze and assess the reasonings in the Communication about the problems of a data economy and the suggested policy proposals of the Commission. (2) It will also try to clarify and extend the discussion about data ownership, rights on data, and data governance, and give some hints for the direction of necessary research. The policy questions raised by the Commission are economic policy questions. From that perspective the question about the need for new property and/or access rights in regard to data always has to start with questions about the existence of market failures and what kind of legal rules and rights can help to remedy these market failures and promote innovation in the digital economy. From a legal perspective this is a functional approach that uses economics as an analytical tool for a better understanding of the effects of different sets of legal rules and rights in regard to data on competition, innovation, and welfare.

The article is structured as follows: Section II. gives a brief overview about the main reasonings, objectives, and policy proposals in the Communication about the data access problem (including an explanation of the suggested ‘data producer right’). Section III. focuses on the reasonings about the basic economic characteristics of data (non-rivalry and excludability) and discusses the reasons for the lack of an incentive problem for the production of data. The problem of potential market failures for trading data and the respective policy proposals in the Communication will be dealt with in Section IV. Since the Commission has large concerns about contractual arrangements in regard to data due to ‘unequal bargaining power’ situations, Section V. will try to clarify this issue and whether a data producer right can help to address those concerns. Section VI. deals with the proposals to solve access problems through the introduction of access rights to privately held data. Especially in multi-stakeholder situations those access rights might play an important role within a broader data governance solution. In Section VII. the conclusion is drawn that a data producer right cannot be recommended but that carefully specified access rights can be important elements within comprehensive sector-specific data governance solutions, for whose design however much more specific research is necessary.

II. Data Access and Data Producer Right in the Communication: An Overview

The main concern of the Commission in the Communication is that the benefits of a thriving data economy cannot be fully realized because most holders of machine-generated data do not make their data accessible but are using them only for themselves (in-house). Empirically, only a limited amount of data-sharing and reuse of data can be observed, and data marketplaces are only slowly emerging and not widely used.⁸ Giving further access to these data for reuse, also in other economic contexts, would lead to more benefits and innovation in the

⁸ Data Economy Communication (n 1) 8, and, in more detail, Data Economy SWD (n 2) 12–19.

digital economy. Therefore the main objective is to increase access and sharing of machine-generated data.⁹ However, the Commission also sees the problem that the manufacturer of a device bought by a user is often in *de facto* control of the data produced by this device, and prevents the user ‘from authorising usage of the data by another party’¹⁰. Therefore the objectives that should be achieved by ‘a future EU framework for data access’¹¹ are much broader. The Communication enumerates five objectives that from an economic perspective can be boiled down to three different main objectives:¹²

- (1) Facilitate access to and sharing of machine-generated data;
- (2) Protect investments, assets, and confidential data, and ensure a fair return on investments that contribute to innovation; and
- (3) Ensure a fair sharing of benefits between data holders, processors and application providers within a value chain, and solve the problem of unequal bargaining power of companies and private individuals, e.g., also through lock-in situations, especially for SMEs and start-ups.

Whereas the first objective concerns the economic benefits of the reuse of non-rivalrous data, and the second objective is indirectly about incentive effects in regard to investing in data and innovation, the third objective seems to focus on distributional questions about fairness in regard to the sharing of the benefits of the data, especially through the problem of unequal bargaining power on markets, both in B2B and B2C contexts.

Based upon these diagnoses and objectives the Communication discusses a number of specific policy options. A closer analysis of the six suggested measures leads to three main policy strategies:¹³

- (1) A first group can be interpreted as measures that facilitate the sharing and trading of data via contracts (guidance about how data control rights are addressed in contracts, fostering technical solutions for reliable identification and exchange of data, and default rules that can serve as benchmarks for contracts related to data).¹⁴ However, they intend not only to reduce transaction costs but also want to deal with fairness and imbalances in bargaining positions (e.g., through an unfairness control also in B2B contractual relationships).¹⁵

⁹ Data Economy Communication (n 1) 8. The Communication addresses this problem both for non-personal data and for anonymized personal data at 9.

¹⁰ *ibid* 10.

¹¹ *ibid* 11.

¹² *ibid* 11–12.

¹³ *ibid* 12–13.

¹⁴ On the topics of contractual concepts see the contributions by Drexler, Janal, Graf von Westphalen, and Dolžan, in this volume.

¹⁵ *ibid* 12.

- (2) A second group of measures intends to solve access problems by introducing mandatory access rights to privately held machine-generated data. Such access rights can refer to public authorities acting in the ‘general interest’ and for scientific purposes, but the Commission also suggests the option of mandatory access to privately held data for other private parties (based, e.g., on FRAND terms), which is *de facto* a ‘compulsory licensing’¹⁶ solution.¹⁷
- (3) An own kind of measure is the option of introducing a new exclusive ‘data producer right’ that is granted to the owner or long-term user of a (data-producing) device, and which gives this ‘data producer’ an exclusive right to use and authorize the use of these non-personal data. The main objective is to clarify the legal situation, give more choice to the data producer how to utilise their data, and contribute to unlocking machine-generated data.¹⁸

Therefore the suggested measures intend to make more data available, on one hand, through reducing transaction costs for trading and licensing of data (facilitative law), and, on the other hand, through mandatory regulations in regard to access rights to data but also by introducing new exclusive rights on machine-generated data.

Since the Commission is well aware of the potentially far-reaching implications of a ‘data producer right’, the Data Economy SWD discusses this issue in a very comprehensive way.¹⁹ The scope of data that should be covered are non-personal or anonymized machine-generated data. However, the information should only be protected on the syntactic level (data as ‘signs’) but not on the semantic level (meaning of data), i.e. data should only be protected as codified information. It is emphasized that it should not be conceived as a ‘super IP-right’.²⁰ A particular difficult problem is the question to whom such a new data producer right should be granted. The Commission discusses several criteria, emphasizing particularly who invests into the creation of data, mentioning both the manufacturer and the economic operators of a device. It is therefore surprising how clearly the Communication identifies the ‘owner or long-term user (i.e. the lessee) of the device’ as the data producer to whom the exclusive right should be granted.²¹ However, there also should be exceptions, particular in respect to ‘the

¹⁶ On the topic of compulsory licences see the contributions by Grützmacher, Mezzanotte, and Weber, in this volume.

¹⁷ *ibid* 12–13.

¹⁸ *ibid* 13.

¹⁹ See Data Economy SWD (n 2) 33–36.

²⁰ *ibid* 34; for the important distinction between the syntactic and semantic level of information, see Zech, ‘Rights to Use Data’ (n 4) 462–463.

²¹ Data Economy Communication (n 1) 13; from an economic perspective, it is not so clear who is investing in these data: If a manufacturer under competition can offer a lower price, because he can use the data from the device, then it is the manufacturer who invests in the data, and not the buyer/lessee of the device. For the difficult problem who should get the data producer right see Zech, ‘Daten als Wirtschaftsgut’ (n 4) 142–144, and Wiebe (n 4) 883; Bundesministerium für Verkehr und digitale Infrastruktur, “‘Eigentumsordnung for Mobilitätsdaten?’ Eine Studie aus technischer, ökonomischer und rechtlicher Perspektive’ (2017). In the Staff Working Document this problem is discussed in a much more open way, especially also the possibility that in many con-

provision of non-exclusive access to the data by the manufacturer or by public authorities’.²² Especially interesting is that the Data Economy SWD explicitly discusses whether a data producer right should be introduced as an absolute right (‘right in rem’) or alternatively only as a set of purely defensive rights that protect the *de facto* ‘possession’ of a data holder, for example against an illicit misappropriation of data in analogy to the protection of know-how as trade secrets. This option is not mentioned in the Communication but is close to the option favoured by many legal scholars not to introduce an exclusive property right but rely and perhaps improve the current legal situation, which already provides a number of (defensive) rights to the *de facto* data holder (primarily also via trade secret protection).^{23 24}

III. Economic Characteristics of Data: Non-rivalry, Excludability and (the Puzzle of) the Lacking Incentive Problem

What are the basic economic reasonings that are explicitly or implicitly used in the Communication in regard to data and the data economy? A crucial question refers to the economic characteristics of data. That data are non-rivalrous in use and therefore should be used as much as possible is one of the core assumptions in the Communication. This is based upon the economic arguments that the marginal costs of an additional use of (non-rivalrous) data is zero, and that the access to more data allows the innovation of new products and services (data-driven innovation).²⁵ From this perspective, the observation that many data that are collected by firms are not used or are only used in-house, and not made available to other firms, raises the question of a possible market failure in regard to the use and trade of these data. The Commission is therefore right in focusing on the problem of data access, data sharing and data transfer. However, before addressing these problems in the next section we also have to take a look at the problem of the production of these data.

The fact that data can be seen as a non-rivalrous intangible good has led the academic discussion to the question whether data should be protected in analogy to IP protection.²⁶ This follows because the non-rivalrous character of data does not support the analogy to physical property; it is the rivalry in use that provides the main rationale for the exclusive rights grant-

texts it might not be appropriate at all to grant a right to one particular firm or person (see Data Economy SWD (n 2) 35, with a reference to OECD, ‘Maximising the economic and social value of data’ (n 7) 25.

²² Data Economy Communication (n 1) 13.

²³ See, e.g., Dorner (n 4) 617; Drexler (n 4) 17–26.

²⁴ However the concept of a set of purely defensive rights for the protection of *de facto* ‘possession’ of data can be understood differently than in the Data Economy SWD (n 2) 33–34. Here the suggested civil law remedies focus on the possibilities to sue others for unauthorized use, which is close to the idea of protecting ‘exclusivity’. Instead, a set of defensive rights for the *de facto* data holder can be understood much more broadly as a set of rights against the destruction, compromising, obstructing the access to, and certain forms of misappropriation of data. See Kerber (n 4) 998.

²⁵ See generally OECD, ‘Data-driven innovation’ (n 7) and OECD, ‘Maximizing the economic value of data’ (n 7) 9, in which data is seen as a ‘(general-purpose) infrastructure’ for Big Data and the digital economy.

²⁶ See Zech, ‘Daten als Wirtschaftsgut’ (n 4) 137; Wiebe (n 4) 877.

ed by physical property. However, the main rationale for protecting innovation and creative works with exclusive IPRs has always been the necessity to solve an incentive problem for producing these intangible goods, which might be caused by too low appropriability of the benefits of the innovation or creative work through too fast imitation via knowledge-spillovers (or direct copying). From an economic perspective it is one of the surprising aspects in the discussion about property rights on data that nobody has claimed that there is a (general) problem of too low incentives for producing data, for instance due to too fast imitation or copying of data. It is interesting that empirically there does not seem to exist a serious problem of copying of data that resembles the well-known problems in regard to IP-protected digital goods. Although the Data Economy SWD briefly mentions investments in regard to data,²⁷ it never addresses the question whether the data economy suffers from too low incentives for producing data and whether therefore property rights on data are necessary to solve such an incentive problem. The Commission implicitly seems to assume – much in line with the academic discussion about data ownership – that there is no market failure in regard to the incentives for producing data, and that therefore the problem is primarily one of access to and sharing of the already existing data. If, however, it is true that there is no incentive problem for producing data, then the main economic rationale for an exclusive IPR for a non-rivalrous intangible good does not work in the case of data, and therefore also the ‘data producer right’ cannot be justified on these grounds.²⁸

I would also agree with the widespread belief that the empirical evidence of an exponential increase of produced data, which also can be expected to continue and even accelerate in the future (especially through sensor data in the IoT),²⁹ does not support serious concerns that the data economy suffers from a general incentive problem for producing data. However, for a proper understanding of the data economy and of the effects of the suggested policy measures, it is important to clarify why such an incentive problem does not seem to exist. Data are produced if the benefits of the data for the data producer are larger than its costs. One possible explanation would be that the costs of producing data are zero or at least very low. This might be true if the data are a by-product of other processes. If there are large economies of scale and scope for data production, then at least the marginal costs of producing additional data might be close to zero. However, in most contexts, there are considerable fixed costs for data collection, for example the production and installation of sensors, or the offering of ‘free’ services in exchange for the possibility to collect data from users. In the recent discussion it also has been emphasized that there might be considerable costs for ‘curating’ data, i.e. costs for ensuring the quality of a data set, for documenting the provenance of data, and bringing it into a format that allows the use of these data (interoperability).³⁰ In regard to personal data, consent has to be given and properly documented, or there are costs for an effective any-

²⁷ Investment in the collecting of data can be found only in the Data Economy SWD (n 2) 22 (in connection with the lack of legal protection for such investments) and 35 (as criterion for allocating the data producer right).

²⁸ See in more detail Kerber (n 4) 993.

²⁹ See, e.g., OECD, ‘Data-driven innovation’ (n 7) 133–143.

³⁰ See for these costs *ibid* 191–198, and OECD, ‘Maximizing the economic value of data’ (n 7) 21.

mization of personal data. Consequently, it cannot be assumed that the costs of the production of data sets are generally zero or at least very small. It can rather be expected that the costs of data production might be very different, depending on the kind of data, and the economic and technological context, in which they are produced. These costs of the production of data also can justify the concerns in the discussion about the protection of investments in data.³¹

If data production is not costless, the lack of a general incentive problem implies that the data producer can reap enough benefits from these data for covering the costs – through using the data herself and/or letting others use these data through data sharing, data trading and data licensing. An important precondition for ensuring the possibility of revenues is the second economic characteristic of data (besides non-rivalry), namely that the data holder is usually capable to exclude others from using her data (excludability), especially through keeping them secret (e.g. by encryption). Empirically, it does not seem to be a problem to keep data secret if they are stored and used in-house.³² The data holder can then commercialize her data, for instance by using them for innovating or improving own products, or for selling additional services (as, e.g., predictive maintenance and data analytics) to their customers. Another important option is that the produced data are only used in-house for developing and offering data-based services to other firms (e.g. targeted advertising). Thus not making data directly available to other firms does not allow the conclusion that the information in these data is not made available to other firms in the data economy. Through the targeted advertising services of Google and Facebook, many advertisers can use the information in the huge data sets of these firms without getting (and needing) direct access to these data. As such, we also should be cautious not to draw too fast conclusions about insufficient use of the data of firms, if they keep them to themselves and make them not directly accessible.

The other possibility for appropriating the benefits of produced data is the direct selling (or allowing the use) of data to other firms. This links the incentive problem of data production to the problem of trading data. If data markets suffer from severe market failures, then the smaller revenues from selling or licensing data might lead to too small incentives for the production of data, and therefore to an under-production of data. One of the potential market failure problems of selling or granting access to data (licensing) is that the buyers (or licensees) should not be able to resell the data to others. This raises again the question whether the excludability of data can be ensured, i.e. that the data holder can control the use of data if she grants access to the data to other firms (via data trading). Both technical and contractual restrictions can be used for solving this problem. As far as these restrictions work effectively, the data holder can ensure the *de facto* control of her data even in those cases, in which she grants other firms direct access to these data. In the Section IV., we will see that there is so far no empirical evidence that technical and contractual restrictions cannot work well enough for ensuring the *de facto* control (and therefore excludability) in the case of trading data. All of this is supported by the empirical fact that we do not observe a serious problem of copying data.

³¹ If there are no costs for data, then there would be no investment into data that has to be protected.

³² However, there are certainly problems of ‘hacking’ and cybersecurity.

What conclusions follow from this analysis? The lack of a general incentive problem despite the existence of costs for the production of data implies that the *de facto* ‘holdership’ of data (through ‘*de facto* excludability’) seems to allow the firms to get enough benefits from their data without the need of the protection through an exclusive property right. However, this result also implies that – due to the costs of data production – such a *de facto* control of privately held data might be necessary for maintaining the incentives for data production. This is an important insight for assessing policy proposals about mandatory access to these data.

However, a deeper analysis into the economics of data would show that the conditions for the costs and benefits of data production can be very different in different economic and technological contexts. Additionally, there also might be considerable positive and negative externalities in regard to the costs and the benefits of the production of data, which lead to differences between, on the one hand, the private and social costs of data production, and, on the other hand, the private and social value (benefits) of data. Depending on these circumstances, the incentives for data production can either be too high or too low. Consequently, our empirical belief that there might be no general under-production of data through too low incentives does not allow the conclusion that the incentives are always optimal and that there might not exist many kinds of data, for which the incentives are either too low or too high, leading to an under- or overproduction of certain kinds of data. As such we also cannot exclude that there are no market failures due to incentive problems in regard to the production of data.³³ Neither the Communication ‘Building a European Data Economy’ nor the accompanying SWD raise the question whether for certain kinds of data there might be incentive problems for the production of data, and whether this might be a problem for a thriving data economy. This is an issue that clearly requires much more additional research.

IV. Market Failures in Regard to Trading of Data? An Open Question

In the academic discussion the most important argument for introducing an exclusive property right on machine-generated non-personal data is the enhancement of their tradability on markets. We know from the discussion on traditional IPRs that patents and copyrights can help to reduce transaction costs for selling and licensing innovations and creative works. Consequently, such a transaction cost reasoning might also provide a rationale for a ‘data producer right’.³⁴ For answering this question we have to analyze from an economic perspective whether markets for the trading and ‘licensing’ of data suffer from market failures that impede the efficiency of these markets. The empirical observation that these markets – despite rapid expansion – are still underdeveloped and fragmented raises concerns in that respect. From that

³³ Therefore, it also might be possible that for certain kinds of data an exclusive property right can be justified with the incentive argument (e.g., if there is a serious copying problem in regard to this specific kind of data). See for a deeper analysis of possibilities for market failures in regard to data (e.g., also fragmentation of data ownership) Duch-Brown/Martens/Mueller-Langer (n 7) 29–32.

³⁴ See Zech, ‘Data as a tradeable commodity’ (n 4) 77.

perspective the authors of the Data Economy SWD were right by starting with a survey of the current knowledge about sharing of data and data markets. However, the problem is that despite some first empirical studies about data trading and data marketplaces, which give some hints in regard to problems of trading data, so far we do not have a comprehensive analysis of the transaction problems on data markets. It is thus not surprising that in the Data Economy SWD only a description of the current state of data markets and the extent of data sharing and data trading can be found, but not an analysis of what the real transaction problems and therefore market failure problems on data markets are.³⁵ Such an analysis, however, is necessary for developing policy measures that can make data markets more efficient.

What do we know about the problems of data markets? First, trading of data can emerge in very different forms, for example via data brokers, data market places, and with many different business models and contractual arrangements. In their theoretical study about data markets Koutroumpis et al. distinguished four types of trading of data: ‘one-to-one’ (bilateral trading), ‘one-to-many’ (dispersal marketplaces), ‘many-to-one’ (harvesting of data, e.g. selling data for ‘free’ services to a single buyer), and ‘many-to-many’ (multilateral marketplaces).³⁶ The revenues from trading data are increasing rapidly, but neither in the US nor the EU well-established (multilateral) data market places exist, on which buyers and sellers of data can directly trade with each other.³⁷ This fits to the insight of Koutroumpis et al. that especially ‘many-to-many’ marketplaces seem to have considerable transaction problems.³⁸ The few empirical studies that exist so far for the EU do not support the thesis that legal problems of data ownership or excludability are among the main problems for data markets. The problems that are mentioned in these studies are insufficient demand for data (due to not enough awareness of the value of data), interoperability problems and lacking standards, problems in regard to the pricing of data, the lack of open platforms for data-sharing, and an insufficient number of skilled data workers.³⁹ The issues of data quality and data provenance have emerged in the most recent discussion. The value of data for buyers critically depends also on the quality and the provenance of data. This encompasses clear information about the type of data, its correctness and completeness, whether it has been collected legally (according to privacy laws or properly anonymized), and the context where and how (i.e., with what kinds of methods) they have been produced etc.⁴⁰ If information about these aspects are not disclosed (or are not veri-

³⁵ Data Economy SWD (n 2) 12–19.

³⁶ See Koutroumpis/Leiponen/Thomas, ‘The (unfulfilled) potential of data market places’ (mimeo, 15 May 2017) 19; they also provide a broad overview about the current literature.

³⁷ See OECD, *Data-Driven innovation*, 2015, 44; Koutroumpis/Leiponen/Thomas *ibid* 3.

³⁸ See the analysis in Koutroumpis/Leiponen/Thomas *ibid*.

³⁹ See Carnelley et al., ‘Europe’s Data Market Places. Current Status and Future Perspectives (European Data Market SMART 2013/0063 D 3.9)’ (2016) 26–27; Cattaneo et al., ‘D8 Second Interim Report (European Data Market SMART 2013/0063)’ (2016) 52–58; OECD, *Data driven Innovation*, 2015, 95–97.

⁴⁰ See Koutroumpis/Leiponen/Thomas (n 34) 11–18 (with more literature); based upon interviews with data scientists and other experts Mattioli even views these problems of data quality and provenance as so serious that he suggests to think about granting an IP-like data right for incentivizing disclosure of all relevant information about the provenance and quality of data, see Mattioli, ‘Disclosing Big Data’ (2014) *Minnesota Law Review* 535.

fiable), then data markets can suffer the typical market failure problems due to information asymmetries (e.g., adverse selection). Whether reputation of the sellers or the data marketplace (if they provide a verification service) can solve these problems, is an open question.

Although a mosaic-like picture of empirical and theoretical insights about trading of data is slowly emerging, a comprehensive analysis of market failure problems is still missing.⁴¹ For our issue of the necessity of an exclusive data producer right, the decisive question is whether these sellers or the data marketplaces can ensure, for example, through monitoring or data traceability, that buyers do not breach their data licensing contracts by leaking the data to third parties. If this turns out to be a serious problem, which cannot be solved, for example, by technical restrictions and traceability of data, then an exclusive property right with its possibility to sue such third parties might emerge again as an interesting policy option. However, it is not clear how the suggested ‘data producer right’ in the Communication would help to solve the other mentioned transaction problems, for example, problems of data quality.⁴²

However, all these questions need much more theoretical and empirical research. This also makes it difficult to assess the other policy proposals about facilitating trading of data in the Communication. Rather unproblematic and potentially helpful might be facilitative measures, as giving guidance how non-personal data control rights can be addressed in contracts, or the development of new default contract rules, because both offer the chance for helping to reduce transaction costs. However, without a clear analysis of the transaction problems it is an open question how such guidance and default contract rules should look like, and whether the Commission and/or legislators would really be capable of offering better solutions than what firms and lawyers (as transaction cost engineers) can achieve in a step-by-step experimental process in the market through developing standard solutions for contracts in regard to data. The brief explanations in the Communication suggest that the guidance and the default contract rules seem to have more the objective of solving fairness problems than focusing really on transaction costs.⁴³ Much more interesting and an important contribution to the entire discussion are the suggestions to foster the development of technical solutions for reliable identification and exchange of data. One of the huge (and not sufficiently discussed) problems of the digital economy is the verifiability of the performance of contracts about data. How can a firm that sells data by granting others access to these data monitor whether the buyer of access really uses the data only in that way that has been specified in the contract, and, especially does not grant other firms access to these data? If compliance with the contract cannot be monitored and verified, then the necessary trust for trading data is lacking. Therefore, tech-

⁴¹ See also Duch-Brown/Martens/Mueller-Langer (n 7) 36–41 (with additional literature).

⁴² In that respect, neither the Communication nor the Data Economy SWD offer a clear reasoning. The only explicit argument is that such a right can reduce legal uncertainty by clarifying the legal situation. However the controversial discussion about the specification and allocation of such a data producer right (see Wiebe (n 4) 883) renders it very unclear whether the introduction of such a right would on balance really reduce legal uncertainties and therefore transaction costs compared to the current situation. See for the new dangers of blocking innovation through such a right also Kerber (n 4).

⁴³ See Data Economy Communication (n 1) 12.

nical solutions for ‘traceability and clear identification of data sources’⁴⁴ (e.g. through ‘watermarking’) might be very important preconditions for enabling effective markets for trading and sharing of data.⁴⁵ This is also linked to questions of interoperability and standardization, which definitely can facilitate the trading of data, and the discussion about Application Programming Interfaces (APIs).⁴⁶ Overall, these policy options might help to solve transaction problems and reduce transaction costs, but clear policy recommendations can only be given after a careful analysis of the real transaction problems, which also might be very different for different kinds of data. Then also entirely new policy options might emerge and become relevant.

V. Distributional and Fairness Problems in Regard to Data?

One important concern in the Communication that is mentioned repeatedly is that market solutions might suffer from ‘unequal bargaining power’ between firms, on the one hand, and private individuals or other firms (especially SMEs), on the other hand.⁴⁷ The Commission emphasized that solutions for facilitating the sharing of machine-generated data should take into account possible differences in bargaining power between the market players, and that a ‘fair sharing of benefits between data holders, processors and application providers within value chains’⁴⁸ should be ensured. In regard to policy solutions the introduction of an unfairness control in B2B contractual relationships is suggested, combined with default contract rules that can serve as a benchmark for balanced solutions for contracts related to data. Although it might be possible to link this issue also to the objective of increasing the sharing and trading of data, in the Communication this issue seems to be seen as an own independent problem. It also clearly relates to similar concerns in the academic literature that private parties and SMEs might not get a fair share (or remuneration) of the benefits of ‘their’ data through the contractual arrangements in the market.⁴⁹ Additionally, it relates to the more general question of who should get the benefits of the data that are collected/produced in certain situations.

The problem is that neither the Communication nor the Data Economy SWD provide a clear analysis of such ‘unequal bargaining power’ situations.⁵⁰ This is a complex topic for a single article. However, distinguishing between three different groups of problems might help to structure the discussion. The first group comprises situations in which one party is weaker

⁴⁴ Data Economy Communication (n 1) 12

⁴⁵ See especially also the discussion of new concepts of decentralized multilateral data market places with ‘distributed ledger technologies’ and the potential role of ‘blockchain technology’ in Kouroumpis/Leiponen/Thomas (n 34) 31–42 (with additional literature).

⁴⁶ See Data Economy Communication (n 1) 16.

⁴⁷ *ibid* 10–12.

⁴⁸ *ibid* 11.

⁴⁹ See, e.g., Zech, ‘Daten als Wirtschaftsgut’ (n 4) 145.

⁵⁰ In fact, the issue of ‘unequal bargaining power’ is not analyzed at all in the section about data access and transfer in the Data Economy SWD.

than the other due to a clearly definable market failure problem. These can be cases in which firms have market power in the way it is defined in competition law. Then ‘unequal bargaining power’ situations can emerge through a lack of competition. In these cases, firms with market power might be capable of concluding contracts about data with private persons or also other firms, especially also SMEs, which lead to ‘unfair’ results in regard to the distribution of the *de facto* control of data or remuneration for data.⁵¹ Other reasons for the relative weakness of a party in a bargaining situation are information asymmetries and behavioural biases (especially in regard to private persons). This can lead to the well-known market failures through information and behavioural problems that are typically dealt with by consumer law (and in the case of personal data, by data protection law, also). For these groups of cases, well-established legal rules (in competition law, consumer law, data protection law etc.) exist for dealing with these problems.

It might be that these legal rules are not regarded as capable of solving all of these problems. For example, the requirements in regard to market dominance according to Art. 102 TFEU can be very high. Policy-makers might look therefore also for legal rules that can deal with weaker forms of market power. In German competition law, § 20 GWB also allows for controlling market power in bilateral relationships between firms. The Commission and also much of the legal literature considers applying the unfairness control of the Unfair Terms Directive⁵² also to B2B-contracts for remedying ‘unfair’ results in contractual arrangements about data.⁵³ However, we should be sceptical to what extent this solution can really help. The main problem is the lack of clear criteria for its application. This might either lead to legal uncertainty and the danger of too much intervention in the markets and/or to false optimism that these policy options (especially extending unfairness control to B2B-contracts) can solve the existing problems. Independent from this proposal there seems to be a broader consensus that the introduction of an exclusive property right on data (as the ‘data producer right’) cannot solve ‘unequal bargaining power’ problems through those market failures because those rights can be easily contracted away in situations with market power or information (or behavioural) problems of one party.⁵⁴ Therefore, this group of ‘unequal bargaining power’ problems should be primarily solved by the regular policy solutions for these market failures, i.e. competition, consumer, and data protection law.

A second group of problems are situations in which specific data are produced or should be used simultaneously by several often clearly definable stakeholders. These can be called multi-stakeholder situations. Such situations emerge, for instance, in the context of smart manu-

⁵¹ This can be exploitative abuse according to Art. 102 TFEU.

⁵² Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts [1993] OJ L95/29.

⁵³ See Data Economy Communication (n 1) 12. See also the contributions by Dolžan, and Graf von Westphalen, in this volume.

⁵⁴ See Drexl (n 6) 342; Kerber (n 4) 996. This important insight can also be found explicitly in the Data Economy SWD (n 2) 36.

facturing, smart agriculture, smart energy, smart health, and also smart cars.⁵⁵ In all of these situations several parties often either contribute jointly to the production of data and/or need the same data for improving their products (and services) and for innovation. The decisive point is that in such complex multi-stakeholder situations, an ‘ownership’ solution (in which only one party has the sole control of these data) might not be an efficient solution. It is important that, in such respect, it does not make a difference whether this exclusive ownership of one party is based upon the ‘*de facto* excludability’ of privately held data or upon an explicit exclusive property right (as, e.g., the data producer right). The problem of granting one private party in a complex multi-stakeholder situation *de facto* or legally an exclusive position in regard to specific data is that this monopolistic control gives them a strong bargaining power in regard to the other stakeholders. Such a position can be used for demanding high ‘licensing’ fees or other advantages from the other stakeholders that can be seen as ‘unfair’ (data hold-up situations). We will discuss this more deeply in the next section in connection with access rights and the need for developing more sophisticated data governance solutions. But one important conclusion can already be drawn here: granting an exclusive data producer right to one party instead of sticking to the current *de facto* exclusivity of the data holder would not solve this problem.

However, there might also be a third group of problems, which I only can hint at here. This one deals with the more general question of who should get the benefits of data from a distributional perspective independently from the first two groups of problems, and how the structure of rights on data might influence the distribution of the huge potential welfare gains that accrue from using the many newly produced data in the digital economy. In the discussion about data ownership this also might be linked to as ‘unfair’ perceived results of markets in regard to who is getting (the benefits of) data. The question whether and under what conditions the granting of property-like positions for data might influence the distribution of the benefits of data has to be analyzed very carefully from an economic perspective. This is also especially interesting for personal data, where EU data protection law grants rights that resemble to some extent exclusive rights.

VI. Data Access Problems and Access Rights to Data

The Communication is correct in its analysis that there might be a problem that privately held data are not made sufficiently accessible to further use. Facilitating trade and licensing of data is one possible solution. The other group of solutions consists of direct measures for making data available to other parties, for instance through introducing mandatory access solutions (or promoting open data solutions). For a deeper analysis, it might be helpful to distinguish two different groups of cases why the exclusive ‘*de facto* control’ of data might be a problem:

⁵⁵ These are the typically discussed ‘Internet of Things’ examples; see, e.g., Data Economy SWD (n 2) 25–28.

(1) The former is based upon the general argument that it is not efficient if non-rivalrous data are not used enough, especially in entirely different economic contexts (Big Data and cross-sector business models). In this case, both facilitating trading (and licensing) and general mandatory access solutions, for example in the public interest (public authorities, public statistics, scientific research) but also for other private firms, can be interesting policy options.

(2) The latter group of cases are the aforementioned multi-stakeholder cases, in which either several parties contribute jointly to the production of data and/or several parties have large economic interests for having access to the same data. In the last section, we already argued that a sole ownership of rights on data (either through *de facto* control or through an explicit property right) might not be an optimal ‘rights on data’ solution due to the danger that this ‘sole owner’ can exploit this monopolistic situation by ‘holding up’ the other stakeholders.⁵⁶ This can not only lead to the above-mentioned ‘fairness’ problems through ‘unequal bargaining power’ situations but also to high transaction costs (through conflicts between the parties) and to reluctance to invest into the cooperation between these firms in value chains and value networks. In that respect, it can also have negative effects on economic efficiency and innovation in the digital economy.

In regard to these multi-stakeholder situations let us take a brief look at two examples. In smart agriculture, sensor data from agricultural machines are an important input both for the farmers and for the manufacturers of machines (and perhaps other specialized service providers). As it is important for the farmers to have access to the sensor data of all their machines from different manufacturers (for improving their agricultural production), it is also important for the manufacturers to have access to the data from all the machines they have sold to different farmers (for improving their machines).⁵⁷ In other contexts, the multi-stakeholder situation might be much more complicated. In the example of connected cars there is a number of stakeholders who have very different interests in the large set of different kinds of data that are produced in connected cars. Important stakeholders are the car owner, the car driver (and other passengers), the car manufacturer, car component suppliers, insurance companies, independent repair services and spare part producers, public authorities (for traffic regulation and road maintenance), and different kinds of firms that would like to offer services (e.g., media content etc.) to the car passengers. In addition, here (as in many future IoT-applications) many of these data are also personal data granting individual persons specific rights according to data protection law.⁵⁸ Giving one party in those multi-stakeholder situations, legally or *de facto*, the sole exclusive right to all of these data might not be the optimal solution, also from an economic perspective. It is therefore not surprising that much more complex solutions about who should have what kinds of rights in regard to these data are discussed, especially in

⁵⁶ Such hold-up problems can be theoretically directly linked to the well-known problem of ‘opportunistic behaviour’ in ‘Hold up’-situations with transaction-specific investments in the governance approach of Williamson, ‘Transaction-cost economics: The governance of contractual relations’ (1979) *Journal of Law & Economics* 233.

⁵⁷ See for smart agriculture Data Economy SWD (n 2) 28.

⁵⁸ See for smart (connected) cars Hornung/Goebel, “Data Ownership” in vernetzten Automobil’ (2015) CR 265–273; C-ITS Platform, ‘Final report’ (January 2016).

the complex case of connected cars. But similar problems also emerge in the case of smart home devices (e.g. in regard to energy efficiency) or health data.

One important solution for dealing with these and similar multi-stakeholder problems is the granting of (non-exclusive) access rights to specific stakeholders. The Commission Communication also suggests a combination of the data producer right with certain non-exclusive access rights (as limitations) to solve problems that emerge if one party is granted an exclusive data producer right.⁵⁹ However, the decisive point is that we do not need an exclusive data producer right for such solutions. It is also possible to define directly access rights for specific stakeholders in such multi-stakeholder contexts that can be enforced against the *de facto* holder of certain data. This is the situation we already have for the special case of independent repair services in regard to access to repair and maintenance information in the automotive industry.⁶⁰ It is also the solution that has been suggested by the Max-Planck-Institute for Innovation and Competition as an alternative to the introduction of a data producer right.⁶¹ The authors correctly argue that the solution to such problems does not lie in the data producer right but in the proper specification of access rights to data for specific stakeholders.⁶² However, from an economic perspective, the technological and economic conditions in different multi-stakeholder contexts as health data, agricultural data, smart home data, energy data etc., are very different, partly also through the very different relevance of data protection in regard to personal data. As such, the question whether specific access rights for certain stakeholders are helpful and can be recommended, requires in-depth legal and economic analyses of these different kinds of situations. Optimal governance solutions for such multi-stakeholder situations might be thus very complex and sector-specific.

However, we also should be cautious not to be too hasty in recommending the legal introduction of specific mandatory access rights in such multi-stakeholder situations. From an economic perspective, firms on competitive markets have incentives to offer their customers and business partners contracts with data governance solutions that solve these problems, for example by offering voluntarily access to the data to the most important stakeholders. Therefore, it is also necessary to show why we cannot expect that market competition leads to appropriate solutions. We also have to look for a market failure problem in such multi-stakeholder situations before we intervene into the market through mandating specific access rights. In

⁵⁹ Data Economy Communication (n 1) 13.

⁶⁰ See Arts 6 and 7 Regulation (EC) No 715/2007 of the European Parliament and of the Council of 20 June 2007 on type approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6) and on access to vehicle repair and maintenance information [2007] OJ L171/1.

⁶¹ See Max Planck Institute for Innovation and Competition, 'Position Statement of Max Planck Institute for Innovation and Competition of 26 April 2017 on the European Commission's Public Consultation on Building the European Data Economy' accessible online under <http://www.ip.mpg.de/fileadmin/ipmpg/content/stellungnahmen/MPI_Statement_Public_consultation_on_Building_the_EU_Data_Eco_28042017.pdf> and, in more detail, Drexl (n 6) 342; see for a deeper discussion of non-exclusive (property) rights on data, and the question of the tradability of these access rights Zech (n 6).

⁶² In the case where several firms have contributed to the production of data, it is also possible to use governance solutions, in which one firm has *de facto* control and the other contributing firms have access rights to these data instead of forms of joint or collective property rights on data.

that respect, the farmer / machine manufacturer example might differ significantly from the example of connected cars because the complexity might be much larger in the latter than in the former case. It can be assumed that an increasing complexity with regard to the number and kind of stakeholders and the interests in using the data might increase the probability of market failures in those multi-stakeholder situations (due to large transaction costs). This, however, is also a topic for future research.

The issue of access rights finally also leads back to our discussion about the incentive problem in regard to data production. If we accept the conclusion in Section III. that our lack of a general incentive problem depends on the possibilities to use and commercialize the data (and is not a consequence of a costless data production), all solutions that focus on granting access to privately held data have to balance the (manifold) benefits of access with possible new incentive problems that might arise if access rights are granted. This is relevant both for solutions that give generally access to certain data (e.g., via public statistics or general mandatory access solutions) and for specific data governance solutions in multi-stakeholder contexts. The (widely neglected) incentive problem for data production thus emerges again in the discussion of access rights as policy solutions. This does not mean that it is always important that a data holder should be well remunerated for the data she controls, if she has to give access to other parties. In regard to traditional IPRs, both economists and lawyers agree that innovators should get enough benefits from their innovations to have sufficient incentives for innovating but they also should not have higher incentives. The same logic also applies to the production of data. If the costs of data production are very low or even zero, granting access to such data might not be a problem, even in the case of no or low remuneration. However, if we think that for certain kinds of data the production costs are high, we should be much more reluctant in granting mandatory access to such data (or require a high remuneration).⁶³

VII. Tailoring Specific Data Governance Solutions Instead of General Property and Access Solutions for Data

With the Communication ‘Building a European Data Economy’ the EU Commission has triggered a very important discussion about the necessary legal and regulatory framework for the huge amount of machine-generated non-personal (and anonymized personal) data that will be produced in the context of ‘smart manufacturing’ (Industry 4.0), connected cars and other devices of the ‘Internet of Things’. In particular, the issue of ‘data ownership’ is an extremely complex problem leading to a lot of puzzling questions. It is the combination of several characteristics of data (data as codified information, non-rivalry in use, excludability with the lack of a general incentive problem, and data as input for Big Data and innovation) that makes data a very special economic good, which does not fit well into traditional economic and legal categories. Trading and sharing of data on markets also seem to be a complex issue whose problems we have not understood well. From that perspective, the Communication is surprisingly

⁶³ For this balancing between the benefits of giving more access and the necessity to protect incentives for data production see from an economic perspective also Duch-Brown/Martens/Mueller-Langer (n 7) 25–27.

successful in its attempts to explore these issues, the potential problems of the data economy, and what kind of policy solutions might help. With its main focus on data access and data sharing, the Commission has certainly chosen a perspective that fits well into the basic functional logic of the digital economy. However, the main problem is that we are still at the beginning of theoretical and empirical research about important characteristics and problems of the data economy, and therefore we should be very cautious about conclusions whether and what type of market failures exists, and what policy measures might be suitable for solving them. Overall, it is even still not clear to what extent we have a problem at all, or whether the markets in the digital economy work well enough also under the current legal and regulatory framework with regard to non-personal data. It should be kept in mind that there is no comparable policy discussion in the US.

The most controversial proposal in the Communication is the introduction of an exclusive ‘data producer right’. Considering our current knowledge, this proposal cannot be supported from an economic perspective. Since data is both a non-rivalrous and excludable good (with no serious copy problem), no exclusive property rights are necessary for ensuring the incentives for data production. So far we also have no evidence that the transaction problems on data markets are caused by lacking property rights. Further, neither the problems of ‘unequal bargaining power’ with regard to contracts relating to data nor access problems in multi-stakeholder situations are solved by the data producer right. If we consider, additionally, the many costs and dangers of introducing an exclusive data producer right for legal certainty, competition and innovation, especially in the digital economy, then the introduction of a data producer right cannot be recommended.⁶⁴ This does not preclude that further research might show that for specific kinds of data, where, for instance, we have well-confirmed evidence of an otherwise unsolvable copy and therefore an incentive problem, such an exclusive property right might be part of a suitable policy solution. In regard to the group of proposals about access rights to privately held data, the non-rivalrous character of data and the crucial importance of access to data for the digital economy seem to favour access rights from an economic perspective. Whereas access solutions in certain cases of public interest and for scientific research are comparably unproblematic, we should be very cautious about general mandatory access regulations to privately held data for other private parties, even if, on first sight, this might have positive effects on competition and innovation. Very careful analysis is needed of the kind of data and the kind circumstances in which such obligations might be appropriate, especially from the perspective of incentives for the production of these data.

This can be different in specific sectors or situations in which several or many stakeholders are in complex relationships with regard to personal and non-personal data. Many of the currently discussed examples (connected cars, smart energy, smart home, smart agriculture, and smart manufacturing) can fit into this category. Here, it can be presumed that we might need specifically tailored complex sets of legal rules and rights in regard to data. Such ‘data governance regimes’ can be seen as the entire set of legal rules and rights that refer to the collec-

⁶⁴ The problems, dangers, and costs of a data producer right have not been discussed in this article. See Kerber (n 4) 996; especially important are the dangers for legal certainty and the public domain (see Wiebe (n 4) 883).

tion/production, processing, storing, using/analyzing, protection, access to and trade/sharing of data in a certain sector. Since the benefits and costs for the stakeholders and the economic and technological conditions (as well as certain normative issues as privacy concerns) might be very different in these different sectors, we would expect from an economic perspective that also the optimal data governance solutions will differ significantly. However, economists would also expect that in many sectors the markets themselves can find the proper data governance solutions and establish them through contractual arrangements with the different kinds of stakeholders. Yet in complex situations (as connected cars) or in the case of serious other market failures, it might be necessary to establish appropriate data governance solutions also through regulatory or legislative measures. From that perspective, economic arguments can support the already wide-spread opinion in the legal discussion that – as a next step – we should focus primarily on the analysis of sector-specific problems in regard to data and develop properly tailored specific data governance solutions.⁶⁵ If we have gained enough knowledge and experiences with sector-specific analyses of data governance problems, it might be possible to develop more general solutions. Such a strategy is certainly in line with the basic ideas and objectives of the Communication ‘Building a European data economy’.

⁶⁵ See, e.g., Drexl (n 6) 419.