Aug 10th, 12:00 AM

From Data Exchanges to Data Markets: An Institutional Perspective

Kai Reimers
RWTH Aachen University, reimers@wi.rwth-aachen.de

Xunhua Guo
Tsinghua University, guoxh@sem.tsinghua.edu.cn

From Data Exchanges to Data Markets: An Institutional Perspective

Completed Research Full Paper

Kai Reimers
RWTH Aachen University
reimers@wi.rwth-aachen.de

Xunhua Guo
Tsinghua University
guoxh@sem.tsinghua.edu.cn

"If it were asked–What discovery has most deeply affected the fortunes of the human race? We think after full consideration it might probably be said–The discovery that a Debt is a Saleable Commodity." (MacLeod 1875, p. 172)

Abstract

It becomes increasingly clear that the objectives of privacy and competition policy are in conflict with one another with regard to platform data. While privacy policies aim at limiting the use of platform data for purposes other than those for which the data were collected in order to protect the privacy of platform users, competition policy aims at making such data widely available in order to curb the power of platforms. We find that current control- and ownership-based approaches are ineffective with regard to their capacity to balance these conflicting objectives and propose an institutional approach which makes platform data saleable. We discuss this approach in view of its capacity to balance the conflicting objectives of privacy and competition policy and with regard to its effectiveness in supporting each separately. Our approach also clarifies the fundamental difference between data markets and other concepts such as data spaces.

Keywords

Data Markets, General Data Protection Regulation, Privacy, Anti-trust, Platforms, Property Rights

Introduction

How to enable the emergence of data ecosystems instead of continuing to foster the rise of digital superstars? This question lies at the heart of the challenges that arise with the expected digital transformation of almost all sectors of the economy. This transformation is characterized by the emergence of platforms that aim to dominate whole industries, often created by outsiders with an intention to 'disrupt' extant forms of organizing economic activities (Chen 2021). While, from a policy perspective, such disruptions are generally welcomed since they are seen as instances of radical innovations, the accompanying rise to dominance of platforms is increasingly a cause for concern (Denemark 2022). In the race to disrupt industry after industry, prospective platform players often straddle, and sometimes cross, legal boundaries which has earned them the nickname of 'guerrilla capitalists' (Chan and Kwok 2021). However, the biggest concern is not with their seeming willingness to operate in a legal grey area, but the prospect of their entrenchment in the industries they aim to disrupt. A monopoly position would be seen as acceptable as long as it can be challenged by newcomers, i.e., as long as it is contestable (European Commission 2020a). But the intense competitive manoeuvers of would-be platform players in the quest for dominance are supposedly made in view of the prospect of high entry barriers which would effectively shield winners in that race from future competition and which, at the same time, are the cause for increasing anti-trust worries and regulatory action.

Network effects are often seen as constituting such entry barriers. However, empirical evidence points to the possibility of effective competition under conditions of network effects where promoters of new

technologies can successfully oust presently dominant platform players (Cennamo 2018). In fact, such a possibility is predicted when both extant and expected network effects are considered since, under suitable conditions, expected network effects can create a self-fulfilling prophecy, described as excess momentum to contrast with the notion of excess inertia that is traditionally seen to underlie entrenchment of monopolistic positions based on network effects (Krugman 1991). Hence, extant network effects may not be as effective entry barriers as is often assumed. Gregory et al. (2021, 2022) have identified a 'data network effect' which points to the potentially crucial role of data collected on platforms for erecting entry barriers. Indeed, recent efforts by the European Commission to curb the dominance of platforms focus on data, not on network effects. The main objective is to prevent platform operators from unfairly benefiting from the data collected on their platforms by limiting what they can do with such data and by enforcing access to such data by other companies through the so-called Digital Markets Act (DMA, European Commission 2020a).

However, in one recent assessment of the DMA, while praising the benefits of access it enforces to platform data for other companies, concerns have been raised whether such regulation might not severely impact incentives of platform operators to innovate by collecting and analyzing such data (Expertenkommission 2023). But how else could these regulatory objectives be reached without diluting the incentives of platforms for data-driven innovation? It seems to us that a promising candidate for such an alternative approach would be to allow platform operators to sell their data to other companies. While the creation of data markets has been recommended for the purpose of fostering innovative use of such data, especially with regard to the application of machine learning techniques (Guse et al. 2022), such efforts have not yet been linked to the regulatory objective of curbing the power of platforms (Abbas et al. 2021). Even the Data Governance Act (DGA), another regulatory proposal by the European Commission published in the same year as the DMA, which is often portrayed as providing the legal basis for the creation of data markets (e.g. Expertenkommission 2023), is not linked to the threat of monopolization of industries through platforms but to the aim of fostering the development of machine learning technologies (European Commission 2020b).[1]

One likely reason for the absence of such a discussion is that the notion of selling platform data seems to be in strong conflict with accepted principles of privacy and data protection. However, our point of departure in this paper is that extant privacy principles and practices—not network effects—hinder the emergence of data ecosystems and lay the foundation for the monopoly power of platforms because they severely restrict the sharing of platform data so that only the platforms themselves can benefit from their full potential (Höppner and Westhoff 2021; Jurcys et al. 2021, Huq 2021). The question then arises how the conflicting objectives of competition and privacy policy can be balanced if allowing platforms to sell the data collected on their platform is seen as an effective means to share such data without diluting incentives to collect them? Currently, both data markets and privacy rules are framed in terms of physical control of platform use data: A data market is understood as a place where data are physically exchanged (Koutroumpis et al. 2020) and privacy rules are based on regulating / limiting physical access to data. Hence, balancing the requirements of competition law and privacy law entails the negotiation and implementation of prohibitively complex configurations of access rights and data partitioning (Lubyová 2022; Jurcys et al. 2021). We therefore in this paper ask: How can data markets be conceptualized so as to simultaneously consider the conflicting demands of competition and privacy policy?

We draw on Commons' (1990) Institutional Economics to argue that answering our research question requires a conceptual move from understanding data markets as the site of the physical exchange of data towards an institutional understanding of such markets where ownership of data is exchanged, not the data themselves. Moreover, we follow Commons' approach to defining ownership in terms of liberty (of the owner) and exposure (of all others) which allows us to encapsulate the trade-off between the conflicting objectives of sharing and protecting data in a single relation: ownership. We do not address the question of how markets for data should be designed in order to become commercially successful (on this question, cf. the systematic literature review by Abbas et al. 2021) or the question of how property rights

[1] It is interesting to note that the DGA is often seen as facilitating the sharing of data between companies when, in fact, it refers only to the sharing of data collected by public bodies; see, e.g., Expertenkommission 2023 and the corresponding German Wikipedia entry.

to data can be monitored and enforced (on this question, cf. the ideas of Koutroumpis et al. 2020 regarding the use of technologies such as distributed ledgers to improve the provenance of data).

We begin by defining our notion of platform use data and distinguish this from the notion of personal data as defined in European law. In the main section, we then first show that present concepts of data markets and platform regulation are indeed based on the notion of physical control over such data. Second, we show how the move from data holdership to data ownership can be made based on Commons' Institutional Economics. Third, we address the question of who should be assigned ownership rights. In the discussion section we then show how our approach can be used to balance the conflicting requirements of data sharing and data protection and why platforms might actually be willing to sell their data.

What are Platform Use Data?

Talk about the creation of data markets or data infrastructures often appears as if it does not matter what kinds of data are to be traded or exchanged (Abbas et al. 2021). But, with equal justification, one may talk about creating 'goods markets' or 'services markets'. While all these categories—data, goods, services—are distinct from one another and thus have a specific meaning, the physical and institutional characteristics of different kinds of goods and services are so variable that finer categories are needed in almost all cases for specifying the institutional characteristics of markets for goods and services. In particular, while it is certainly true that data are different from goods in that the costs of copying are negligible for data but not for goods, this characterization is not sufficient for deriving the institutional requirements of data markets. For example, markets for consumer profiles certainly need to meet very different institutional requirements than markets for stock prices. While markets for consumer profiles must, above all, be concerned with privacy rules, the most important concern for markets for trading stock prices is their currentness (Koutroumpis 2020).

In this paper, we are concerned with the kind of data that need to be shared in order to curb the power of platforms. What kind of data are these? We submit that these are the data that are generated on the platforms themselves, since such data cannot be obtained other than from the platform operators. For example, connections between data objects are often created on a platform such as when a user 'puts' a 'product' into a 'shopping basket'. Here, data representing the user, an IP number or other identifying data, are connected with data representing a product. By contrast, many kinds of data collected by platforms can also be obtained from other sources. For example, data that connect a person with a bank account are typically not created on a platform but may have to be provided by a user in order to complete a transaction. Data generated on a platform therefore do not comprise data that already exist elsewhere but are now also provided to the platform. In short, we define platform use data as a record of connections between data objects when such connections are created through actions of users on a platform. Such records can be created, among other ways, through clicking on elements of a website / app; moving the cursor across a website / app; opening a website / app; entering search terms / questions and queries. This definition excludes, among others, the following kinds of data: address data, financial data, health data, and data secretly recorded by the device about the user.[2]

Our definition intersects with the notion of 'personal data' as defined by European Law in the General Data Protection Regulation (GDPR), which defines personal data as

> "… information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person …" (GDPR, Article 1, Point 1)

The sole criterion in this definition is whether data can be related to a natural person. Our definition of platform use data includes personal data when one of the data objects connected through such a record

[2] Thouvenin and Tamò-Larrieux (2021) have also called for a more differentiated approach to defining platform data. However, their proposal—to limit private property rights to those data that can be stored in a private cloud—is ultimately tautological.

refers to an identifiable natural person. It excludes personal data if connections between data objects contained in such records have been established elsewhere. And lastly, records may be created about connections between data objects where all links to personal data have been removed, e.g. when a connection has been established between two products by having both 'placed' in a 'shopping basket' without also recording the link to the user. The relationship between our definition of platform use data and that of personal data as given by the GDPR is depicted in Figure 1.

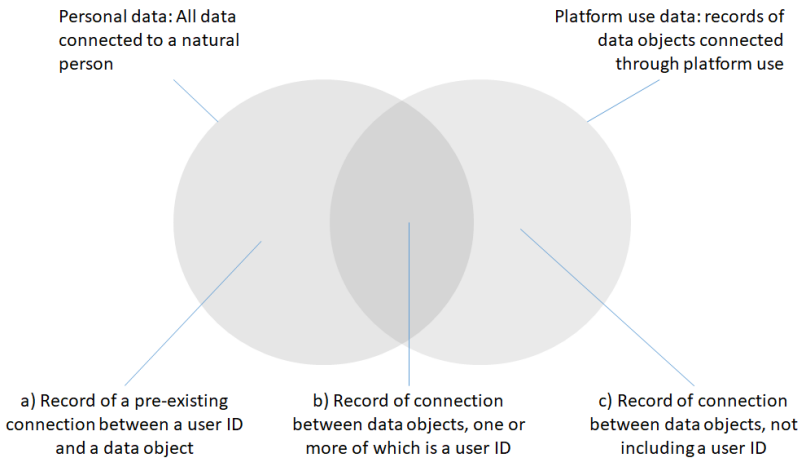

Figure 1: Relationship between the EU definition of personal data and our definition of platform use data

From Control to Ownership

Current efforts to curb the power of platforms, as they have become manifest in the Digital Markets Act (DMA), focus on regulating physical control of data by limiting what platform operators can do with the data they have collected and, conversely, by enforcing access to subsets of such data by other firms, including competitors. Likewise, regulating physical control of data is also the basis of efforts to protect consumers from harmful effects of such data, manifest in the General Data Protection Regulation (GDPR), and even of regulatory initiatives to create data exchanges, manifest in the Data Governance Act (DGA). We begin this section by briefly illustrating this claim—that all three regulatory initiatives are based on a preference for physically controlling what can be done with data—and then show how the notion of physical control differs from the institutional notion of ownership. We will also discuss the question of who should be assigned ownership of platform use data.

The control approach

The point of departure of the Digital Markets Act (DMA) is the idea that platforms act as gatekeepers between businesses and their customers. This is seen as a potential problem because platform operators may engage in the same commercial area as their business users, i.e. they may also act as their competitors. Since business users and their customers must, as part of doing business through the platform, provide data about themselves and their transactions, platform operators may use their privileged access to such data to compete with their business users, which is seen as unfair (European Commission 2020a, p. 15). To prevent such behavior, gatekeepers must not use data "generated" by business users or their customers on the platform to offer services which compete with those of their customers if such data are not publically available (Article 6a). Conversely, business users and their customers should be given effective and immediate (real time) access to data that they "generated" on the platform free of charge (Article 6h, i). Further versions of these general rules are given for more specific cases.

First, it is noteworthy that the DMA does not define the notion of user-generated data, even though the terms 'data' (any digital representation of acts, facts, and information), 'personal data' (by reference to the GDPR) and 'non-personal data' (all other data) are defined (Article 2, Points 19-21). Hence, it is not clear to which sub-set(s), as defined above (see Figure 1), these rules refer. It is possible to argue that all three subsets are, in some way, generated by users. The rules furthermore comprise positive and negative elements, i.e., they prescribe certain actions and prohibit others. The prescriptions specify the modalities of physical access to data. The prohibitions cannot likewise deny physical access to data because platform operators, by definition, have access to them. They therefore prohibit use of the data under certain conditions, which amounts to denying access to these data because, for the platform operator, the effect is the same. The DMA thus regulates access to data by opening and closing access routes.

The General Data Protection Regulation (GDPR) makes control of personal data by users its cornerstone: "Natural persons should have control of their own personal data" (The European Parliament and Council of the EU 2016, p. 2). To this effect, it focuses on regulating the modalities of "processing" personal data (Article 1, Point 1), where processing "… means any operation or set of operations which is performed on personal data or on sets of personal data …" (Article 4, Point 2). It is directed towards "controllers" defined as a "… natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data …" (Article 4, Point 7) and operates by restricting processing through the controller, where "… 'restriction of processing' means the marking of stored personal data with the aim of limiting their processing in the future …" (Article 4, Point 3). Hence, regulating physical access is, again, the dominant mode of operation. It is noteworthy that the GDPR informally alludes to ownership in the above cited phrase ("… their *own* personal data …"), which is seen to rest with the natural person about who some data are.

The Data Governance Act (DGA) aims to make "… public sector data available for re-use, in situations where such data is [sic!] subject to rights of others" (European Commission 2020b, p. 1). It does so through the creation of a "… harmonized framework for data exchanges" (ibid., p. 10). However, the term 'exchange' is not understood as a market but as a "… governance framework for data access and use, in particular regarding the re-use of certain types of data held by the public sector … (ibid., p. 11). The terms 'data holder', 'data user', 'data sharing', and 'access [to data]' are all defined in terms of physical control over data (Article 2, Points 5-8). By contrast, the notion of ownership of data is not defined but indirectly thematized in the sense that the regulation should not affect any extant intellectual property rights (European Commission 2020b, p. 14) while the granting of exclusive use rights to data is prohibited (Article 4, Point 1), unless that would serve a "general interest" (Article 4, Points 2 and 5). Thus, the aims of the regulation are again meant to be achieved through controlling physical access and institutional mechanisms, such as exclusive use rights, are expressly prohibited or limited.

The ownership approach

From an institutional perspective, for something to be tradeable it needs to be owned since selling and buying things means transferring ownership of something (Commons 1990). Data can, in principle, be owned (Jurcys et al. 2021; Thouvenin and Tamò-Larrieux 2021) and property in price data has been established more than a hundred years ago when another technological innovation disrupted then extant forms of doing business: the telegraph (Mulherin et al. 1991).[3] Data ownership of personal data by data subjects is also increasingly seen as a more effective means of control over personal data than oversight by public authorities in their stead (Huq 2021; Jurcys et al. 2021), although this position has also been challenged (see for a summary of that critique Thouvenin and Tamò-Larrieux 2021).

But what is ownership? Thouvenin and Tamò-Larrieux (2021) argue with regard to data ownership that there are two main legal understandings: ownership as private property and ownership as control. However, from an institutional perspective, ownership is property (legal control), not physical control. In fact, the main argument by Commons is that ownership and (physical) control have not been sufficiently separated by the classical and neo-classical economists and that clearly separating these concepts is the

[3] Huq (2021) quotes another 100 years old court ruling which suggests that ownership of data is not recognized by US courts. However, this ruling refers to ownership of news information and thus concerns a special case. The comparison with price data generated on a stock exchange seems to offer a much better model for platform use data.

hallmark of Institutional Economics. The most fundamental difference between these concepts is that physical control refers to the relationship between man and nature whereas ownership refers to the relationship between man and man. Ownership is often characterized as the right to exclude others from using something (e.g. Jurcys et al. 2021 with regard to data ownership). But Commons argues that, since ownership characterizes a relationship between man and man, it always implies two legal moments. To the right to exclude others from using something corresponds a no-right of all others (the whole world) to mobilize collective action—through a court decision—against the owner to do what he or she likes with what is owned (Commons 1990, p. 82). This two-way relationship creates two economic states: A state of exposure of the whole world to the consequences of the actions of the owner, which they have to tolerate, and a state of liberty on the part of the owner to do what he or she likes with what is owned. These two states of exposure and liberty are the legal basis for any market since they create the possibility of selling the right in question. The institutional approach does not ignore the aspect of physical control—the relationship between man and nature—but places it in the future: "Legal control is future physical control." (ibid., p. 87). Ownership, and the transfer of ownership through selling it, therefore always look to the future. A thing is evaluated on the basis of an expectation of the beneficial consequence—income, profit, pleasure—it will spawn in the future (ibid., p. 408).

The states of liberty and exposure created through ownership are not unlimited and a large part of the institutional work underpinning the functioning of markets consists of defining these limits and continuously adjusting them to changing circumstances. For example, in Germany it has been decided that the liberty of the owner of a forest is limited by the rule that he or she must not prevent the public from entering it because a forest also serves as a recreational area for city dwellers. Likewise, laws have been created at the state level in Germany that limit the liberty of owners of a house or apartment by the rule that they must not let it unoccupied for a longer period of time in order to avert speculation. In a similar manner, ownership of data can and should be limited, for example in order to protect data subjects from harmful effects of using such data. However, since the value of something owned consists of the expected benefit derived from it in the future, the tightening of these limits amounts to a taking of property in the sense that the value of what is owned will be diminished. A fundamental characteristic of the US-American institutional system therefore consists of the principle of due process: the definition of ownership must not be changed without due process of law (Commons 1990, p. 63).

Who should own platform use data?

When ownership has been proposed in the context of personal data, authors typically recommend that ownership should be assigned to data subjects. Two recent papers offer alternative reasons for this proposal. Huq (2021) argues that data subjects, by owning data about them, can control such data even when the data are physically held by others elsewhere, e.g. by platform operators on their servers. By contrast, Jurcys et al. (2021) argue that even when data are physically controlled by data subjects, e.g. in a so-called "private data cloud", assigning ownership to data subjects would strengthen their degree of control even further while also providing a more flexible approach to sharing such data with service providers who could "run" their services "on top" of personal data held in private data clouds (Jurcys et al. 2021).

However, these proposals do not aim to make personal data tradeable. In fact, the possibility that data subjects might sell their personal data is seen as a strong argument against assigning ownership rights to them because this would diminish the control rights that they currently have under modern data protection regimes (Thouvenin and Tamò-Larrieux 2021). By contrast, our main question is how platform use data can be made saleable. We therefore propose to assign ownership rights to platform operators, not data subject, because otherwise transaction costs for platform operators to acquire such data from data subjects would be prohibitive (ibid.), but limit this proposal to platform use data, i.e. to subsets b) and c) as shown in Figure 1, which excludes a large part of personal data as defined by the GDPR (subset a) in Figure 1). The rationale of our argument is that, in this way, platforms may have an effective incentive for sharing their data (below, we will discuss the strength of this incentive and also propose how the degree of exposure of data subjects created through such ownership could be limited). But apart from this functional legitimation, what moral basis could be quoted in support for this argument?

So far, two moral reasons have been given to support ownership by data subjects (see Jurcys et al. 2021). On the one hand, it has been argued that such data are created by the labor of users, manifest in the often

used term 'user-generated data', and therefore users have a right to the fruit of their work. This is a version of the labor theory-justification of private property originally proposed by John Locke (ibid.). On the other hand, it has been suggested that personal data are a virtual extension of the body and personality of the data subject and therefore belong to them to the extent that one has a natural property right to one's body and personality (ibid.). We propose that a moral basis for our proposal to endow platform operators with property rights to platform use data could be provided if the issue is broadened to include the ownership of platforms. We cannot elaborate this idea and instead refer to Mansell and Sison (2020) who argue, based on an analysis of the Medieval corporation, that corporations should be owned by their members on the grounds that they are all, to variable degrees, involved in producing the collective good that is the corporation's output. We think that such an argument is even stronger for the case of platforms whose value, as is widely acknowledged, is based on positive network externalities produced by their users.

Discussion

The institutional understanding of ownership, as laid out above, integrates two conflicting objectives: that of granting liberty to the owner to do with what is owned as she or he likes, including selling it, and that of limiting the exposure of all others to the consequences of such actions. Applied to the realm of platform use data, these objectives are the enabling of sharing platform use data more widely in order to curb the monopoly power of platform operators and the protection of platform users from the detrimental effects of data use by businesses. By defining the limits of both the liberty of owners and the exposure of all others these objectives can be traded off against each other. By contrast, the current control-based approach prevents any such trade-off by taking the protection of personal data as an absolute cornerstone of platform regulation so that any data sharing is possible only within the constraints of such protection (Thouvenin and Tamò-Larrieux 2021). Similarly, current property rights-based approaches to both data protection and data sharing define property right as the right to exclude others from using data. While this is certainly an important aspect, focusing on this aspect alone glosses over the possibility of tempering ownership rights which therefore appear to give absolute liberty to the owner and to leave all others completely exposed to the actions of owners. By conceptualizing ownership as a liberty / exposure-relation the possibility of limiting both, and thus of tempering the otherwise absolute right of exclusion, comes into view.

Both current control- and ownership-based proposals to platform use data are thus not effective with regard to reconciling these conflicting objectives. But, in our view, they are unlikely to even achieve each objective taken separately. The control-based approach gives rise to a highly complex web of contingent access rights which tends to cement the dominant position of platforms since only they have the resources to cut through this web of rules effectively (Jurcys et al. 2021; Huq 2021). Conversely, it does not prevent that platforms themselves use user data with detrimental effects for platform users (Huq 2021), mostly through forms of price discrimination, as has poignantly been shown by Clough and Wu (2022). Current ownership-based proposals, which assign ownership of platform use data to data subjects, also seem ineffective. On the one hand, the transaction costs of selling, and thus sharing, such data are prohibitive (Thouvenin and Tamò-Larrieux 2021). On the other hand, ownership of personal data by platform users does not prevent businesses from training their algorithms on other data and then using the results against those users who have not allowed businesses to use their data, for example for purposes of price discrimination, as has been argued by Jurcys et al. (2021).

By contrast, our proposal to assign ownership rights to platform use data to platform operators while limiting both liberty and exposure created through such property rights promises to be effective with regard to both objectives. On the one hand, any limits imposed on property rights to platform use data would apply to both businesses who buy such rights and platform operators who collect such data and possibly sell the rights to them. It would thus effectively protect users from harmful actions also of platform operators, for example by prohibiting to use platform use data for purposes of price discrimination. On the other hand, it would enable effective sharing of such data through selling them and thus contribute to curbing the monopoly power of platform operators.

However, would platform operators actually want to sell platform use data if they could? There are two reasons which support such an expectation. On the one hand, it is likely that there are firms which can make better use of platform use data than the platforms themselves, i.e. that there are benefits of

specialization. In an ingenious study, Agogo (2021) has empirically demonstrated widespread informal data sharing among websites which he convincingly explains by such specialization benefits. However, even if there are indeed such effects, platform operators would still have to trade off such benefits against monopoly rents lost through selling platform use data. Yet, they may very well decide in favor of selling platform use data. On the one hand, monopoly rents can only be maintained if all platforms refrain from selling platform use data because algorithms can be trained on one set of use data and then used against other users. This would require some form of informal cooperation and thus be highly risky in view of increasingly suspicious and alert anti-trust authorities. But by the same token, selling platform use data would be an effective means to alleviate any suspicion of anti-trust authorities that platforms are misusing their monopoly power.

The proposal to assign ownership of platform use data to platform operators seems, at first glance, counter-intuitive with regard to the objective of curbing platform monopoly power and absurd with regard to protecting platform users from detrimental effects of using such data. But, on closer look, these objectives can be effectively achieved through our proposed approach while this approach also allows for trading off these objectives against each other. A crucial step in seeing this possibility, however, consists of adopting a more nuanced approach to defining platform data. The concept of 'personal data' is far too broad for allowing effective regulation of platforms because it mixes highly sensitive data such as health and financial information—subset a) in Figure 1—with data that are less prone to be used against users while they also are of utmost importance for unleashing the expected benefits of applying new techniques such as machine learning in terms of improving resource allocation across almost all sectors of the economy, namely those which we have defined as subset b) in Figure 1.

Our proposal, however, is not meant to suggest that the definition of property rights to platform use data and their assignment to platform operators will suffice to curb the monopoly power of platforms and to protect platform users from harmful effects of the use of such data. For this to be the case, markets for platform use data must be successfully established and currently many obstacles remain in this regard, as shown by Abbas et al. (2021). Moreover, adequate market designs and governance structures have yet to be found. While some forms of data sharing services have been successfully established, so-called many-to-many markets with full control over ownership rights have yet to emerge (Koutroumpis et al. 2020). Rather, our aim was to show that, based on an institutional understanding of markets, it is possible to envision data markets that could be effective with regard to both curbing the monopoly power of platforms and adequately protecting platform users from the detrimental effects of the use of such data.

Conclusion

In this paper, we have addressed the question: How can data markets be conceptualized so as to simultaneously consider the conflicting demands of competition and privacy policy? We have shown that current control- and ownership-based approaches are insufficient in this regard. Control-based approaches tend to take the objective of data protection as absolute and then seek to enable data sharing, through defining further access rules, within the constraints of such protection. But even if data protection was seen as a variable objective that needs to be balanced with the conflicting objective of curbing the monopoly power of platforms, any effort of re-balancing would likely increase the complexity of the already complex rules governing access to platform data and thus ultimately, albeit unintendedly, strengthen the monopoly power of platforms because only they have the resources to cope with this complexity effectively. Current ownership-based approaches assign property rights to platform data to platform users, but not with an intention to make such data saleable but to increase factual control users have over such data, either as an alternative to physical control possibilities or as an extra layer of legal control to reinforce the physical control users should be given through technical means. In this way, assigning property rights to platform users is seen as a countervailing power to the monopoly power of platform operators.

Our proposal, counter-intuitively, is to assign property rights to platform data to platform operators in order to curb their monopoly power. However, for this to be seen as an effective measure, two conceptual moves are necessary. On the one hand, a more nuanced approach to defining platform data is required. The current approach, based on the concept of 'personal data', is far too broad because it mixes highly sensitive data, which should not be shared under any circumstance, with data the sharing of which is necessary in order to unleash the potential efficiency gains promised by the combination of big data and

machine learning techniques. On the other hand, the liberty created through assigning property rights needs to be tempered since to that liberty corresponds an exposure of 'the whole world' to the actions of the owner. Thus, property rights should not be seen as absolute either but as limited by ongoing institutional work which adjusts the liberty granted to owners and the exposure of all others to changing circumstances. One clear candidate for such limitation is a prohibition to use platform use data for purposes of price discrimination. Such a prohibition would apply to platform operators and other businesses alike and thus effectively protect platform users from one of the most detrimental and insidious effects of exploiting platform use data, a level of protection which current privacy regimes fail to offer. Moreover, our approach would effectively allow, based on a due process of law, for the continuous re-balancing of the conflicting objectives of data protection on the one hand and of curbing monopoly power on the other by adjusting the limits imposed on the liberty of data owners and thus by limiting the exposure all others, especially platform users, have to endure.

Our proposal also makes a conceptual contribution by clearly distinguishing between data markets and other forms of data sharing such as data spaces and cooperatives. This distinction is that between data exchanges and data markets. The word 'exchange' has acquired a double meaning of a place where commodities are physically exchanged and of a site where ownership in things is transferred. We suggest that this difference needs to be made explicit by distinguishing physical control from ownership (legal control). Concepts such as data spaces and data cooperatives are based on physical control while a data market is a site where ownership of data, not the data themselves, is transferred. While concepts based on physical control do not imply or require the definition of property rights, data markets come into existence only if ownership of data is defined.

REFERENCES

Abbas, A. E., Agahari, W., van de Ven, M., Zuiderwijk, A., and Reuver, M. de. 2021. "Business Data Sharing through Data Marketplaces: A Systematic Literature Review," *Journal of Theoretical and Applied Electronic Commerce Research* (16:7), pp. 3321-3339.

Agogo, D. 2021. "Invisible Market for Online Personal Data: An Examination," *Electronic Markets* (31:4), pp. 989-1010.

Cennamo, C. 2018. "Building the Value of Next-Generation Platforms: The Paradox of Diminishing Returns," *Journal of Management* (44:8), pp. 3038-3069.

Chan, N. K., and Kwok, C. 2021. "Guerilla Capitalism and the Platform Economy: Governing Uber in China, Taiwan, and Hong Kong," *Information, Communication & Society* (24:6), pp. 780-796.

Chen, A. 2021. The Cold Start Problem: Using network effects to scale your business, London: Random House Business Books.

Clough, D. R., and Wu, A. 2022. "Artificial Intelligence, Data-Driven Learning, and the Decentralized Structure of Platform Ecosystems," *Academy of Management Review* (47:1), pp. 184-189.

Commons, J. R. 1990. *Institutional Economics -- Its Place in Political Economy, Vol. 1*, New Brunswick, London: Transaction Publishers.

Denemark, J. 2022. "Google, Antitrust, and Digital Market Act: Is There New Hope for the AdTech Market?" in *EU Antitrust: Hot Topics & Next Steps*: *Proceedings of the International Conference held in Prague on January 24–25, 2022*, V. Šmejkal (ed.), Prague, pp. 40-53.

European Commission. 2020a. Proposal for a Regulation of the European Parliament and the Council on Contestable and Fair Markets in the Digital Sector (Digital Markets Act), Brussels.

European Commission. 2020b. Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act), Brussels.

Expertenkommission Forschung und Innovation. 2023. Gutachten zur Forschung, Innovation und technischer Leistungsfähigkeit Deutschlands: Gutachten 2022, Berlin.

Gregory, R. W., Henfridsson, O., Kaganer, E., and Kyriakou, H. 2022. "Data Network Effects: Key Conditions, Shared Data, and the Data Value Duality," *Academy of Management Review* (47:1), pp. 189-192.

Gregory, R. W., Henfridsson, O., Kaganer, E., and Kyriakou, S. H. 2021. "The Role of Artificial Intelligence and Data Network Effects for Creating User Value," *Academy of Management Review* (46:3), pp. 534-551.

Guse, R., Thiebes, S., Hennel, P., Rosenkranz, C., and Sunyaev, A. 2022. "Datenmarktplätze für Künstliche Intelligenz im Gesundheitswesen: Potenziale, Herausforderungen und Strategien zur Bewältigung," *HMD Praxis der Wirtschaftsinformatik* (59:6), pp. 1527-1544.

Höppner, T., and Westerhoff, P. 2021. "Privacy by Default, Abuse by Design: EU Competition Concerns About Apple's New App Tracking Policy," *SSRN Electronic Journal*.

Huq, A. Z. 2021. "Who Owns Our Data?: We need a model of ownership that recognizes our collective interests," *Boston Review*.

Jurcys, P., Donewald, C., Fenwick, M., Lampinen, M., Nekrošius, V., and Smaliukas, A. 2021. "Ownership of User-Held Data: Why Property Law is the Right Approach," *Harvard Journal of Law and Technolgoy, JOLT Digest*.

Krugman, P. 1991. "History versus Expectations," *Quarterly Journal of Economics* (106), pp. 651-667.

Koutroumpis, P., Leiponen, A., and Thomas, L. D. W. 2020. "Markets for Data," *Industrial and Corporate Change* (29:3), pp. 645-660.

Lubyová, L. H. 2022. "Digital Markets Act: A Fair Framework for the Online World?" in *EU Antitrust: Hot Topics & Next Steps*: *Proceedings of the International Conference held in Prague on January 24–25, 2022*, V. Šmejkal (ed.), Prague, pp. 54-64.

Macleod, H. D. 1875. *The Theory and Practice of Banking*, London: Longmans et al.

Mansell, S. F., and Sison, A. J. G. 2020. "Medieval Corporations, Membership and the Common Good: Rethinking the Critique of Shareholder Primacy," *Journal of Institutional Economics* (16:5), pp. 579-595.

Mulherin, J. H., Netter, J. M., and Overdahl, J. A. 1991. "Prices are Property: The Organization of Financial Exchanges from a Transaction Cost Perspective," *Journal of Law and Economics* (34), pp. 591-644.

The European Parliament and the Council of the European Union. 2016. "Regulation (EU) 2016/679 of the European Parliament and of the European Council of 27 April 2016 on the Protection of Natural Persons with regard to the Processing of Personal Data and on the Free Movement of such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)," *Official Journal of the European Union* (4.5.2016).

Thouvenin, F., and Tamò-Larrieux, A. 2021. "Data Ownership and Data Access Rights: Meaningful Tools for Promoting the European Digital Single Market?" in *Big Data and Global Trade Law*, M. Burri (ed.), Cambridge University Press, pp. 316-339.