

RESEARCH OUTPUTS / RÉSULTATS DE RECHERCHE

Data sharing for digital markets contestability

Feasey, Richard; DE STREEL, Alexandre

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (HARVARD):
Feasey, R & DE STREEL, A 2020, *Data sharing for digital markets contestability: towards a governance framework*. CERRE, Bruxelles.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

cerre

Centre on Regulation in Europe

REPORT

September 2020

Richard Feasey
Alexandre de Stree

**DATA SHARING
FOR DIGITAL MARKETS
CONTESTABILITY
TOWARDS A GOVERNANCE
FRAMEWORK**





The project, within the framework of which this report has been prepared, has received the support and/or input of the following organisations: ARCEP, BIPT, Facebook, Mediaset, Microsoft, Snap Inc. and Telefónica.

As provided for in CERRE's by-laws and in the procedural rules from its "Transparency & Independence Policy", this report has been prepared in strict academic independence. At all times during the development process, the research's authors, the Joint Academic Directors and the Director General remain the sole decision-makers concerning all content in the report.

The views expressed in this CERRE report are attributable only to the authors in a personal capacity and not to any institution with which they are associated. In addition, they do not necessarily correspond either to those of CERRE, or to any sponsor or members of CERRE.

© Copyright 2020, Centre on Regulation in Europe (CERRE)

info@cerre.eu

www.cerre.eu

Table of contents

About CERRE	5
About the authors	6
Executive summary	7
1 Introduction	14
1.1 Aim and scope of the report	14
1.2 Types of data	15
1.2.1 Personal data	15
1.2.2 Non-personal data	17
1.3 Types of data sharing	18
1.4 Types of interoperability	19
2 Incentives to share data	21
2.1 Incentives of the individual users to share their data	21
2.2 Incentives of firms to share aggregated user data	26
2.3 Barriers to data sharing and market failures	29
3 Competition policy and data sharing	33
3.1 Compulsory access under Article 102 TFEU	33
3.1.1 Relevant case-law	33
3.1.2 Conditions of essential facilities and application to data	36
3.2 Compulsory access under Merger Regulation	38
3.3 Limits of data sharing by Article 101 TFEU	40
4 Horizontal and sectoral EU laws for data sharing	44
4.1 Existing legislations enabling data sharing	44
4.1.1 EU rules on data portability	45
4.1.2 EU rules on data sharing	48
4.2 Existing legislations limiting data sharing	51
5 Regulatory governance of data sharing	55
5.1 The scope of data sharing	55
5.1.1 Which digital platforms should be obliged to provide access to data?	55
5.1.2 Which organisations are entitled to obtain access to data (and the conditions to meet before they can do so)?	59
5.1.3 The types of data to be shared, the geographic scope from which it is drawn, and the conditions under which sharing can occur, including on whose initiative?	62
5.2 The conditions of sharing	70
5.2.1 Technical standardisation	71
5.2.2 The data transfer process	73
5.2.3 Rights of redress and other contractual matters	75



5.2.4	Prices for data access.....	77
5.2.5	Promoting disruptive business models	82
5.2.6	Exiting from data sharing arrangements	83
6	Policy recommendations.....	86
	References	92

About CERRE

Providing top-quality studies and dissemination activities, the Centre on Regulation in Europe (CERRE) promotes robust and consistent regulation in Europe's network and digital industries. CERRE's members are regulatory authorities and operators in those industries as well as universities. CERRE's added value is based on:

- its original, multidisciplinary and cross-sector approach;
- the widely acknowledged academic credentials and policy experience of its team and associated staff members;
- its scientific independence and impartiality;
- the direct relevance and timeliness of its contributions to the policy and regulatory development process applicable to network industries and the markets for their services.

CERRE's activities include contributions to the development of norms, standards and policy recommendations related to the regulation of service providers, to the specification of market rules and to improvements in the management of infrastructure in a changing political, economic, technological and social environment. CERRE's work also aims at clarifying the respective roles of market operators, governments and regulatory authorities, as well as at strengthening the expertise of the latter, since in many Member States, regulators are part of a relatively recent profession.

About the authors



Richard Feasey is a CERRE Senior Research Fellow, an independent consultant and a Senior Adviser to the Payment Systems Regulator in the UK. He lectures at University College London and Kings College London. Richard was previously the Public Policy Director of Vodafone Group plc from 2001 until 2013. In October 2017, he was appointed to the panel of the UK's Competition and Markets Authority and in October 2018 to the National Infrastructure Commission for Wales.



Alexandre de Streel is a Professor of European law at the Universities of Namur and Louvain in Belgium. He is Academic Co-Director at the Centre on Regulation in Europe (CERRE) and the Director of the Research Centre for Information, Law and Society (CRIDS), focusing his research on Regulation and Competition Law in the network industries. He is a member of the Scientific Committee of the Florence School of Regulation (FSR) at the European University Institute in Florence and Research Fellow at European Institute of Public Administration (EIPA) in Maastricht. Professor de Streel regularly advises international organisations (such as the European Commission, European Parliament, OECD, EBRD) and national regulatory authorities on regulatory and competition issues in network industries. He is also an Assessor (member of the decisional body) at the Belgian Competition Authority.

Executive summary

There is today growing momentum behind proposals for 'data sharing' as a remedy for competition concerns in digital markets, as well as efforts by the European Commission (including in the forthcoming Data Act of 2021) to promote the sharing of data more widely in the European economy. However, there is as yet less focus on the practical challenges that will need to be overcome to implement data sharing arrangements that effectively promote innovation and competition in or preserve the contestability of digital markets. There is also limited experience of regulated data sharing in practice. This report aims to address that gap and offers a series of recommendations on what authorities will need to do if data sharing by digital platforms is to occur at scale in Europe. The report complements another CERRE Report on the role of data for digital market contestability.

I. Incentives to share data

Data is already being shared voluntarily under a wide variety of conditions and for a variety of reasons. One category of data sharing arrangements is those initiated by individuals, which normally involve personal data, and in which the benefits of sharing generally accrue to that individual. Examples include digital platforms that allow individuals to download their data to better understand what has been collected about them; laws such as the General Data Protection Regulation (GDPR) which allow individuals to initiate a transfer of data from one organisation to another to switch service provider or 'multi-home' across several providers; other regulations such as the Second Payment Services Directive (PSD2) or the Open Banking regime in the UK which allow individuals to initiate the continuous sharing of data with providers of complementary services; and 'Personal Data Stores' who manage data on behalf of individuals and provide greater security, convenience or opportunities to monetise data.

Although there are many such opportunities for an individual to share data today, **very few appear in fact to do so.** Some major governmental initiatives to promote data sharing, such as Midata in the UK or the smart energy meter data programmes in Europe, have failed to meet expectations. This seems to be because users have a low level of trust in the arrangements, find the process complex and time consuming, or find it difficult to evaluate the benefits they might obtain from doing so.

Data can also be shared between organisations voluntarily, normally in bulk and without first requiring the consent of individual users. In this case, the benefits of data sharing are likely to be enjoyed by a large number of users rather than being confined to a specific individual. Organisations such as insurance companies may share data when there is a mutual advantage to doing so (to identify fraudulent activity), digital platforms, such as Facebook, may allow others to access data they hold to encourage complementary innovation or may provide 'ancillary' services such as identity management in return for access to the data of those other firms that use them. Firms like Bloomberg or Nielsen may collect and sell data, or, like MasterCard, may donate data to support research or other causes. Firms may not share the data itself, but may allow others to interrogate it through 'sandboxes' or 'trusted intermediaries'. Many public organisations share significant volumes of data generally without charge and the Open Data Directive imposes significant data sharing obligation of public sector data.

Although more data is shared between organisations, **they encounter similar issues to individuals.** It can be complex and difficult to agree on the technical standards required for data to flow smoothly between them. Firms may be uncertain about the legal status of the data over which they exercise control and whether sharing may expose them to unforeseen liabilities. They may also be uncertain about the credentials of the intended recipients or their capacity to keep data secure, or about the kinds of reputational risks which became apparent, for example, after Cambridge Analytica obtained access to data on Facebook users. For these and other reasons, many organisations may conclude that the costs and risks of data sharing outweigh the potential benefits.

Some form of regulation may be required to overcome some of the barriers to sharing data, even when both parties otherwise have incentives to share data and would benefit from doing so. This report is, however, concerned with circumstances in which one of the parties has powerful incentives not to share data because it is a significant source of a competitive advantage which is difficult or impossible for others to replicate, and which therefore allows that platform to preserve its 'gatekeeper' position in its core market and at the same time to leverage these advantages into other markets. The objective of imposing an obligation to share data in these circumstances is therefore to preserve the contestability of adjacent markets as well as, more speculatively, to support rivalry in the core market and ultimately, to promote data driven innovation in the EU.

II. EU legal framework for data sharing

Competition laws may impose data sharing obligation under some strict conditions. If data could be considered as essential facilities, the refusal to share such data may be considered as an abuse of dominant position prohibited by Article 102 TFEU. However, the conditions of the essential facilities, even when they are adapted to take into account the specific characteristics of data, are difficult to meet and, to date, very few refusals to share data have been considered as abusive. Moreover, when two data-rich firms merge, the competition authority may impose some remedies if the combination of previously separate data sets would significantly impede effective competition. In such circumstances, the authority may either impose the merging parties to share data with their competitors (as has been the case in *Thomson/Reuters*) or impede the combination of data sets by the merging parties (as may be the case in the yet to be decided *Google/Fitbit* case).

Alongside competition law, the **EU horizontal or sector laws also contain several rules that stimulate or impose data sharing and data portability.** Concerning horizontal rules applicable to all sectors of the economy, the obligations focus mainly on the portability of personal data (with the 2016 General Data Protection Regulation and the 2019 Digital Content Directive). The portability of non-personal data is encouraged, but not imposed, by the 2018 Free Flow of Data Regulation. Although steps in the right direction, these rules have several limits and shortcomings and they do not provide for a fully-fledged data sharing framework. The most comprehensive data sharing obligation and governance framework is imposed by the 2019 Open Data Directive which applies to data owned by public sector bodies and public undertakings in the EU. There are also extensive data sharing obligations in several sectoral legislation, for instance in the financial sector (2015 Second Payment Services Directive), the automotive sector (2018 Motor Vehicle Regulation), and the energy sector (2019 New Electricity Directive).


Thus, **while the EU legal framework contains some rules imposing data sharing, rules are in general limited and do not provide for a comprehensive and effective governance framework to share data.**

III. Recommendations for an effective governance framework in case data sharing is imposed

This report does not recommend that a particular institution, whether at the national or European level, is given the task of regulating data sharing. Whatever the precise institutional arrangements, we identify several challenges which a regulator will need to overcome. These include determining the identity of the digital platforms that will be obliged to share data; deciding the conditions under which data is shared and the obligations of recipients; the user experience (if user consents are required); the scope and other characteristics of the data to be shared; arrangements for the governance of data sharing and the resolution of disputes and errors; and the commercial or other terms under which data is shared. The report presents several conclusions and makes several recommendations.

Regulating recipients as well as donors

The report concludes that **regulation for data sharing should not be viewed as being limited to the oversight of a small number of large platforms that might be obliged to share data.**



That is because it will also require strict oversight of potentially a very large number of smaller firms that might seek access to such data and which may then rely upon it to provide services of various kinds. Given the potentially wide range of uses to which data could be applied, and the wide range of organisations which may require access to such data, individual users will not consent to the sharing of data unless they can be confident that any recipient of the data will keep it secure and will adhere to other conditions of sharing, so as to preserve trust in, and the integrity of the overall data sharing process. The controllers of commercial data will also be rightly concerned about bulk sharing obligations if misuse by others puts their reputation or commercial position at risk. Recipients of data may be also putting themselves in a position of acute dependency (since they may rely upon uninterrupted data sharing to sustain their services for users) and will not enter into such arrangements unless they consider that they have adequate protections and rights of redress in the event of any disruption or interruption in supply. A comprehensive system of regulation of both donors and recipients of data will be required to guard against misuse and to ensure trust on all sides.

It follows that if regulated data sharing is to be adopted at a significant scale, regulators will need to establish an effective regime for overseeing those in receipt of data and for enforcing the rules effectively on an ongoing basis. This will need to include rules governing the resolution of disputes and determining how liabilities fall if consumers or other firms are harmed. Since many of those who share or receive data are unlikely to hold market power or otherwise to be guilty of any abuse, we consider that oversight of such arrangements is unlikely to be an appropriate task for a competition authority and will instead require a dedicated regulatory body.


Extensive obligations to adopt common technical standards

All forms of data sharing will require the **adoption of common technical standards by both those sharing data and those in receipt of it.** The same standards should be adopted for all the different forms of data sharing that we propose. We consider that potential recipients of data have sufficient incentives to adopt the standards since they would not otherwise obtain access to the data they require. Those platforms that have been directed to share data will need to be obliged to adopt the relevant standards, such that data can be shared in a form and manner which supports the regulatory objectives. In the early stages of regulation, this may impose additional costs on the newly regulated entities as they have to restructure the way they manage their existing data assets or adopt new external interfaces. This may also contribute to delay in the implementation of new data sharing obligations, which will be a particular concern if the objective of data sharing is to prevent leveraging into emerging digital markets. In the longer term, we conclude that data sharing regulation should promote the very extensive adoption of common technical standards by organisations which may not currently have obligations to share data (but which might be required to in the future), those who may not currently request access to data (but will want to preserve the option to do so in the future), and in relation to forms of data which may not currently be shared (but which may be required to be shared in future). **This 'anticipatory' approach to technical standards means that regulators should consider the application of common technical standards to data sharing in sectors well beyond the existing scope of large digital platforms,** as has been proposed in Australia. In short, **we recommend regulators should decouple requirements to adopt common technical standards from obligations to share data in the expectation that the former will be much more extensive than the latter.**

The most important and difficult role for regulators will lie in determining the type and scope of data that is to be shared and which organisations should be obliged to share it. **We conclude that two forms of regulated sharing are likely to dominate.**

Recommendations on sharing of data about individual users

The first form of sharing – and **the one which is likely to be capable of being implemented first** – will be the **sharing or porting of data about individual users.** This mode of sharing is likely to be appropriate when the individual concerned will benefit directly from the sharing process, usually through the provision by the recipients of complimentary services in adjacent markets. The



value of the data, in this case, lies in its depth and personalised nature, rather than in its volume. The process to enable the sharing of the data will generally require that the user consent to the transfer, and the process by which these user consents are obtained and authenticated will have a significant impact on the effectiveness of this remedy. Technologies such as biometric IDs will have a significant role to play.

The data to be transferred would be data provided by the user to the platform and data derived from observations of that individual's interactions with the platform. It would exclude 'inferred data' that is created by the platform itself (as well as excluding third party data that is purchased from other sources). The presumption should be that all relevant data about an individual would be shared.

The overall competitive impact of these data sharing arrangements will necessarily be limited, given the relatively high transaction costs associated with first obtaining individual consents from every user and the relatively small volumes of data that will be transferred each time consent is obtained. Over time, however, data that is obtained in this way could accumulate and be used for other purposes. For this reason, **we recommend that obligations to share data about individual users in the way we propose should be quite extensive and apply to digital platforms which we would describe as meeting the 'gatekeeper minus' threshold.** This would mean a strong presumption that the obligation to share would apply to all platforms which the regulator had determined as having 'gatekeeper' or equivalent status and to some others as well. However, **this obligation would not apply to every platform or firm**, and so would be less extensive than, for example, **the 'data portability' obligations which apply under the GDPR (which are narrower in scope).** We do not recommend that the European Commission seek to expand the existing GDPR data portability requirements to address the competition concerns we consider in this report and conclude that a separate regime, specifically designed for this purpose, is the better approach.

We consider that there is a case for a regulator to require the sharing of individual user data without any form of payment passing between the donor and the recipient. Each party would be expected to bear its costs to the transfer.

It is unclear at this stage how effective the arrangements for the sharing of individual data outlined above would prove to be. However, there is a risk that the high transaction costs and uncertain benefits continue to deter users and render this approach relatively ineffective in preserving the contestability of the markets we are concerned with. **In such circumstances, we recommend the European institutions should consider more radical approaches, including changes to the GDPR which would allow for individual users to 'opt out' their data** (rather than requiring them to 'opt in') when transfers of their data are initiated - provided always that the recipients of the data comply with the relevant regulatory conditions.

We recognise that this may represent some loss of consumer sovereignty over their data, but consider that such a trade-off may need to be made if data sharing arrangements are to achieve their aim of ensuring contestability in digital markets. It is far from clear that the interests of European consumers are better served by preserving rights to consent whilst allowing new digital markets to be dominated by existing 'gatekeeper' platforms. Indeed, in the long run, the privacy rights of European consumers may be better served by measures that more effectively promote competition. At the very least, the debate should be had and **we, therefore, recommend the European Commission consider provisions in the forthcoming Data Act to enable the use of 'opt out' arrangements for the sharing of personal data to preserve market contestability under certain prescribed conditions.** There is certainly a precedent for such arrangements, since control of personal data sets has often changed without individual user 'opt ins' when one firm acquires another firm or when one firm acquires another's data assets.

Recommendations on the bulk sharing of user data

The second form of sharing will be the bulk transfer of aggregate user data. As with the first category, this would involve sharing data provided by individual users or arising from their interactions with the platform but would exclude inferences that are generated by the platform itself. This mode of sharing is likely to support entry into adjacent or emerging markets, with such entry being supported by insights derived from large data sets. It may even to support competition in some or all of the core market activities from which, or by means of which, the data has been derived.


The overall competitive impact of these data sharing arrangements could be significant – likely more significant than for individual user data - since the volume of data to be shared is likely to be very substantial and may represent a significant proportion of the donor platform's data assets. In some circumstances, it may be necessary for the data to be shared without first anonymising it to allow recipients to effectively rival the incumbent platform. **Since obtaining individual consents from every user would not be feasible in these circumstances, we recommend that regulators and policymakers consider other mechanisms to enable the bulk sharing of non-anonymised user data.**

Alternatively, regulators should consider requiring the platform that controls the data to allow third party access to the full data set for training algorithms or otherwise deriving the same sorts of insights from the data that are available to the incumbent. The terms under which such access is provided would also need to be carefully regulated since those seeking access to the data sets would remain dependent upon the owner of the assets providing full and unrestricted access. Similar challenges arise even the data is held by a 'neutral' intermediary. Such arrangements are therefore likely to require a high degree of regulatory oversight (and associated cost), although they also have considerable attractions if non-anonymised data is important to preserve contestability or if very large data sets are involved.

Although we would expect all the relevant data about an individual user to be shared with every recipient, **there is likely to be much greater heterogeneity of demand amongst potential recipients of bulk transfers of aggregate data.** Some potential recipients may require (or may only be able to handle) relatively small volumes of data, representing only a fraction of that held by the donor. Others may require the sharing of much larger data sets. There may also be questions about the geographic scope of the data to be shared. This will present two challenges. First, the regulator will need to ensure that a **suitable menu of data options** is developed, preferably collaboratively and inclusively, to ensure that the needs of as wide a range of potential recipients as possible will be met as far as possible. This is likely to involve a degree of compromise on the part of some parties, with the regulator adjudicating between conflicting demands.

Second, **we consider there is a strong prima facie case for assuming that recipients of aggregated data should be required to pay the data, with the payment varying by the volume and value of the data being shared (and not simply the costs of implementing the data sharing arrangements or storing the data).** The primary concern here is to preserve incentives for both parties in the sharing arrangement to innovate and invest in existing or new digital services to acquire additional data for themselves. We do not want data sharing arrangements to crowd out other forms of commercial activity from which users derive significant benefits, particularly in many digital markets.

However, we do not make firm recommendations as to how these prices should be derived because we have yet to find a well-developed methodology for doing so. Requiring firms to agree with terms on a 'FRAND' basis may not be adequate in several circumstances. **We recommend that a study be undertaken by the Commission to consider how regulators would establish wholesale prices for data that was to be shared.** The methodologies and the practices which were developed for the public open data framework to calculate the marginal costs and to recover costs developed for public data may feed this study. We also consider that setting appropriate wholesale prices for the receipt of aggregate data will also be necessary to ensure that recipients have appropriate



incentives to reassess their data requirements as they grow and develop their businesses, allowing for the possibility that they would terminate existing data sharing arrangements once they have acquired, or are in a position to acquire, sufficient data for themselves from their users. Otherwise, extensive data sharing arrangements could likely become a permanent feature of European digital markets in the years to come.

The final question in this context concerns the identity of the platforms that would be obliged to share aggregated personal data on a bulk basis. **We conclude that this should be a much-limited set of entities than we recommend for the porting of individual data** and would not necessarily be a requirement of every platform that was found to hold 'gatekeeper' status under the European Commission's latest proposals, although we think a designation of 'gatekeeper' status should establish a rebuttable presumption. **We, therefore, characterise this sub-set of entities as being those that meet a (more demanding) 'gatekeeper plus' threshold.** The analysis required to demonstrate this would need to be undertaken on a case by case basis.

The challenge ahead

The recommendations in this report, if adopted, would represent an extensive programme of regulatory activity that would need to be undertaken by bodies with responsibilities for implementing data sharing which have yet to be assigned in Europe. **Establishing the institutional and regulatory framework to deliver data sharing at scale will require legislation.** Moreover, we recommend that the European policymakers consider further legislative changes in the forthcoming Data Act to enable the sharing of personal data on an opt out basis under certain narrowly prescribed circumstances and to ensure contestability in digital markets.

Finally, we are mindful that data sharing remedies that we have considered in this report arise from the assumption that digital platforms will continue to derive significant market power from their centralised control of big data sets which they have accumulated by enabling diffuse groups of users to transact with each other through the platform. This may be the case, but regulators and policymakers should also keep an eye on (and potentially take steps to promote) **new technologies and architectures which might in the future enable a much greater degree of decentralisation and wider distribution of data**, thereby removing the very sources of market power which this report has sought to address.

01



INTRODUCTION



1 Introduction¹

1.1 Aim and scope of the report

This report seeks to engage with the practical challenges which public authorities will face if they were to decide to oblige digital platforms to share their data with potential competitors. There is today considerable and growing momentum behind proposals for 'data sharing' as a remedy for competition concerns in digital markets, but as yet little detail as to what they might consist of or how they might work. This report can be read alongside other research that has been undertaken by colleagues at CERRE² which has produced detailed case studies of data used by specific types of digital platforms (online search, e-commerce, and media platforms) and detailed proposals for how data sharing might promote market contestability, competition, and innovation in each case. This report complements their work by considering what practical steps public authorities would need to take, and what issues they would need to address if the sorts of data sharing arrangements which they propose were to be effectively implemented³.

We start by reviewing how data is shared in other contexts and what might be learned from the arrangements that have been instituted to facilitate such sharing. Unsurprisingly, we observe that data is shared when parties to the arrangement derive a benefit from doing so. For example, we suggest that if the sharing of data is to require the prior consent of a particular user then such arrangements are only likely to work if that user perceives a direct benefit (or 'private value') from the arrangement. If both all parties to the arrangement stand to benefit, then sharing may occur voluntarily without the intervention of public authorities, although there may still be significant obstacles and transactions costs which explain why, in the view of many, the potential benefits of voluntary data sharing are a long way from being fully realised in Europe today.

If some parties stand to gain from sharing, but others not, then some form of intervention is likely to be required before it occurs. We assume such an intervention would be required to oblige a large digital platform to share data which it holds with a potential competitor. We, therefore, consider the various existing legal instruments which have been or could be, used to require firms (and public bodies) to port or share data. Some of these instruments, such as the General Data Protection Directive, require data portability to fulfil goals other than the promotion of competition, whereas others, such as competition law, clearly have competition as their primary focus.

From this review, we identify several practical questions that will need to be addressed by any data sharing regime which seeking to promote or safeguard market contestability, competition and innovation in digital markets. The answers to some of these questions depend upon the form of competition which any measures might be intended to promote. In many cases, we would expect data sharing to be employed as a means of promoting 'complementary innovation' and preventing foreclosure in adjacent markets rather than having the more ambitious aim of displacing a dominant digital platform from the core market from which it derives its data advantages. Sometimes these arrangements might allow data intermediaries or data aggregators to help individual users enjoy greater benefits from multi-homing. Sometimes they may instead enable innovation and competition in adjacent markets that benefit users in general.

There are some signs that European policymakers are prepared to incentivise more data sharing. Since at least 2018, the Commission has sought to facilitate arrangements that allow firms to share

¹ The authors are grateful to Bruno Liebhaber, Jan Krämer and Thomas Tombal for their helpful comments and suggestions.

² Krämer, Schnurr and Broughton Micova (2020).

³ For a very recent paper on the same theme, see Kerber (2020).

data amongst themselves or with public bodies⁴. In February 2020, the Commission adopted 'A European strategy for data'⁵ which identified, amongst other things, existing barriers to the sharing of data between businesses, including: 'a lack of economic incentives (including the fear of losing a competitive edge), lack of trust between economic operators that the data will be used in line with contractual agreements, imbalances in negotiating power, the fear of misappropriation of the data by third parties, and a lack of legal clarity on who can do what with the data (for example for co-created data, in particular, IoT data)⁶. Many of these issues are directly relevant to the data governance arrangements we discuss in this report and the Commission currently intends to address many of them by way of a new Data Act in 2021.⁷

In addition to facilitating data sharing more broadly, the Commission also engaged in what it describes as 'broader fact-finding around the high degree of market power of certain platforms and also in the context of the Commission's work on the Digital Services Act package. Based on this fact-finding, the Commission will consider how best to address more systemic issues related to platforms and data, including by ex-ante regulation if appropriate, to ensure that markets stay open and fair.'⁸ Recent papers produced by the Commission indicate that relevant cases studies will include "situations, where the market does not provide for a market based solution and these platforms are unwilling to share their data, it may be necessary to require such platform through ex ante regulatory measures to offer access to the required data on reasonable, standardized and non-discriminatory terms."⁹ The Commission is also consulting on proposals to adopt what it refers to as a 'new competition tool' and/or an 'ex-ante regulatory instrument of very large online platforms acting as gatekeepers', either or both of which, if adopted, would be likely to support the imposition of obligations to share data¹⁰.

This report discusses the practical challenges that would need to be overcome if these or other initiatives are to lead to data sharing remedies that ensure digital markets remain contestable.

1.2 Types of data

1.2.1 Personal data

Before considering the many different types of existing data sharing and portability arrangements that can be observed in Europe today, it is important to introduce some of the terminologies we employ in this report. There are many different types of data but no comprehensive or authoritative taxonomy. For our purposes, a distinction should be drawn between what is commonly referred to as 'personal' data and 'non-personal' data. This is important because (as we explain further in section 3), under the General Data Protection Directive (GDPR), individual users have certain legal claims over and rights concerning personal data that do not apply to non-personal data. **Personal data** is

⁴ In 2018, the Commission published a Staff Working Document 'Guidance on sharing private sector data in the European data economy, SWD(2018) 124 which provided firms with basic guidance on the contractual issues which typically need to be addressed, including the scope and nature of the data to be shared, the use to which it can be put, the parties with which it is to be shared, how the data is to be protected, the technical means by which access is to be provided, liabilities if data is of poor quality or is false, supply is interrupted or data is destroyed or lost. The guidelines encourage (but clearly cannot oblige) European firms to adopt standardised APIs in order to facilitate data transfers. In the same year, the Commission adopted a Communication, 'Towards a common European data space' (COM(2018) 232), in which, amongst other things, it proposed that 'Support Centre for data sharing under the Connecting Europe Facility programme will put in place a set of measures to make it easier to share private sector data in addition to public sector data'. It offers know-how and assistance on data sharing by providing best practice examples and information on APIs, existing model contracts and other legal and technical aspects⁴. The Commission also proposed to promote the adoption of common APIs.

⁵ Communication from the Commission of 19 February 2020, A European strategy for data, COM(2020) 66.

⁶ Ibid p.8

⁷ In July 2020, the Commission services consulted on proposals for a regulation governing 'Common European Data Spaces', Ares (2020)3480073.

⁸ Ibid p.14

⁹ European Commission, Inception Impact Assessment for Digital Services Act Package: ex ante regulatory instrument

¹⁰ Commission launches consultation to seek views on Digital Services Act package', press release, IP/20/962

defined as 'any information relating to an identified or identifiable natural person'¹¹. This could encompass a very wide range of data in the context of a digital platform. It is therefore common to sub-divide such personal data into at least four categories (an approach which has been employed by, amongst others, the OECD, European data protection authorities, and advisers to the Commission¹²).

First, there is '**provided**' data, which is data about an identifiable user that has been provided to the platform by the user themselves, often (but not always) at the time when they first sign up for a particular digital service. Data may also be provided, for example, every time a user posts new photos or comments on their social media service. Such data may have considerable value for digital platforms but is commonly regarded as having been provided by, and so remaining under, the control of the user who provided it. It is replicable in the sense that if a user can supply data to one digital platform then they could, in principle, supply the same data to other platforms as well. However, if users have provided a lot of data over a long period, or if they use the digital platform itself as a means of storing data, it may be more difficult to replicate provided data unless the user can themselves download a copy from the platform in question. Accordingly, the GDPR provides users with the rights to download data which they have themselves earlier provided to the platform without incurring costs by doing so.

Second, there is '**observed data**', which is data that is generated through interactions between an identifiable user and the platform. We consider this category of data is often the most important when it comes to sources of competitive advantage in digital markets since it is by observing the responses and actions of individual users that digital platforms are often able to generate predictions about how individual users will interact in future or what might best or more closely fulfil their needs. A standard example is the observed propensity of individual users to click on links that are served in response to search queries. This data will provide a search engine with information both about the likely preferences and needs of that individual user, but also with valuable inferences, when combined with observed data from other users, about what all users with similar characteristics might want¹³. Importantly, for our purposes, observed data is difficult to replicate since its creation relies upon what is likely to be a very large volume of interactions between the individual user and the digital platform itself, often undertaken over an extended period. To replicate this data, a user would not only have to switch from one digital platform to another but would then have to engage with the new provider for an extended period to allow them to generate a similar profile and data set. However if, as is often the case, observed data helps to ensure higher quality interactions with the platform, the user may find themselves having to forgo these benefits when they switch to a new provider. Moreover, a large number of users may need to co-ordinate their switching to a new provider if they are collected to benefit from the insights that a digital platform obtains from being able to analyse aggregated sets of observed data.

In one sense, observed data is co-created by the user and the platform since it cannot arise without the presence of both parties. The GDPR considers that users should retain similar rights over observed data as they do over provided data, including the right to download it at no cost and to require that it be transferred to another platform if technically feasible. The platform's rights over observed data will depend upon the consents it has obtained from users and the contractual obligations it is subject to.

¹¹ Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46 (General Data Protection Regulation), OJ [2016] L 199/1, Article 4

¹² See Guidelines of 13 April 2017 of Working Party 29 on the right to data portability, WP242 rev.01. Also Cremer et al.(2019) and OECD (2019).

¹³ For a detailed discussion, see Krämer, Schnurr and Broughton Micova (2020).

Third, there is '**inferred data**', which is data that is derived from the analysis of other data that is undertaken by the platform itself. The creation of inferred data does not, therefore, require any further interaction with the user even though the inferences may relate to that user. A digital platform's capacity to generate insightful, innovative, or useful inferences from the data that it has accumulated may be an important source of competitive differentiation and advantage in markets where such insights can be translated into differentiated services. Inferred data that is valuable may, therefore, be difficult to replicate and may reflect genuine innovation on the part of the platform that holds it. At the same time, it may be sufficient for a rival platform to obtain access to the underlying inputs - the provided and/or observed data from which the inferences are drawn - to be able to compete effectively. Inferred data is not, therefore, commonly regarded as being under the control of the individual user, although a digital platform may still require their consent if such data were to be shared.

Fourth, a digital platform may be able to **acquire personal data from a third party** rather than obtaining it from the user or their interactions with the user. Data of this kind may be acquired to complete profiles of users when they have neglected to provide certain personal details, or for other purposes. As we note in the next section, such data is likely to be available for sale from a variety of specialist data providers and so easily obtained by rival platforms as well. Interventions by public authorities are unlikely to be necessary unless a large digital platform were able to acquire competitively valuable data on an exclusive basis. We are not aware of this having been a concern to date, and so the sharing of acquired personal data is not considered further in this report.

1.2.2 *Non-personal data*

Having discussed four discrete categories of personal data, the **same categories can be applied for our purposes to non-personal data**, being information that does not relate to an identifiable individual. Thus, non-personal provided data may be information that was originally provided by an identifiable user but which has subsequently been reconfigured, or anonymised, so that the identity of the user who provided it is no longer apparent. Similarly, non-personal observed data will have been generated by interactions between the platform and identifiable users, but subsequently manipulated so that the identities of those users are removed.

The issue of **anonymisation** is a highly technical one and has become a subject to some controversy. On the one hand, the anonymisation of data may remove privacy concerns and allow data to be shared under less onerous conditions, including without the need to obtain consent from individual users before doing so. This should facilitate data sharing. Indeed, by its nature, anonymisation requires the aggregation of large volumes of data. On the other hand, anonymisation involves the extraction of information that may otherwise have significant value. An anonymised set of data that is shared is likely to be less competitively valuable than the source data which is retained by the digital platform that obtained it directly from individual users. Anonymisation, therefore, facilitates the sharing of data in bulk, but with a necessary loss of valuable information in the process.

Aside from personal data that has been anonymised, non-personal data may of course also refer to data that has not been acquired from or is the result of interactions with, any identifiable users. Digital platforms may obtain information from a vast range of sources, including their own and third party private and public organisations, to support the provision of relevant services to their users. Much of this data (such as weather forecasts or traffic or footfall data) is available from third parties, easily replicated by all platforms, and not a key focus in this report. We should, however, note that concerns have increasingly been raised about the terms of access to data that is increasingly generated by sensors and other Internet of Things devices and which is critical to the provision of maintenance, diagnostic, and other ancillary services relating to machinery. We discuss in section 3, for example, the European regulation which was introduced to enable independent garages to obtain access to the data that is required to undertake the repair and servicing of modern automobiles.

Similar issues have arisen concerning access to and the ownership of data that is generated by tractors in the agricultural sector¹⁴.

1.3 Types of data sharing

It is also important to explain what we mean by the term '**sharing**' when we employ it in this report. Again, there does not appear to be an authoritative taxonomy but we take 'sharing' for our purposes to involve the provision by a digital platform of access to specified categories of data to third parties continuously. This distinguishes 'data sharing' from a process that more often involves the one-off transfer of a specified set of data from one firm to another, generally at the initiative of the individual user¹⁵. The distinction is important because provided data is invariably provided only once by a user, which means that a one-off transfer of that data to another platform is sufficient to replicate the data. However, as noted earlier, observed data is generally accumulated through the interaction between a user and the platform over time, with more recent data likely to be more valuable and competitively significant than older interactions. If new entrants seek to compete directly with an incumbent digital platform, it seems likely that users will seek to multi-home across both platforms, at least until the new competitor has established its credentials. That requires an effective data access remedy to allow for the continuous transfer of observed data from the incumbent platform to the entrants, at least for some time. However, we have also explained that we consider that data sharing remedies are often likely to be used to safeguard or promote competition in adjacent markets rather than displacing the incumbent platform altogether. In these circumstances, the user will continue to interact with the incumbent platform in the upstream market irrespective of whether there is an entry in the downstream or adjacent market. The incumbent platform will therefore continue to obtain access to observed data and effective competition in the adjacent markets may require that such data continues to be shared.

Although there are many possible forms of data sharing, we follow our CERRE colleagues in considering that two are most likely to be relevant when the objective is the promotion of competition in digital markets.


- The first of these involves the sharing of personal data about an individual user to facilitate the provision of services in adjacent markets from which they stand to benefit. In this case, arrangements will be needed to obtain appropriate consent from the user. We refer to this as **(continuous) sharing or portability of individual user data**.
- The second involves the sharing, in bulk, of data about aggregate user behaviour or preferences to facilitate the provision of matching and other services that require access to such data to compete. We refer to this as **(bulk) sharing of aggregate user data**. A different set of arrangements (and different thresholds for intervention) will be appropriate in this case.¹⁶

The sharing arrangements should be informed by the type of data sharing that is required to effectively remedy the competition concern. We consider how existing European laws, particularly those relating to the safeguarding of rights to privacy, might be accommodated within the arrangements we envisage and what might need to change for any measures to be fully effective. We also identify other aspects of European law which could and should be improved upon. Some of these changes may be controversial or difficult, but we consider that data sharing remedies are very

¹⁴ OECD (2019), p.99

¹⁵ Cremer et al (2019) p.83

¹⁶ A further (but in our view difficult) distinction might be drawn between data that has been acquired as a by-product of use of the platform (such as search queries) and data which forms an integral part of the service itself (such as content posted by users on social media). For further discussion of these distinctions, see Krämer, Schnurr and Broughton Micova (2020), Section 3 and 4.4.2



unlikely to achieve their competition objectives in Europe unless policymakers are prepared to commit seriously to the effort and to overcome some formidable obstacles.

1.4 Types of interoperability

Several other studies also introduce distinctions between different forms of **'interoperability'**.¹⁷ One important case is what the Commission's advisers describe as **'full protocol interoperability'**, which refers to arrangements under which the services of competing platforms will interwork or interoperate with each other. A standard example would involve the users of one digital messaging service being able to communicate directly with the users of another, rival messaging service without them having to subscribe to the other service. Alternatively, users on a particular social messaging platform might be able to post photos or comments on the pages of users on another social messaging platform (and vice versa). 'Full protocol interoperability' might be required as a competition remedy in markets exhibiting strong direct network effects, particularly if users single-home rather than multi-home. In such markets, new entrants may be otherwise unable to persuade users to forgo the benefits of network effects by switching away from the incumbent platform.

There are many interesting issues associated with the imposition of 'full protocol interoperability' remedies of this kind, but they are not the focus of this report. Our interest is with the sharing of data held by digital platforms rather than the interworking of services that they may provide. As we explain in more detail in Section 5.2.1, the effective sharing of data between firms continuously is likely to require a significant degree of interoperability, likely facilitated by the provision of various **standardised technical interfaces and the adoption of common standards for data models** (which reference how data and meta-data are structured and organised so that it can be analysed, manipulated and extracted for relevant purposes). There are various ways in which such 'interoperability' might be accomplished, each involving different costs and benefits, and so we envisage the development of technical arrangements as being more of a process than an event. Their form will be determined to a large degree by prior decisions about the type of data that is to be shared and the terms under which this is to be done. We do not, therefore, consider it necessary to associate our consideration of data sharing with any particular form of interoperability or set of technical arrangements.

¹⁷ Kerber and Schweitzer (2017).

02



INCENTIVES TO SHARE DATA

2 Incentives to share data

Data sharing is undertaken for a wide range of purposes, and has been, or could be, implemented under a wide range of different arrangements, some of which involve regulation but many of which are entirely voluntary. In this section, we review some of the arrangements that exist and consider what might be learned from them. In subsequent sections, we discuss existing examples of European rules of data sharing. Later we consider the implications of our findings for the rules and governance of regulated data sharing to promote competition.

As we noted in the introduction, data sharing arrangements are likely to emerge when both parties to the arrangement derive some benefits from participating¹⁸. Economists often remind us that data has certain properties that may make it easy to share. Data is non-rivalrous, meaning that the same data can be used by different parties for different purposes without the activities of one party affecting the opportunities of others. There also appear to be economies of scope in aggregation, meaning that data from different sources can be combined to yield insights and benefits which could not have been obtained from either of the sources independently¹⁹. This is in addition to the additional benefits which might be derived from simply increasing the volume of data that is available to different parties. It is these unusual properties of data that may provide strong incentives to share data for mutual benefit, but which may equally provide incentives for firms and others to deprive potential competitors of access to data.

Sometimes the beneficiaries of data sharing are particular individuals, sometimes they are organisations such as firms or public organisations. In the latter case, individuals who obtain services or products from firms or citizens who rely on public organisations are also likely to share in the benefits which those organisations derive from having access to data, or from sharing it. Nonetheless, we think a useful distinction can be drawn between data sharing which benefits a specific individual – and which is therefore generally initiated by that same individual – and data sharing which has broader benefits and which is generally initiated by an organisation. The data that is accessed at the initiative of a user is generally data about that specific user and is generally required to provide services for their benefit²⁰. Data that is accessed at the initiative of an organisation may involve data that has been acquired from a large number of users and subsequently anonymised or it may be non-personal data that has been derived from other sources (including being created by the organisation itself). In these cases, the data is required to provide services from which a large number of users, rather than any particular user, are likely to benefit²¹.

2.1 Incentives of the individual users to share their data

In Europe (and increasingly in other parts of the world), individual users are considered to have certain rights over data that they have themselves provided to organisations, including both firms and public bodies. These rights arise from the conviction that citizens and consumers should exercise control over data that relates specifically to them (and by which they can be identified), even if they have chosen to share it with other parties. Under the GDPR, control involves rights to require the deletion of data which has previously been provided or which has been observed or inferred from their interactions with the organisation in question, but also rights to require the return of data back to the same user (irrespective of whether or not the data is subsequently deleted by organisation that held it) in a standardised format and, 'where technically feasible', the direct transfer or 'porting' of the data to another organisation. We note that the first review of the implementation of the GDPR

¹⁸ See Martens et al. (2020).

¹⁹ Krämer et al. (2020), Section 2.

²⁰ Cremer et al (2019:25)

²¹ Cremer et al (2019:25) refer to this as 'anonymous use of individual-level data', 'aggregated data' and 'contextual data'.

(after 2 years) found that the application of the right to port individual user data by users had been very limited²².

We have identified several reasons why an individual user might wish to obtain access to personal data and/or to share it:

- **An individual user may require access to the data that has been accumulated by the platform about him to better understand what it is** (and perhaps subsequently to make changes to what data is collected from them in the future). This may be particularly important with observed data that is acquired and accumulated every time the user engages with a digital platform, such as data about search terms that have been entered or web sites that have been visited. Data access may therefore serve to reduce the information asymmetries between users and digital platforms and, potentially, ensure that users can then alter their privacy settings to ensure that the data disclosed and retained is the minimum necessary to meet their requirements. The OECD (2019:43) refers to this motivation as 'informational self-determination'. The early data access arrangements that were developed by digital platforms like Google ('Google Take Out') and Facebook appear to have been primarily intended to serve this purpose, allowing data to be provided to the user on request but not, at least easily, to be transferred to another organisation. Evidence of user demand for such capabilities is not publicly disclosed by either Google or Facebook (in contrast, Google does disclose the demands it receives from Governments and other organisations for their users' data²³), although Google referred to 'millions of users' having downloaded their data in 2016²⁴. It is reported that user demand for access to data held by Facebook increased significantly following the revelations in 2018 that Facebook had shared user data from over 50 million accounts with Cambridge Analytica²⁵.
- **An individual user may require access to the data held about her because she wishes to share it with another service provider.** As noted earlier, this could be done either as a single transfer or continuously into the future. The former could be the case if a user wishes to switch from one provider to another (in a 'single homing' environment) or if the data is confined to the provided data. The latter is more likely to be appropriate if a user wishes to 'multi-home' with several providers simultaneously²⁶ and if the data included observed data. In the former case, the individual user is likely to be switching because they perceive some benefit for themselves in doing so. It may be possible for them to switch without obtaining access to their data from the existing platform, but they may then incur costs in terms of time and effort in replicating the data for their new provider. If the data is observed data, then it will take a new platform time to accumulate the data (and the insights it derives from it), and the service the user receives from the new provider until sufficient data is accumulated may be inferior. Some users may not be willing or able to switch without access to the data, and so would be deprived of the service altogether. Individual users might, in theory, be prepared to download data from one platform and upload it to another,

²² 'The Staff working Document accompanying the Communication (COM (2020) 264) found 'The right to data portability is not used to its full potential. The European Strategy for Data (hereafter Data Strategy), adopted by the Commission on 19 February 2020, emphasised the need to facilitate all possible uses of this right (e.g. by mandating technical interfaces and machine-readable formats allowing portability of data in (near-to) real-time). Operators note that there are sometimes difficulties in providing the data in a structured, commonly used machine-readable format (due to the lack of standard). Only organisations in particular sectors, such as banking, telecommunications, water and heating meters, report having implemented the necessary interfaces. New technological tools have been developed to facilitate the exercise by individuals of their rights under the GDPR, not limited to data portability (e.g. personal data spaces and personal information management services).', SWD (2020) 115

²³ <https://transparencyreport.google.com/user-data/overview>

²⁴ House of Lords (2016), para 245

²⁵ <https://www.vox.com/2018/11/20/18105541/facebook-user-data-request-download-delays-high-volume>

²⁶ For example, the average UK user multi-homed across 7.1 social media platforms in 2019, see <https://wearesocial.com/uk/blog/2019/03/digital-in-the-uk-data-and-learning-for-2019>

as the GDPR appears to contemplate²⁷. But in practice, the sharing of data on any scale is likely to require the user (or its agent) to initiate a direct and continuous transfer of data from one platform to another using APIs. There are several examples of initiatives that are intended to help individual users share 'their' data in this way:

- **Several Government initiatives have sought to facilitate the sharing of data between providers of utility, financial or medical services to facilitate switching in markets where users tend to single-home.** In Europe, initiatives such as the UK Government's Midata project, launched in 2011 or the French Mesinfos programme²⁸ have sought to promote data access voluntarily. More successful have been the so-called Green and Blue Button programmes in the United States, launched in 2010 and 2012 respectively, which have allowed, in the Green Button case, over 60 million households to download information about their energy consumption from 150 energy providers (and to facilitate data access for third parties) and, in the Blue Button case, 150 million persons to access their medical records from 16,00 organisations²⁹.
 - **Several private sector initiatives to facilitate the transfer of data.** The most relevant of these for our purposes is the Data Transfer Project (DTP) in which Google, Facebook, Apple, Microsoft, and several other (mainly large) digital platforms are currently developing a set of 'adaptors' which would enable the transfer of user data via the proprietary APIs of each platform from one to another. The DTP was founded in 2018 and is still under 'active development'. One of the challenges is, as the DTP notes, the formats for data in different 'verticals' (emails, photos, or music each represent a different vertical in this context) 'have emerged organically in a largely disconnected ecosystem'³⁰. This means that each organisation may use a different proprietary 'data model' to organise the data it acquires, as well as proprietary APIs which already allow that data to be shared with third party developers. The DTP aims to address these issues by developing 'adaptors' which allow data to be converted from one proprietary model to another and by encouraging the future development and adoption of common data models in the future³¹. It is unclear at this stage when the full DTP functionality will be made available to users, nor what it might comprise of, although we know Facebook has adopted DTP functionality to enable the sharing of photos with Google³².
- **Individual users may require access to data about them because the sharing of that data may allow them to benefit from 'complementary innovation' in other (data-dependent) markets.** In this case, the user does not seek access to data to substitute one service for another of a similar kind but to obtain a new service. This may arise if data from a single organisation is shared with one or more third parties, but it may also arise if the data from one or more organisations with whom that user interacts is then shared with a

²⁷ In addition, Article 16 of the Content Services Directive 2019 (2019/770) allows users to request a digital platform to provide any non-personal digital content which they created or provided whilst being provided with services upon the termination of the relationship with the service provider. However, there is no obligation for the platform in question to facilitate the transfer of this data to another provider.

²⁸ Ctrl-Shift (2018), p.30

²⁹ OECD (2019), p.126, Ctrl-Shift (2018), p.82-3

³⁰ <https://datatransferproject.dev/documentation>

³¹ See Gal and Rubinfeld (2019). The authors suggest that interim solutions, such as 'data translators' which are 'algorithms, which relate the data attributes of one data set to those of another data set, can (partly) solve some of the data integration problems outlined above while significantly reducing intervention in the choices of market player', p.766

³² Kramer, Senellart and de Streel (2020) report: 'A number of import or export connectors have been implemented to interface with various platforms, but there are very few public-facing sites that do use the DTP (the most prominent being a specific use case on Facebook: users have been able very recently to use it to transfer their photos to Google Photos', p.47. On Facebook's most recent data portability plans, see Facebook (2020)

single third party that is then able to combine or aggregate the data in ways which the donors cannot:

- **Some significant complementary innovation initiatives have been undertaken in the financial and payment services industry in recent years.** One of the most advanced of these is the *Open Banking initiative* in the United Kingdom, where a common set of APIs has been developed under the guidance of a new regulatory body (the Open Banking Implementation Entity) to allow third-party financial service providers to obtain access, at the initiative of the user, to the banking data of accounts held by that user.³³ The model is intended to allow third parties to develop and offer new types of services, such as advisory or recommendation services (including price comparison websites) which analyse user data to produce recommendations for appropriate services or ways in which users may reduce their charges or otherwise improve their financial health, or for third parties to themselves transfer funds between different accounts on the users' behalf. The implementation of Open Banking is expected to proceed in stages, with access to new types of data becoming available for transfer (e.g. pensions, investments, and other accounts) and new and more complex payment functions being enabled over time, although some banks have taken the opportunity to supplement data which they supply to fulfil regulatory obligations with 'premium APIs' which they offer on a purely commercial basis³⁴. The Open Banking APIs were launched in 2018 and currently has around 2 million monthly users, with the UK Government also looking to adopt a similar approach to other sectors under the 'Smart Data' initiative³⁵. The European Commission adopted the *Second Payment Services Directive* in 2015, which introduces similar provisions to Open Banking for the rest of Europe³⁶. Similarly, the Australian Government introduced a 'Consumer Data Right' (CDR) under which, initially, accredited financial service providers will have access to both product data and, with the user's consent, account data held by authorised deposit takers. Regulations for financial services have been introduced and the first transfers are expected to be undertaken in late 2020³⁷. Standards for APIs are being developed by the Data Standards Body, some of which are expected to be applicable for data sharing more widely as the Consumer Data Right is extended to other parts of the economy in the future.
- **Data access and interoperability provisions have also been adopted at the European level to facilitate access to data about household energy consumption in smart electricity³⁸ and gas meters³⁹,** which are currently being deployed in energy networks throughout Europe. The Commission had initially expected around 80% of European households to have a smart electricity meter and smart gas meter by 2020. A 2019 report estimated that the Member States would achieve closer to 40% penetration of smart electricity and 23% of smart gas meters

³³ See CMA Final Report of 9 August 2016 on the Retail Banking Investigation and CMA, pp. 441-460 and CMA Order of 2 February 2017 on the Retail Banking Investigation, Sect. 10 to 14 and the Associated Explanatory Note, paras.28-39. All documents are available at: <https://www.gov.uk/cma-cases/review-of-banking-for-small-and-medium-sized-businesses-smes-in-the-uk>.

³⁴ Fingleton/ODI (2019) propose that the UK Open Banking programme supports the development of standard APIs which banks could use to offer 'premium services' for which they could charge, p.36. For a current example, see

³⁵ UK Government, Smart Data: putting consumers in control of their data and enabling innovation', June 2019 and 'Next Steps for Smart Data', September 2020

³⁶ Directive 2015/2366 of the European Parliament and of the Council of 25 November 2015 on payment services in the internal market, OJ [2015] L 337/35.

³⁷ <https://www.accc.gov.au/focus-areas/consumer-data-right-cdr-0>

³⁸ Directive 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27, OJ [2019] L 158/125, Art 23 and 24

³⁹ Directive 2009/73 of the European Parliament and of the Council of 13 July 2009 concerning common rules for the internal market in natural gas, OJ [2009] L 211/94, amended by Regulation 2018/1999 and Directive 2019/692.

by that date⁴⁰. Implementation of smart metering projects has been beset by various delays, most of which relate to the development of the accompanying metering infrastructure, technical standards, and the willingness of households to adopt the technology. The Commission has taken an increasingly prominent role in specifying the functional requirements of smart meters, standardised communications interfaces, security, and data protection considerations⁴¹. The cost benefit analyses undertaken by the Member States⁴² have attached much greater weight to changes in household consumption patterns arising from the direct feedback that the meter provides, cost savings for suppliers who no longer need to read meters at premises, and revenue protection from obtaining accurate readings. There is, so far as we are aware, no research to date on either the volume of user data that has been transferred to third parties as a result of the implementation of smart meters or its impact on innovation or competition in the provision of energy-related services such as smart appliances, 'smart home' management services or other forms of complementary innovation which might be enabled by the sharing of data. It seems likely that further action will be required – as evidenced by the UK Government's and energy regulator's efforts to extend the Midata programme to the energy sector and require energy suppliers to provide access to data via standardised APIs⁴³ – if the potential for data access to promote competition in the energy sector is to be realised. This would include addressing many of the same issues relating to third party data access that we identify later in this report.

- **Individual users may seek access to their data in the expectation that they will be able to monetise it and generate an income (even if they have not been able to do so to date).** This is an idea often associated with 'Personal Data Stores' (PDS) or Privacy Information Management Systems (PIMS) which operate as intermediaries between users who control their data and digital platforms and other organisations who derive value from being able to access it and who may therefore be willing to pay to do so. The development of such organisations, and users' motivations for initiating a transfer of data to them, has been rather mixed to date⁴⁴. Some PDS providers emphasise that users benefit from enhancing security and protection of their data if it is held by them and provided, at the users' request, to other organisations as and when it is appropriate to do so rather than being held in a central repository by a large digital platform, which may be vulnerable to hacking or may be regarded as having poor data governance or inappropriate incentives. Others refer to the convenience that arises from avoiding having to re-enter data to multiple platforms, or the same platform on multiple occasions. Users may also benefit from arrangements that allow a third party to obtain relevant information to enable the provision of services without disclosure of sensitive information. Other providers emphasise the opportunity for users to monetise their data by 'selling' it back to the platform⁴⁵. On this view, users themselves may be unable or unwilling to engage in market transactions involving their data with digital platforms, given information asymmetries (how much is the data worth?) and transactions costs (whom to negotiate with or how to obtain payment?), but intermediaries can do so on their behalf and earn a commission in the process. The market for PDSs remains at an early stage of development and they face the familiar scaling

⁴⁰ Tractebel (2019), p.9

⁴¹ Ibid, p.18

⁴² Ibid, p. 34/5. This is mainly so for electricity, since many Member States are yet to undertake a CBA for smart gas meters

⁴³ Department for Business, Energy & Industrial Strategy (2018)

⁴⁴ See, for example, the study by the Judge Business School (2015). This study also hypothesizes benefits for businesses as well as users if PDSs allow them to reduce compliance costs and notes that some PDS generate revenues from payments received from businesses rather than commissions from users.

⁴⁵ Examples of such providers include tide.org and swashapp.io

challenges of any business engaged in a two-sided platform market. Adoption by users to date appears to be very low. A study for the European Commission found that the potential market for PIMS/PDS services in Europe depends greatly on the proportion of transfers that users would be willing to authorise via a PIMS, with the resulting range (between €1 billion p.a. and €90 billion p.a.) reflecting a very high level of uncertainty about its prospects.⁴⁶ Some observers have concluded that the PIMS market would be unlikely to develop without greater Government involvement to coordinate the development of common standards⁴⁷.

The evidence we present here shows that some individual users have perceived tangible benefits from being able to share data, with the data shared invariably being personal data that requires the user themselves to initiate the process. However, whilst there is little or no empirical evidence as to the volumes of data that is shared in this way in Europe today, the evidence points to it being very low. The downloading of data from digital platforms by individual users appears to be confined to a small minority and the monetisation of personal data by individuals remains more of an aspiration than a reality, many years after it was first proposed. Some large government initiatives, such as the US Green Button programme have been widely adopted by users, but others, such as the Midata and European smart meter programme have fallen well short of expectations or been frustrated by the organisations that were involved in delivering them. This suggests that barriers to sharing, such as concerns about privacy or security which we discuss later, may cause individual users to decline to authorise access to data, even when it seems clear they would otherwise stand to benefit from doing so⁴⁸.

We note that one study for the UK Government concluded that data sharing to facilitate switching between service providers did not appear to be particularly attractive for users (which may account for what appears to be relatively low levels of use of existing services such as Google Takeout and the limited progress of other Government projects to promote switching). The same report concluded that data sharing to enable complementary innovation was of much greater interest to users (although it seems to us that significant investments in user education would be required for these benefits to be properly understood)⁴⁹. To date, the most promising recent application of data sharing appears to be the Open Banking initiative in the UK, although adoption remains at a relatively early stage.

With this in mind, we now consider arrangements for the provision of access to data that do not involve users but instead are undertaken for the benefit of and at the request of other organisations.

2.2 Incentives of firms to share aggregated user data

Although our focus in this report is with access to data under conditions in which firms are unlikely to volunteer to share data with potential competitors, it is important to acknowledge that there are many instances in which organisations do share data voluntarily.⁵⁰ Examples include:

- **Commercial organisations exchange data when there is mutual commercial benefit from doing so.** The nature of these benefits should not include restrictions on competition or co-ordinated conduct, and so data sharing or pooling arrangements amongst firms who

⁴⁶ Judge Business School (2015), p. 35

⁴⁷ Ctrl-Shift(2018), p.7. Judge Business School (2015) came to a similar conclusion, p.24, as do Kramer, Senellart and de Streeel (2020), p.66-72

⁴⁸ We refer to other cases of 'data philanthropy' later in this section, but such philanthropy is generally undertaken by organisations rather than by individual users

⁴⁹ Ctrl-Shift (2018), 'However, the net benefits from government encouraging mobility as a means of supplier switching (following the model of the current account switch service) were assessed in this economic analysis as low. Such an intervention would be unlikely to address the fundamental behavioural issues underlying a lack of switching; cognitive limitations prevent consumers from recognising the value of switching'. p.43

⁵⁰ See de Streeel and Tombal (2020), Expert Group for the Observatory on the Online Platform Economy (2020), Everis (2018), IDC and Lisbon Council (2020), OECD (2020).

would otherwise be competitors are often subject to review by competition authorities⁵¹. Sometimes commercial organisations in the same sector share or pool data on a bi- or multi-lateral basis, without payments flowing between them, as when data is shared for credit reference purposes or to detect fraud (e.g. with insurance claims)⁵². Other examples include Nallian, a third-party platform that facilitates the sharing and analysis of data by suppliers of air freight services⁵³, and agrirouter, a platform that was intended to facilitate data for precision farming in Germany and elsewhere⁵⁴.

- **Commercial organisations may provide unilateral access to data, generally on a restricted basis via APIs, to promote complementary innovation and grow digital 'eco-systems'**. This strategy has been extensively studied and is commonly pursued by digital platforms, such as Facebook, Twitter, Google or Apple, all of whom provide applications developers with access to data to facilitate innovation which increases the value of the platform to users⁵⁵ and/or which enable the platform to augment their own 'first party' data with 'third party data' which they obtain via these applications⁵⁶. In such circumstances, the data that can be accessed and the terms on which access is offered are determined by the platform itself and can vary over time as the platform's strategy evolves and/or it internalises some applications into its own business (often through the acquisition of other service developers).
- **Commercial organisations may provide access to data in return for 'ancillary services'**. Krämer, Schnurr, and Broughton Micova (2020) also explain how large platforms such as Google and Facebook also offer 'ancillary services' such identity management services ('Login with Facebook'), payments services, or tracking technology ('Google Analytics') to other digital services providers. This may be a valuable service for those providers and their users, but it also requires the sharing of valuable data (about user engagement and purchase history in the wider digital environment) with the large platform that provides the service⁵⁷.
- **Commercial organisations may provide access to data in return for payment.** Data brokers such as Bloomberg, Oracle, Acxiom, and Nielsen acquire and aggregate huge volumes of data from a wide range of sources, some public and some private, and sell data sets to other firms for marketing and other purposes, such as the tracing of persons, verification of identity or credit checking⁵⁸. Data may include personal details, tax, and court records, as well as records of purchases and other profiling data, including location data, which is relevant to those who engage in targeted advertising. Customers will often combine and match such data with other data that they have themselves acquired directly from users. Brokers generally compete on the accuracy and quality of the data they supply, its scope, and price. Data broking markets are expected to be worth over \$10 billion by 2022⁵⁹, but have remained largely unscrutinised to date. Data market providers also provide commercial

⁵¹ See Lundqvist (2018).

⁵² Data pools of this kind might be viewed as horizontal arrangements. The sharing of a users' debt repayment history amongst organisations might not be considered beneficial for the individual concerned (who may be denied further credit as a result) but is generally considered by competition authorities to be justified on the grounds that it removes the information asymmetry that otherwise exists between a user and potential lender, and prevents other users from unfairly subsidising bad debtors. There are also many vertical sharing arrangements as when data is transferred between airlines and suppliers of aero engines in order to improve diagnostics, performance management or predictive maintenance, or data is transferred from farmers to suppliers of agricultural machinery. A large number of IoT services are expected to involve the sharing of data between those using the technologies and those responsible for its supply.

⁵³ <https://www.nallian.com/>

⁵⁴ <https://my-agrirouter.com/en/>

⁵⁵ <https://developers.google.com/products/develop>; <https://developers.facebook.com/>; <https://developer.twitter.com/en>

⁵⁶ Krämer, Schnurr and Broughton Micova (2020), p.55

⁵⁷ Ibid p.70

⁵⁸ Federal Trade Commission (2014)

⁵⁹ See Ram and Murgia (2019)

platforms that allow firms to trade their data, without the data market itself engaging in the acquisition or supply of data⁶⁰. Such providers often specialise in the type of data that is traded, or the types of firms engaged in doing so.

- **Rather than share underlying data, organisations may restrict access to the proprietary data they hold by using it to respond to specific queries in return for payment.** Microsoft provides syndicated search services to third parties such as Yahoo, Ecosia, and DuckDuckGo which produce results (and serve adverts) in response to user queries which third parties can then re-present⁶¹. Telecoms providers such as Telefonica and Orange have developed commercial services that use their location and other data to address the queries of third parties without sharing the underlying data itself⁶², as has Uber⁶³. Such arrangements may be motivated by the wish to monetise data without ceding the competitive advantage to rivals⁶⁴ (as might occur if access were provided to the underlying data), but they might also allow the provider to obtain benefits for its services by realising economies of scale from the additional user interactions. Such arrangements might also arise because of privacy concerns, including the risk that anonymised data sets might subsequently be reconfigured to reveal personal identities. To address this, several firms provide 'trusted intermediary' or 'sandbox' services (a kind of PDS for business) which are intended to allow private organisations to monetise their data assets securely and without sharing the underlying data⁶⁵.
- **Commercial organisations may engage in 'data philanthropy'**, where they provide access to data, normally to specific charitable or public organisations, to enable them to pursue particular social or economic objectives. Mastercard runs such a programme⁶⁶. A significant number of digital platforms and telecommunications operators are providing location data to health authorities during the COVID crisis and have previously provided aggregated data about population movement and other aspects of user behaviour to statistical offices and planning authorities⁶⁷.
- **Public organisations account for a significant proportion of data sharing initiatives**⁶⁸. The OECD estimate that 65% of all data sharing initiatives were undertaken by public sector organisations, either to enable sharing with other public organisations or to allow sharing with private organisations. Examples include the sharing of geospatial (maps) or transport data to enable complementary innovation or spillovers, such as the development of new smartphone apps⁶⁹, but also the sharing of health and other scientific data on a more restricted basis (often in research repositories) between public institutions (often across borders) in the pursuit of collaborative research⁷⁰. In the former case, data is completely 'open', unencumbered by intellectual property or other rights, and available to any person or organisation to download and use without requiring any kinds of consents (and often free of

⁶⁰ An example is the French firm, DAWEX, see <https://www.dawex.com/en/> or Whoapi, see <https://whoapi.com/>

⁶¹ <https://about.ads.microsoft.com/en-gb/resources/training/syndicated-partner-network>

⁶² OECD p.39

⁶³ <https://movement.uber.com/?lang=en-GB>

⁶⁴ In the Microsoft case, the syndication revenues are likely to be less significant than the additional search query data which Microsoft can then use to improve its search algorithm.

⁶⁵ As example of a firm providing such services is Aircloak, see <https://aircloak.com/solutions/how-it-works/>

⁶⁶ OECD (2019), p.48

⁶⁷ On Covid, see European Parliament Briefing (2020).

⁶⁸ Although, as the OECD (2019:28) note, the distinction between 'public data' and data held by the private sector is not always clear-cut.

⁶⁹ See, for example, OECD p.66: 'a major part of the benefits of open data by TfL [Transport for London] were realised thanks to the development of apps that used TfL open data to provide real-time traffic information for more accurate navigation systems (Table 3.1). More than 80 data feeds were made available for developers through a free unified application programming interface (API),15 which ensured accurate real-time data for over 13 000 registered developers and more than 600 apps. This generated a gross value added of GBP 12 million to GBP 15 million per year for businesses and led to the direct creation of more than 500 jobs and more than 230 indirect jobs across the supply chains and the wider London economy'

⁷⁰ OECD (2019), p.36

charge). In the latter case, data may be sensitive and subject to significant access controls. The European Union has adopted the Open Data Directive which promotes the sharing and re-use of data held by a wide range of public bodies⁷¹.

2.3 Barriers to data sharing and market failures

Despite the many motivations that organisations have for sharing data and the many benefits that have been identified from doing so, a common theme from the literature is that, as with the sharing of data for the benefit of specific individuals, the economic and social benefits from sharing large volumes of aggregated data remain unrealised relative to its potential, both with the public and with the private sector and both in Europe and elsewhere in the world⁷².

There seems to be no doubt that, in an increasingly data driven economy, the potential gains from sharing data are large. The OECD finds:

'Evidence shows that data access and sharing can generate positive social and economic benefits for data providers (direct impact), their suppliers and data users (indirect impact), and the wider economy (induced impact)....Recently available studies by sector (public vs. private sector) further discussed below provide a rough estimate of the magnitude of the relative effects of data access and sharing. They suggest that data access and sharing can increase the value of data to holders (direct impact), but it can help create 10 to 20 times more value for data users (indirect impact), and 20 to 50 times more value for the wider economy (induced impact). In some cases, however, data access and sharing may also reduce the producer surplus of data holders. Overall, these studies suggest that data access and sharing can help generate social and economic benefits worth between 0.1% and 1.5% of gross domestic product (GDP) in the case of public-sector data, and between 1% and 2.5% of GDP (in few studies up to 4% of GDP) when also including private-sector data'⁷³

The inability of users and organisations to fully realise benefits from data sharing and the barriers that exist to their doing so, ought to yield important insights for any policymaker or regulator that is proposing to rely upon data access to promote market contestability. Some of these challenges may not be specific to data sharing, such as the failure of markets to discover prices at which data can be traded efficiently or the presence of externalities which neither party can capture but which represent a gain for society. Others may be specific to the activity of sharing data, although the degree of risk may also depend on the type of data that is shared and how data is shared. We focus primarily on the supply-side barriers to the sharing of data but recognise that there may also be demand-side issues which mean that data which is available nonetheless remains underutilised or underexploited. Challenges include:⁷⁴

- **Co-ordination issues, such as the agreement on common standards and infrastructure to facilitate the transfer of data between organisations** which are diverse and which may not have (indeed, can be expected not to have when data sharing is

⁷¹ Directive 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information, OJ [2019] L 172/56. For further details of data sharing initiatives in the public sector see OECD (2019), p.117-121

⁷² Deloitte (2016) observe 'These are markets that are still in their infancy, i.e. what is known as an 'emergence phase'. To be entirely active in these markets, EU companies need to be intensive data users, but that is the case of only 6.3% according to a study for the European Commission. The fact that most companies have not yet engaged with these markets has been borne out by the qualitative assessment of the business models of more than 100 European firms as part of this study. Most companies have not yet completely integrated these new realities into their business models and approaches. But for the small number of companies which are currently proactively engaged in the data economy, there are genuine uncertainties and barriers to them moving forward, and which may well be acting as deterrents to companies want to enter the market.'

⁷³ OECD (2019) p. 60

⁷⁴ See Deloitte, et al. (2018) and Martens et al. (2020) for the market failures in B2B data sharing.

being used to promote competition) incentives or objectives that are well aligned. Recognising this, most researchers agree that some form of 'co-ordinating entity' or regulator is required to oversee the development and implementation of new data access arrangements⁷⁵. Much of the rest of the report is devoted to the consideration of the issues and challenges which such a body will need to address.


- **Organisations that provide third parties with access to their data may expose themselves to a variety of legal risks.** This is self-evident with data which, under the GDPR, is considered to be personal data and for which user consent is required before disclosure. As we explain later, legal uncertainty may also arise with other proprietary data which may be subject to various contractual safeguards or over which third parties may also claim rights. One of the features of data is that its legal status, and the rights of those in possession of it, is often unclear, at least in comparison with many tangible assets.
- **Even if an organisation considers that it has legal authority to share data with third parties – or is obliged to do so – it may be concerned about reputational risks or fines that may arise from the uses to which it is put by others.** Risks may arise because a third party uses the data in an unanticipated way or in breach of its contract, as appears to have been the case when Cambridge Analytica was given access to data acquired by Facebook in 2018 for 'academic research purposes' but was subsequently found to have been using it to support its political consultancy business which was assisting campaigns in the United States and, allegedly, other countries⁷⁶. They may also arise because a third party is using the data for legitimate purposes, but is then subject to a security breach or cyberattack which results in the unlawful disclosure of data. A third-party may also sell data to another party despite contract provisions which may seek to limit their rights to do so. Alternatively, a third-party may fail to delete personal data once it no longer has legitimate grounds for retaining it. An inevitable consequence of providing third-party access to data is that, unless the terms are highly restricted (e.g. to the provision of responses to queries), the data is distributed amongst organisations, each of which is potentially vulnerable to security breaches or mismanagement of the data. Organisations may use contracts to avoid legal liability in such circumstances, but it is difficult to avoid the reputational consequences of errors. Facebook's market value fell by 17% following the revelations about Cambridge Analytica.

Although different organisations are subject to different motivations, most will only share data if they perceive the potential benefits of doing so to exceed the kinds of costs and risks outlined above. Many studies of the barriers to data sharing have found that the safety and security of data is the primary concern for organisations (as well as for individual users)⁷⁷. In some cases, the result is that data is not shared and there is under-provision of access to data in the economy. In other cases, organisations will adopt strategies to minimise risks or costs, or to limit them sufficiently to ensure that they are exceeded by the benefits of sharing. One means of doing this is to limit or reduce the scope of the data to be shared to avoid the legal risks arising from the GDPR or similar legislation elsewhere in the world. This is often done by anonymising data so that it can no longer be associated with an identifiable individual. In response, new technologies and techniques are constantly being

⁷⁵ For example, Furman et al (2019) propose the establishment of a new 'Digital Markets Unit' to, amongst other things, oversee the implementation of data access arrangements in the UK. An earlier study by Ctrl-Shift came to similar conclusions, Ctrl-Shift (2018). See also Gal and Rubenstein (2019).

⁷⁶ Cadwallar, 'The Cambridge Analytica Files', The Guardian, available at <https://www.theguardian.com/news/series/cambridge-analytica-files>

⁷⁷ Ctrl-Shift (2018) p.3.



developed to allow for the recreation or recombination of the data to be able to extract the identities of individuals and, thereby, greatly enhance its commercial value⁷⁸.

Alternatively, as noted earlier, organisations and individual users may retain control of the underlying data but offer answers to queries from third parties by interrogating the data on their behalf. Or they may use trusted intermediaries, or intermediary infrastructure, to safeguard their data whilst obtaining the benefits of sharing it. The PDS discussed earlier offer such a service for individual users, whilst 'sandboxes' may allow organisations to share data with intermediaries who then manage the risks associated with sharing or onward transfer on their behalf.

In addition to restricting the scope of the data to which access is granted, organisations may limit who they are prepared to share the data with. They may be organisations with a common set of interests, as in the case of data shared amongst public health institutions or credit reference agencies, or they may be organisations that meet other criteria. To meet this need, initiatives such as the Industrial Data Space have been developed to provide accreditation and licensing services for organisations who wish to share data⁷⁹. As we discuss below, most of the regulatory initiatives to promote data sharing, such as the Open Banking initiative or the Australian Customer Data Right initiative, limit access to organisations that have been accredited and whose activities are overseen by some form of regulatory body.

Since the sharing of data, whether at the initiative of users themselves or organisations, remains relatively underdeveloped in relation to the potential benefits which we might expect to obtain from it, there are many questions which have yet to be resolved. Before turning to those, we discuss the various ways in which regulation has been used to promote data sharing in Europe to date - and what conclusions might be drawn from them.

⁷⁸ OECD (2019) p.29

⁷⁹ Fraunhofer (2017)

03

**COMPETITION
POLICY AND DATA
SHARING**

3 Competition policy and data sharing⁸⁰

Competition law imposes data sharing obligation under very specific conditions, either when the refusal to give access to the data could be considered as an abuse of dominant position under the so-called essential facilities doctrine or as a remedy to a merger between two data-rich firms that could significantly impede effective competition. Conversely, competition law may also limit data sharing, either when data sharing would amount to an anti-competitive agreement or when data siloing is imposed as remedy to a merger between two data-rich firms.

3.1 Compulsory access under Article 102 TFEU

3.1.1 Relevant case-law

Among the main essential facilities cases decided by the Court of Justice of the European Union, three are related to data and information. The first case is **Magill**. In this case, Radio Telefis Eireann (RTE) had a statutory monopoly over television broadcasting in Ireland and BBC and IBA a statutory duopoly in the UK (thus including Northern Ireland). RTE and BBC owned the copyright in their programme listing (thus some data) for their respective channels and ITP the copyright for the listings of IBA. Each of them published a weekly guide for their own programmes but none of them published a comprehensive weekly guide with the programmes of the three channels for Ireland. They refused to give a licence to Magill which was willing to publish such a comprehensive guide.

Upon complaint, the Commission⁸¹ condemned the channels for abuse of dominant position. This was confirmed by the General Court.⁸² On appeal, the Court of Justice confirmed again the Commission decision and judged that, although a refusal to grant a licence in respect of an intellectual property right cannot in itself constitute an abuse of a dominant position, the *'exercise of an exclusive right by the proprietor may, in exceptional circumstances, involve abusive conduct.'*⁸³ The Court of Justice identified three conditions for those circumstances:

- The dominant firms *'reserve to themselves the secondary market of weekly television guides by excluding all competition in that market (...) since they denied access to the basic information which is the raw material indispensable for the compilation of such a guide.'*;
- The refusal to provide basic information by relying on national copyright provisions prevented the appearance of a new product, a comprehensive weekly guide to television programmes, which the channels did not offer and for which there was a potential consumer demand;
- There was *'no justification for such refusal either in the activity of television broadcasting or in that of publishing television magazines'*.⁸⁴

The second case is **IMS-Health**. In this case, IMS-Health collected pharmaceutical sales information from wholesalers in Germany, structured them with the so-called 1860 brick structure (linked to the German postal codes) developed with pharmaceutical companies and then provided sales reports to those pharmaceutical firms. IMS-Health had an intellectual property right on the 1860 brick structure and refused to licence it to NDC-Health which wanted to compete on the downstream pharma sales reports. Upon complaint by NDC-Health, the Commission had ordered interim measures forcing IMS to licence its brick structure that was found indispensable to carrying on business in the downstream

⁸⁰ This section is partly based on Graef, Tombal and de Stree (2019).

⁸¹ Decision of the Commission of 21 December 1988, Case IV/31.851, *Magill TV Guide v. ITP, BBC and RTE*.

⁸² Case T-69/89 *RTE v. Commission*, EU:T:1991:39; Case T-76/89 *ITP v. Commission*, ECLI:EU:T:1991:41.

⁸³ Joint Cases C-241/91P et C-242/91P, *Radio Telefis Eireann (RTE) and Independent Television Publications (ITP) v. Commission*, EU:C:1995:98, para 50.

⁸⁴ Respectively, paras 56, 52 and 55 of the Case.

market.⁸⁵ In the meantime, a litigation took place before a German Court which made a preliminary reference to the Court of Justice.

In its reply, the Court of Justice decided that the refusal to licence an intellectual property right constitutes an abuse of a dominant position where the following conditions are fulfilled:

- *'the refusal is such as to reserve to the owner of the intellectual property, the right to market for the supply of data on sales of pharmaceutical products in the Member State concerned by eliminating all competition on that market;*
- *the undertaking which requested the licence intends to offer, on the market for the supply of the data in question, new products or services not offered by the owner of the intellectual property right and for which there is a potential consumer demand;*
- *the refusal is not justified by objective considerations.*⁸⁶

Moreover, the Court of Justice decided that: *'the degree of participation by users in the development of that structure and the outlay, particularly in terms of cost, on the part of potential users in order to purchase studies on regional sales of pharmaceutical products presented on the basis of an alternative structure are factors which must be taken into consideration in order to determine whether the protected structure is indispensable to the marketing of studies of that kind.*⁸⁷

The third case is **Microsoft**. In this case, Microsoft had a near monopoly on the PC operating system market and was providing interoperability information to the producers of workgroup servers.⁸⁸ However, when Microsoft decided to enter the workgroup server market, it stopped giving interoperability information. Upon complaint of Sun Microsystems, a workgroup server producer, the Commission condemned Microsoft and forced it to resume the provision of interoperability information.⁸⁹ On appeal by Microsoft, the General Court confirmed the Commission's decision and summarised the case-law in the following way: ⁹⁰

331. It follows from the case-law cited above that the refusal by an undertaking holding a dominant position to license a third party to use a product covered by an intellectual property right cannot in itself constitute an abuse of a dominant position within the meaning of [Article 102 TFEU]. It is only in exceptional circumstances that the exercise of the exclusive right by the owner of the intellectual property right may give rise to such an abuse.

332. It also follows from that case-law that the following circumstances, in particular, must be considered to be exceptional:

- *in the first place, the refusal relates to a product or service indispensable to the exercise of a particular activity on a neighbouring market;*
- *in the second place, the refusal is of such a kind as to exclude any effective competition on that neighbouring market;*
- *in the third place, the refusal prevents the appearance of a new product for which there is potential consumer demand.*

⁸⁵ Case 38 044.

⁸⁶ Case C-418/01, *IMS Health v. NDC Health*, EU:C:2004:257, para 52 with a re-ordering of the conditions.

⁸⁷ Para 30 of the Case.

⁸⁸ Work group server operating systems are operating systems running on central network computers that provide services to office workers around the world in their day-to-day work such as file and printer sharing, security and user identity management.

⁸⁹ Decision of the Commission of 24 March 2004, Case 37.792 *Microsoft*.

⁹⁰ Case T-201/04, *Microsoft v. Commission*, EU:T:2007:289.

333. Once it is established that such circumstances are present, the refusal by the holder of a dominant position to grant a licence may infringe [Article 102 TFEU] unless the refusal is objectively justified.

334. The Court notes that the circumstance that the refusal prevents the appearance of a new product for which there is potential consumer demand is found only in the case-law on the exercise of an intellectual property right.

Thus, in **Magill**, the Court of Justice validated the compulsory access to programme listings, data for which there was a legal barrier (the copyright), and which was a by-product of the main activities of the broadcasters. In **IMS-Health**, the Court of Justice set the conditions to impose access to a structure for data which was a *de facto* industry standard. In **Microsoft**, the General Court validated the compulsory access to interoperability information which was also close to *de facto* industry standard.

Next to those EU cases, two non-digital national cases, which are very similar, are interesting. In both cases, a **firm uses a customer list developed when it enjoyed a legal monopoly to promote a new service allowing it to compete unfairly through data cross-subsidisation which "un-levels" the playing field between the former monopolist and the new entrants.** The first case was decided by the French competition authority against the previous gas monopolist *Gaz de France* (now Engie) which was using its customers list to promote a new gas service. In an interim decision, the authority forced *Gaz de France* to share the list with its competitors on the gas market as such a database was developed under a legal monopoly and was not easily reproducible by new entrants.⁹¹ In the final decision, the authority imposed a fine of €100m on GDF.⁹² The second case was decided by the Belgian competition authority against the National Lottery which was using its customers lists to send a one-off promotional email to launch its new sports betting product.⁹³ Given its nature and size, the authority concluded that the contact details could not have been reproduced by competitors in the market under reasonable financial conditions and within a reasonable period of time.⁹⁴

In the digital sector, two American cases are also interesting. In both cases, a **small firm was relying on the data of a bigger digital platform to provide data analytics services and then, at some point, was cut off from the access to that data.** In the first case, *PeopleBrowsr* analysed Twitter data to sell information about customer reactions to products or about Twitter influencers in certain communities. At some point, Twitter decided that its data will not anymore be accessible directly, but should be bought from certified data resellers. Following a complaint by *PeopleBrowsr*, a Californian Court ordered, with interim measures, that Twitter had to continue to provide its data directly. Then the parties settled the case deciding that after a transition period, *PeopleBrowsr* will get the data from the certified data resellers.⁹⁵ In the second case, *hiQ* analysed LinkedIn public available data to provide information to businesses about their workforces. At some point, LinkedIn limited access to this data by legal and technical means, because it wanted to provide similar services itself. Following a complaint by *hiQ*, a US federal district judge ordered LinkedIn to resume the supply of its data.⁹⁶

⁹¹ Decision 14-MC-02 of 9 September 2014 of the French Competition Authority, *Direct Energie and UFC Que Choisir v. Engie*. This decision is based on the opinion that the French competition authority had adopted in 2010: Opinion 10-A-13 of the French Competition Authority of 14 June 2010 on cross-use of customers database.

⁹² Decision 17-D-06 of 31 March 2017 of the French Competition Authority, *Direct Energie and UFC Que Choisir v. Engie*.

⁹³ Decision 2015-P/K-27 of 22 September 2015 of the Belgian Competition Authority, *Stanleybet Belgium/Stanley International Betting and Sagevas/World Football Association/Samenwerkende Nevenmaatschappij Belgische PMU v. Nationale Loterij*.

⁹⁴ *Ibidem*, par. 69-70.

⁹⁵ <http://blog.peoplebrowsr.com/2012/11/peoplebrowsr-wins-temporary-restraining-order-compelling-twitter-to-provide-firehose-access/> and <http://blog.peoplebrowsr.com/2013/04/peoplebrowsr-and-twitter-settle-firehose-dispute/>

⁹⁶ *hiQ Labs v. LinkedIn*.

Lastly, an ongoing case regarding access to financial data is interesting. In October 2017, the Commission ran inspections in several banks and bank association in Poland and in the Netherlands because it had concerns that those banks may have engaged in anti-competitive practices aimed at excluding non-bank owned providers of financial services by preventing them from gaining access to bank customers' account data, despite the fact that the respective customers have given their consent to such access.⁹⁷ As we will explain in the next section, there is no sector specific rules which imposes the sharing of such financial account data.

3.1.2 Conditions of essential facilities and application to data

The three main conditions of the essential facilities doctrine, which is the test to impose access in EU competition law, are summarised in the Commission Guidance on the application of Article 102 TFEU to exclusionary abuses of dominant position.

(i) Condition 1: Indispensability of data

When access to raw data is requested, assessment of the indispensability condition implies an enquiry as to whether an alternative raw dataset is available or could be collected by a firm having the same size as the data owner (e.g., assessed by market share in the consumer market). This is an empirical analysis that should be examined on a case-by-case basis. The wide availability and the non-rivalry of data often do not make them indispensable as the Commission has concluded in several past merger cases. However, in some cases, data collection may be subject to legal, technical, and economic barriers which may make them indispensable. Besides, many collected data are often generated by the users themselves⁹⁸ which may facilitate the finding of indispensability (as decided in *IMS-Health*). Finally, the fact that the requested data have not already been traded, which is very often the case in practice, should not be an obstacle to imposing sharing as it suffices that there is demand and that such demand can legally and practically be met (as it has also been decided in *IMS-Health*).

When access is about data structure, the assessment of the indispensability condition implies an enquiry as to whether the same information (not necessarily derived from the same raw data sets) is available or could be built by a firm having the same size as the data structure owner. Again, this is an empirical issue, but data structuring may show important network effects and become a *de facto* industry standard (as it was the case in *IMS-Health*).

(ii) Condition 2: Elimination of effective competition in the downstream market

The assessment of the elimination of downstream competition is very complex in case of data. First, the downstream market is not always known, as one of the main features of big data and AI is to experiment, crunch a lot of data without knowing in advance what information or knowledge will be found and what action might be taken. Therefore, the refusal to share data may lead to the possible elimination of a competitor on a not-yet-defined and future market. This requires a more dynamic analysis, better in line with market realities, but is more difficult to do and possibly increasing the risks of antitrust errors.

Second, the data owner is often not (yet) active on the downstream market because, as explained by Drexel (2017): "a typical feature of the data economy is that data is collected for one purpose but may turn out to be interesting for very different purposes pursued by other firms of very different sectors." The evolution of digital industries is quick and uncertain, and many firms are 'paranoid' about the next disruptive innovation.⁹⁹ Thus, a data owner may refuse to share data with a firm that

⁹⁷ Commission Press Release of 6 October 2017 MEMO/17/3761; also Borgogno and Colangelo (2020a).

⁹⁸ For instance, in the connected cars, most data are generated by the driver: Kerber (2018).

⁹⁹ Andy Grove, the iconic founder of Intel, wrote in 1999 a book that he famously titled: *Only the paranoid will survive*. On disruptive innovation, see Gans (2016).

is not (yet) a competitor either because it plans to enter in the downstream market (future offensive leverage) or because it fears that the data seeker will disrupt its business (defensive leverage). In short, given the characteristics of the data economy, refusal to deal while not being active on the downstream market may be anti-competitive exclusionary conduct.

(iii) Condition 3: New product and consumer harm

The interpretation of this condition is not very clear. As explained above, the EU Courts link this condition to the protection of the facility by an intellectual property right but have applied it more strictly in some cases than in others. The Commission integrates this condition into a more general consumer harm assessment. Taking the Courts' interpretation, the first issue is thus to determine whether the data to which access is required are protected by intellectual property (IP) rights. If there is IP protection, the next issue is whether the product that the access seeker aims to bring on the downstream market is sufficiently new or, at least improved, compared to the data owner's products. Drexl (2017) is doubtful that this will often be the case as he considers that the generation of new information due to data sharing is often not sufficiently innovative to justify the compulsory licensing of the intellectual property right.

However, more fundamentally, the assessment of this condition faces the same two difficulties analysed previously for the second condition, i.e., the product to be offered by the access seeker is often still unknown and the facility owner is often not (yet) providing a competing product. Therefore, the more general consumer harm approach proposed by the Commission is more appropriate to the characteristics of the data economy. Thus, the competition authority will have to examine whether, for consumers, the likely negative consequences of the refusal to share data outweigh, over time, the negative consequences of imposing data sharing.

(vi) The adaptation of the essential facilities conditions to the characteristics of data and the digital economy

The **key issue is to determine whether the benefits of compulsory data sharing outweigh its costs**.¹⁰⁰ The benefits are created by the entry of the data access seeker that may bring more competition, innovation, diversity, and choice to the secondary market. The costs are the reduced investment incentives for the facility owner and for the potential access seeker and the operation costs of the antitrust enforcers and the dominant firms that have to implement the access obligations.

Those benefits and costs largely depend on the characteristics of data. The benefits of data sharing may be higher than for other (single-purpose) inputs because data are general-purpose and may be used and re-used in several contexts to build different information and knowledge.¹⁰¹ Conversely, the costs of data sharing on investment incentives may be lower than for other (rival) inputs because data are non-rivalrous and the data owner may keep them while sharing them, hence its incentives to collect, structure or analyse them remain unchanged. Such incentive costs may even be zero when the data were obtained as a by-product of another activity done independently of the data collection, as was the case in *Magill*.¹⁰² In this hypothesis, the value of the data amounts to a windfall gain for its owner. The incentive costs will also be reduced when the data were constituted under the protection of a legal monopoly as in *Gaz de France* or in the *Belgian National Lottery*.¹⁰³

The cost and benefit analysis also depends on the competitive dynamics in the data economy. Data markets show important economies of scale and scope on the supply side and massive direct and

¹⁰⁰ Kerber (2018).

¹⁰¹ Also Abrahamson (2014:879); Meadows (2015).

¹⁰² Graef (2016), Prufer and Schottmuller (2017), Rubinfeld and Gal (2017:377); Schweitzer et al. (2018).

¹⁰³ In those two national cases, we may also claim that the customers list was also a by-product to the core activities of the data owner.

indirect network effects on the demand side.¹⁰⁴ This leads the markets to tip more often than in other sectors of the economy, which implies that competition enforcement should focus on preserving the contestability of those markets for which data sharing may be key. Data markets also show rapid and uncertain innovation often after extensive experimentation. This requires a better understanding of the firms' strategies that may, for instance, terminate data sharing to free ride on the experimentation costs or refuse to share data to alleviate the risks of future disruption.

Therefore, **applying the same cost-benefit analysis which is at the core of the antitrust case-law of duty to deal in light of the different characteristics of data and the competitive strategies and dynamics of the data economy, leads us to suggest that the threshold for imposing data sharing under Article 102 TFEU should be lower than the threshold to impose access to other products.**¹⁰⁵ However, a lower threshold does not mean no threshold, as the freedom to contract and the right of propriety still need to be protected in the data economy. As in the other sectors of the economy, the antitrust agency should convincingly demonstrate that the benefits of sharing data outweigh its costs.

If the benefits outweigh the costs of data sharing, and if exceptional circumstances thus justify the imposition of data sharing, the competition authorities should then determine the **quantitative and qualitative conditions for such access**. The *Microsoft Compliance* case confirmed that a competition authority should not impose a specific price but may rely on more open terms provided they are sufficiently precise for the data owner to determine with enough legal certainty the price to charge. Thus, as suggested by Rubinfeld and Gal (2017), the authority may require that data should be shared on FRAND terms. It would then be up to the data owner to propose a price that complies with this obligation, applying the basic principles proposed in the Commission SEP Communication. In case of disagreement between parties, the framework for good faith negotiation provided in *Huawei* could be imposed.¹⁰⁶ Moreover, the Support Centre for data sharing to be set up by the Commission could facilitate those negotiations as this Centre should focus on the ways and means by which data are exchanged and provide support to make the exchanges easier, in particular by developing model contract terms for B2B data sharing.¹⁰⁷

3.2 Compulsory access under Merger Regulation

The European Commission has analysed competition concerns relating to the combination of data several times in merger decisions. So far, the Commission has not yet blocked a merger on the ground that the combination of data would give rise to competition concerns.

In its 2007 *Google/DoubleClick* merger decision, the Commission argued that the combination of Google's data on users' search behaviour with DoubleClick's data on web-browsing behaviour of users would not give the merged entity a competitive advantage that could not be matched by competitors.¹⁰⁸ According to the Commission, such a combination of information was already available to several Google's competitors, including Microsoft and Yahoo which both ran search engines and offered ad serving at that time as well. Besides, the Commission argued that competitors could purchase data or targeting services from third parties including portals, other major web publishers, and internet service providers.¹⁰⁹

¹⁰⁴ As explained in Bourreau and de Stree (2019), such characteristics of the data markets amplify the pro and anti-competitive effects of firms' behaviours such as refusal to share.

¹⁰⁵ Also calling for lower threshold, Abrahamson (2014), Kerber (2018:328), Meadows (2015); Schweitzer et al. (2018). OECD (2015) goes also in the same direction.

¹⁰⁶ Case C-170/13, *Huawei v. ZTE*, EU:C:2015:477.

¹⁰⁷ Commission Staff Working Document of 25 April 2018, Guidance on sharing private sector data in the European data economy, SWD(2018) 125, p.6.

¹⁰⁸ Case COMP/M.4731 – *Google/ DoubleClick*, 11 March 2008, par. 366.

¹⁰⁹ Case COMP/M.4731 – *Google/ DoubleClick*, 11 March 2008, par. 269-272 and 365.

In 2008, the Commission analysed the **Thomson/Reuters** merger¹¹⁰ and had concerns that such merger would reduce significantly the competition in the markets for the distribution of aftermarket broker research reports, of earning estimates, of fundamental financial data of enterprises and of time series of economic data. To remove those concerns, the merging parties committed to divest copies of the databases containing the content sets of such financial information products, together with relevant assets, personnel and customer base as appropriate to allow purchasers of the databases and assets to quickly establish themselves as a credible competitive force in the marketplace in competition with the merged entity, re-establishing the pre-merger rivalry in the respective fields. The parties could also continue to use these databases in the future to commercialise the respective data to their own customers. With those data sharing remedies, customers of such financial information products therefore would continue to have sufficient alternatives post-merger.

In 2014, the Commission analysed data-related competition concerns in **Facebook/WhatsApp**. According to the Commission, the acquisition of WhatsApp would not increase the amount of data potentially available to Facebook for advertising purposes because WhatsApp did not collect data valuable for advertising purposes at the time of the merger.¹¹¹ The Commission also investigated possible theories of harm relating to data concentration to the extent that it might strengthen Facebook's position in the market for online advertising. In that regard, the Commission argued that the merger would not raise competition concerns even if Facebook would introduce targeted advertising on WhatsApp or start collecting data from WhatsApp users to improve the accuracy of the targeted ads served on Facebook's social networking platform.¹¹² In the Commission's view, there would continue to be a sufficient number of alternative providers to Facebook for the supply of targeted advertising after the merger, and a large amount of internet user data that are valuable for advertising purposes were not within Facebook's exclusive control. In particular, the Commission considered Google, Apple, Amazon, eBay, Microsoft, AOL, Yahoo!, Twitter, IAC, LinkedIn, Adobe, and Yelp as market participants that collect user data alongside Facebook.¹¹³

Whereas the Commission in *Facebook/WhatsApp* did not define a possible market for data or data analytics services on the ground that "*neither of the Parties is currently active in any such potential markets*",¹¹⁴ an evolution is visible in **Microsoft/LinkedIn**. Under the assumption that such data combination is allowed under the applicable data protection legislation, the Commission distinguished two main ways in which the *Microsoft/LinkedIn* merger could raise competition concerns as a result of the combination of data. First, the Commission acknowledged that the combination of two datasets as a result of a merger may "*increase the merged entity's market power in a hypothetical market for the supply of this data or increase barriers to entry/expansion in the market for actual or potential competitors, which may need this data to operate on this market*". Second, the Commission made clear that, even if there is no intention or technical possibility to combine the two datasets, "*it may be that pre-merger the two companies were competing with each other based on the data they controlled and this competition would be eliminated by the merger*".¹¹⁵

¹¹⁰ Commission Decision of 19 February 2008, Case M.4726 *Thomson/Reuters*.

¹¹¹ Case COMP/M.7217 – *Facebook/WhatsApp*, 3 October 2014, par. 166.

¹¹² In May 2017, the Commission imposed a 110 million euro fine on Facebook for providing misleading information during the merger investigation. While Facebook had informed the Commission that it would be unable to establish reliable automated matching between Facebook users' accounts and WhatsApp users' accounts, WhatsApp announced updates to its terms of service in August 2016 including the possibility of linking WhatsApp users' phone numbers with Facebook users' identities. However, the fact that misleading information was given did not impact the 2014 authorisation of the transaction as the decision was based on a number of elements going beyond automated user matching and the Commission at the time carried out an 'even if' assessment assuming user matching as a possibility (Press release European Commission, 'Mergers: Commission fines Facebook €110 million for providing misleading information about WhatsApp takeover', 18 May 2017, available at http://europa.eu/rapid/press-release_IP-17-1369_en.htm).

¹¹³ Case COMP/M.7217 – *Facebook/WhatsApp*, 3 October 2014, par. 187-190.

¹¹⁴ *Ibidem*, par. 72.

¹¹⁵ Case M.8124 – *Microsoft/LinkedIn*, 6 December 2016, par. 179.

Despite this evolution in approach, the Commission nevertheless came to the same conclusion in *Microsoft/LinkedIn* as in *Google/DoubleClick* and *Facebook/WhatsApp*, namely that the combination of data-enabled did not raise serious doubts as to the merger's compatibility with the internal market concerning online advertising.¹¹⁶ The Commission specified three grounds for this. First, Microsoft and LinkedIn did not make available their data to third parties for advertising purposes, with only very limited exceptions. Second, the combination of their respective datasets did not appear to result in raising the barriers to entry/expansion for other players in this space, as there would continue to be a large amount of internet user data that were valuable for advertising purposes and that were not within Microsoft's exclusive control. Third, the merging parties were small market players and competed with each other only to a very limited extent in online advertising and its possible segments.¹¹⁷

In its 2018 ***Apple/Shazam*** merger decision, the Commission again concluded that the combination of the datasets of the two companies would not give rise to competition concerns. This time the focus was not on online advertising but on digital music streaming apps where Apple is active with its Apple Music service and Shazam offers a leading music recognition application. According to the Commission, it would be unlikely that the merged entity would have the ability to foreclose competing providers of digital music streaming apps even if Shazam's data would be integrated into Apple's dataset. Shazam's data, in the Commission's view, does not seem to be an important input to improve existing functionalities or offer additional functionalities within digital music streaming apps and does not appear to be unique based on a comparison with other alternative datasets in relation to the metrics associated with the Four V's of big data, namely the variety of data, the speed at which the data is collected (velocity), the size of the dataset (volume), and its economic relevance (value).¹¹⁸

In the ongoing ***Google/Fitbit*** case, the Commission opened a Phase II investigation as the merger may entrench Google's market position in the online advertising markets by increasing the amount of data that Google could use for ads personalisation.¹¹⁹ During the first phase of the investigation, Google submitted commitments consisting in the creation of a data silo - which is a virtual storage of data - where certain data collected through wearable devices would have been kept separate from the other datasets within Google. Hence, the data in the silo would have been restricted from usage for Google's advertising purposes. However, the Commission considered that the data silo commitment was insufficient as it did not cover all the data that Google would access as a result of the merger and would be valuable for advertising purposes. This case is interesting as it shows that the Commission may prefer data siloing over data sharing to remedy some competition concerns when two data-rich are merging.

3.3 Limits of data sharing by Article 101 TFEU

If competition law may impose data sharing to remedy an abuse of dominant position or a merger between two data rich firms, competition law may also limit the sharing of commercially sensitive information among competitors. Article 101(1) TFEU can limit data sharing in situations where the exchange of information among competitors gives rise to collusion and the resulting restriction of

¹¹⁶ The Commission reached the same conclusion in *Verizon/Yahoo*, see Case M.8180 – *Verizon/Yahoo*, 21 December 2016, par. 89-93.

¹¹⁷ Case M.8124 – *Microsoft / LinkedIn*, 6 December 2016, par. 180. The Commission similarly investigated a hypothetical market for data in two other instances in the *Microsoft/LinkedIn* decision, namely in the context of customer relationship management software solutions and with regard to the use of data for machine learning in productivity software solutions. For both instances, the Commission concluded that the transaction did not raise serious doubts as to its compatibility with the internal market. See Case M.8124 – *Microsoft / LinkedIn*, 6 December 2016, par. 253-264 and 373-381. For an analysis, see Graef (2018b: 79-81).

¹¹⁸ Case M.8788 – *Apple/Shazam*, 6 September 2018, par. 313-330.

¹¹⁹ Press Release of the Commission of 4 August 2020, IP/2020/1446.

competition cannot be justified under Article 101(3) TFEU by showing that the procompetitive effects outweigh the anticompetitive effects.

In ***Asnef-Equifax***, the Court of Justice was asked to assess the compatibility with Article 101 TFEU of a register set up by financial institutions in Spain involving the exchange of solvency and credit information about their customers to evaluate the risks of engaging in lending and credit activities. The Court argued that in order not to restrict competition under Article 101 TFEU: (i) the relevant market at stake should not be highly concentrated; (ii) the register should not be capable of revealing the identity of the lenders, as this could help to identify the market position or commercial strategy of competitors; and (iii) the register should be accessible in a non-discriminatory manner to all operators active in the relevant sphere so that some operators are not put at a disadvantage if they do not have access to information needed for risk assessment.¹²⁰ These three conditions could also be applied to assess data pooling arrangements where information about individual customers is shared among market players.

The Commission ***Horizontal Agreements Guidelines*** pay attention to how certain factors can make the exchange of information among competitors more problematic, including the strategic nature of the information, the market coverage of the firms involved, the individualised or aggregated nature of the company information exchanged, the age of the data, the frequency of the information exchange, the public or non-public nature of the information, and whether the exchange of information is public or non-public.¹²¹

Most illustrative is the investigation the Commission opened in May 2019 into the data pooling system of ***Insurance Ireland***, which is an association bringing together companies active in the insurance sector in Ireland. As part of its activities, Insurance Ireland administers a database to which member companies contribute insurance claims data on an ongoing basis. According to the Commission's press release, the objective of the system is "*to facilitate the detection of potentially fraudulent behaviour by insurance claimants and to ensure the accuracy of information provided by potential customers to insurance companies and/or their agents*".¹²² While the Commission acknowledges that such a data pooling system may benefit consumers by ensuring more suitable products and competitive prices, it is concerned in particular about whether the conditions of access to the system of Insurance Ireland restrict competition and thereby reduces Irish drivers' choice of insurance policies.¹²³

Apart from listing the factual circumstances of the case, the press release also provides a more general background as to the Commission's current thinking about data pooling. The Commission states that data pooling arrangements are often pro-competitive: (i) they directly benefit consumers by enabling effective competition on the market, as service providers may be able to offer better prices and services to consumers by accessing and participating in a data pool; and (ii) access to data in a data pool may enable effective market entry, resulting into the improved choice of services and suppliers to the benefit of consumers. However, the Commission also points out that data pooling arrangements may in some situations lead to restrictions of competition, for instance when: (i) the conditions of access to and participation in a data pool result in placing certain market players at a competitive disadvantage; or (ii) the data pooling system enables market players to become aware of the market strategies of their competitors.¹²⁴


¹²⁰ Case C-238/05 *Asnef-Equifax*, EU:C:2006:734, par. 58-61.

¹²¹ Commission Guidelines on the applicability of Article 101 of the Treaty on the Functioning of the European Union to horizontal co-operation agreements [2011] OJ C 11/1, par. 86-94.

¹²² Press release European Commission of 14 May 2019, IP/19/2509.

¹²³ *Ibidem*.

¹²⁴ *Ibidem*.



Beyond restrictions on access to the pool on which the investigation of the Commission against Insurance Ireland focuses, another question is when the existence of data pooling arrangements in themselves can breach Article 101 TFEU through the exchange of commercially sensitive information among competitors. In this regard, Lundqvist (2018) makes a distinction between three situations. Whereas the exchange of technical information for the development of new products or interoperability among existing products through a data pool seems largely unproblematic, data pooling arrangements where parties share strategic and competitive information regarding prices or innovations have to be considered as potentially breaching Article 101 TFEU. Within those two extremes lie data pools in which not directly commercially sensitive information is shared but where information is exchanged about a large number of customers in a way that may ultimately enable a member to the pool to extract competitive insights based on data analytics. In its ongoing revision of the Horizontal Guidelines, the Commission is expected to clarify how these existing indicators have to be applied to assess the more complicated data pooling arrangements where possible anticompetitive effects are less pronounced than in previous cases.

Cremer et al. (2019) made interesting suggestions as to how the assessment of R&D agreements or patent pools could inspire the competition analysis of data pooling. As regards the pooling of inferred data, the limits set on coordination in the context of R&D agreements are argued to be relevant because the sharing of inferred data may decrease incentives to engage in independent data processing and thus reduce competition in the field of data analytics. Concerning possible analogies with patent pooling, the point is made that it is much more difficult for data, especially observed data, as compared to patents to be categorised as either substitutable/non-substitutable or essential/non-essential. Interestingly, reference is also made to mandated access as a remedy to prevent data pools with market power from restricting competition. This would mean that data needs to be shared with third parties, for instance under FRAND terms. According to the experts, such a duty to give others access to the pool should be proportional to the pool's market power: "*a group of smaller players pooling their data to gain a competitive advantage should not be forced to give their pooled data to a much larger player*".¹²⁵

¹²⁵ Crémer, de Montjoye and Schweitzer (2019:96-97).

04

HORIZONTAL AND SECTORAL EU LAWS FOR DATA SHARING

4 Horizontal and sectoral EU laws for data sharing¹²⁶

Next to competition law which may impose or limit data sharing under specific circumstances, other EU laws complement antitrust rules and impose further data sharing obligation which go beyond what is imposed under antitrust law or limit data sharing in a more restrictive manner than what is prohibited under antitrust law.¹²⁷

4.1 Existing legislations enabling data sharing

The EU legal framework contains several rules imposing or encouraging the portability and the sharing of data.¹²⁸ The rules imposing data portability tend to be general and apply to all sectors of the economy. They are mainly composed of: (i) for personal data, the *General Data Protection Regulation* (GDPR)¹²⁹ and (ii) for non-personal data, the *Digital Content Directive* (DCD)¹³⁰ applicable in a B2C relationship and the *Free Flow of Data Regulation* (FFDR)¹³¹ applicable in a B2B relationship.

The rules imposing data sharing are general for public data and imposed by the *Open Data Directive* (ODD).¹³² For private data, data sharing rules tend to be sector-specific and are mainly composed of the *Second Payment Service Directive* (PSD2) imposing access to payment account data;¹³³ the new *Motor Vehicle Regulation* imposing access to some vehicle data;¹³⁴ the new *Electricity Directive* imposing access to some customers data;¹³⁵ the *European Electronic Communications Code* (EECC) for access to directory data;¹³⁶ the *Postal Services Directive* for access to postal address.¹³⁷

Table 1: EU legal framework for data portability and sharing

	Personal data	Non-personal data
Data portability	- GDPR (2016)	- DCD (2019) in B2C - FFDR (2018) in B2B
Data sharing	<ul style="list-style-type: none"> - Public data: ODD (2019) - Financial: PSD2 (2015) and UK Open Banking (2016) - Automotive: Motor Vehicle Regulation (MVR) (2018) - Energy: Electricity Directive (2019) - Electronic Communications: EECC (2018) - Postal: Postal Services Directive (2008) 	

¹²⁶ This section is partly based on Graef, Tombal and de Streel (2019) and Kramer, Senellart and de Streel (2020).

¹²⁷ See also Costa-Cabral and Lynskey (2017).

¹²⁸ See Support Centre for Data Sharing (2020).

¹²⁹ Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46 (General Data Protection Regulation), OJ [2016] L 199/1.

¹³⁰ Directive 2019/770 of the European Parliament and of the Council of 20 May 2019 on certain aspects concerning contracts for the supply of digital content and digital services, OJ [2019] L 136/1.

¹³¹ Regulation 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union, OJ [2018] L 303/59.

¹³² Directive 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information, OJ [2019] L 172/56.

¹³³ Directive 2015/2366 of the European Parliament and of the Council of 25 November 2015 on payment services in the internal market, OJ [2015] L 337/35, arts.66-67.

¹³⁴ Regulation 2018/858 of the European Parliament and of the Council of 30 May 2018 on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles, OJ [2018] L 151/1, arts.61-66.

¹³⁵ Directive 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity [2019] OJ L 158/125, art.23.

¹³⁶ Directive 2018/1972 of the European Parliament and of the Council of 11 December 2018 establishing the European Electronic Communications Code, OJ [2018] L 321/36.

¹³⁷ Directive 97/67 of the European Parliament and of the Council of 15 December 1997 on common rules for the development of the internal market of Community postal services and the improvement of quality of service, OJ (1997) L 15/14, as amended by Directive 2002/39, Regulation 1882/2003 and Directive 2008/6.

4.1.1 EU rules on data portability

4.1.1.1 GDPR

Data portability aims to strengthen the data subject empowerment, i.e. the power of control that the data subjects have on their data and to re-balance the relationship between data subjects and data controllers.¹³⁸ To do that, two specific rights are given to the data subjects:

- First, the data subject has the right to receive the personal data concerning him which he has provided to a controller (the data giver) and to transmit those data to another controller (the data seeker) in a B2C2B relationship (art. 20(1) GDPR). For instance, a data subject can receive his current playlist from a music streaming service to find out how many times he listened to specific tracks or to check which music he wants to purchase and to port it to another platform to listen to music from there.¹³⁹
- Second, a data subject has also the right to have his data transmitted directly from one controller to another in a more direct B2B relationship (art. 20(2) GDPR). In essence, this means that a data seeker can import data directly from the data giver with the consent of the data subject.

The first portability right (B2C2B) is the strongest as it should be exercised without hindrance from the data giver. According to the European Data Protection Board (EDPB), such hindrance could be 'fees asked for delivering data, lack of interoperability or access to a data format or API or the provided format, excessive delay or complexity to retrieve the full dataset, deliberate obfuscation of the dataset, or specific and undue or excessive sectorial standardization or accreditation demands'.¹⁴⁰ The second portability right (B2B) is weaker as it can only be exercised when technically feasible, which is assessed on a case-by-case basis. Those two (new) portability rights complement and go further than the (old) data access right given by Article 15(3) of the GDPR.

The scope of the portability right is limited to certain categories of personal data. The GDPR mentioned the data **provided by the data subject**. In its interpretative Guidelines, the EDPB mentions three categories of data:¹⁴¹

- The *data actively and knowingly provided* by the data subject such as name, age, an email address;
- The *observed data* provided by the data subject by virtue of the use of the service or the device, such as search history, traffic and localisation data, the heartbeat tracked by a wearable device;
- The *inferred data and derived data* created by the data controller based on the data provided by the data subject such as the outcome of an assessment regarding the health of a user or the profile created in the context of risk management and financial regulations to assign a credit score.

The EDPB notes that the portability right should be interpreted broadly and covered the first two categories, i.e. the data that have been actively provided by the data subject but also the observed data, and only the third category (the inferred data) should not be covered. However, it remains to be seen whether the EU judges will follow such a broad interpretation.

The scope of the portability right is also limited by the type of processing and covers only personal data whose **processing is based on consent or contract**. There is thus no general right to data

¹³⁸ Guidelines of 13 April 2017 of Working Party 29 on the right to data portability, WP242 rev.01, p. 4.

¹³⁹ *Ibid.*, p. 5.

¹⁴⁰ Guidelines of 13 April 2017 of Working Party 29 on the right to data portability, p.15.

¹⁴¹ *Ibidem*, p.10.

portability as it does not apply to processing operations necessary for the performance of a task in the public interest vested in the controller, nor to processing operations necessary for the compliance with a legal obligation to which the controller is subject. For instance, a financial institution has no obligation to respond to a portability request relating to personal data that has been collected in the context of compliance with its legal obligation to fight money laundering.¹⁴²

Finally, the right to data portability only applies if the data **processing is carried out by automated means**, and therefore does not cover most paper files.

Article 20 of the GDPR imposes that the data have to be provided in a **structured, commonly used, and machine-readable format**.¹⁴³ Recital 68 of the GDPR clarifies further that data controllers are encouraged to develop interoperable formats that enable data portability. According to the EDBP, “the terms structured, commonly used, and machine-readable are a set of minimal requirements that should facilitate the interoperability of the data format provided by the data controller. In that ways, ‘structured, commonly used and machine-readable’ are specifications for the means, whereas interoperability is the desired outcome”. However, such interoperability goals should not go as far as imposing technical compatibility, as it is clarified by Recital 68 of the GDPR.

According to the EDPB, ‘the most appropriate format will differ across sectors and adequate formats may already exist, and should always be chosen to achieve the purpose of being interpretable and affording the data subject with a large degree of data portability. As such, formats that are subject to costly licensing constraints would not be considered an adequate approach’ and ‘where no formats are in common use for a given industry or given context, data controllers should provide personal data using commonly used open formats (e.g. XML, JSON, CSV,...) along with useful metadata at the best possible level of granularity while maintaining a high level of abstraction’¹⁴⁴

The EDBP also encourages cooperation between industry stakeholders and trade associations to work together on a common set of interoperable standards and formats to deliver the requirements of the right to data portability as is done by the European Interoperability Framework (EIF)¹⁴⁵ which creates an agreed approach to interoperability for organizations that wish to jointly deliver public services.¹⁴⁶

Article 12(3) of the GDPR requires that the data giver provides information on action taken to the data subject **without undue delay and in any event within one month** of receipt of the request. This one month can be extended to a maximum of three months for complex cases if the data subject has been informed about the reasons for such delay within one month of the original request.

Article 12(5) of the GDPR provides that data should be ported **free of charge**, unless the data controller can demonstrate that the requests are manifestly unfounded or excessive, in particular, because of their repetitive character. In this case, the controller may either charge a reasonable fee taking into account the administrative costs of porting the data, or refuse to port the data. The EDPB notes that: ‘for information society services that specialise in automated processing of personal data, implementing automated systems such as Application Programming Interfaces (APIs) can facilitate the exchanges with the data subject, hence lessen the potential burden resulting from repetitive requests. Therefore, there should be very few cases where the data controller would be able to justify a refusal to deliver the requested information, even regarding multiple data portability requests’.¹⁴⁷

¹⁴² *Ibidem*, p. 8.

¹⁴³ Machine-readable format is not defined in the GDPR but is defined in the Open Data Directive as ‘a file format structured so that software applications can easily identify, recognise and extract specific data, including individual statements of fact, and their internal structure’: art.2(13) of the Directive 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information, OJ [2019] L 172/56.

¹⁴⁴ Guidelines of 13 April 2017 of Working Party 29 on the right to data portability, p.17-18.

¹⁴⁵ https://ec.europa.eu/isa2/eif_en

¹⁴⁶ Guidelines of 13 April 2017 of Working Party 29 on the right to data portability, p.18.

¹⁴⁷ *Ibidem*, p.15.

The EDBP also specifies that 'the overall system implementation costs should neither be charged to the data subjects, nor be used to justify a refusal to answer portability requests'.

4.1.1.2 Digital Content Directive

Like personal data protection law, consumer law also enables data portability, notably through Article 16 of the Digital Content Directive (DCD) of May 2019. The Directive applies to all suppliers of digital content or services (i.e., virtually any firm in the digital economy that collects data) when dealing with a consumer (i.e., any natural person who is acting for purposes which are outside that person's trade, business, craft, or profession).¹⁴⁸ The DCD grants a **form of portability right for the non-personal data provided or created by the consumer**. However, this right for consumers does not apply in several situations when the content is of little practical use to the consumer, who therefore has a limited interest in the portability of such data, particularly since requiring such a mechanism is costly for the trader.¹⁴⁹

The DCD is only an indirect enabler of data sharing as it solely provides the consumer with a right to retrieve some of its non-personal data. It does not allow the direct transmission of data between two traders. Nevertheless, the underlying idea of the DCD is to allow the consumers to retrieve their data to then share this data with other traders. This new right ensure that consumers can easily switch content providers, by reducing legal, technical, and practical obstacles, such as the inability to recover all the data that the consumer has produced or generated through his use of digital content.¹⁵⁰

Unlike the GDPR, the DCD provides that, when the consumer terminates the contract, the trader must refrain from using the non-personal data provided or created by the consumer.¹⁵¹ The fate of the data held by the original controller/trader, therefore, differs in the two regimes, as the GDPR does not prevent the original controller from continuing to use the ported data, while the DCD provides that the trader must refrain from using the data in the future unless it has been generated jointly by the consumer and others, and other consumers can continue to make use of the content.¹⁵² This difference can be explained by the fact that data can be ported at any time under the GDPR, while data portability is only made possible after the termination of the contract by the consumer in the DCD.

While the GDPR applies to personal data that has been provided by or observed on the data subject, the DCD applies to any **content other than personal data, which was provided or created by the consumer** when using the digital content or digital service supplied by the trader. The scope of application of the DCD is thus complementary to that of the GDPR.¹⁵³ This is welcome as the distinction between personal and non-personal data might be difficult to draw in practice. Indeed, given the GDPR's broad definition of personal data and the technological progress in big data and AI for identification, the vast majority of the data provided or created by the consumer are likely to be considered as personal data. In any case, it should be underlined that the "inferred and derived" personal data, which are not considered as data "provided" by the data subject, are neither covered by the GDPR nor by the DCD, and thus cannot be ported.

¹⁴⁸ Art.2(6) DCD.

¹⁴⁹ Recital 71 of the DCD.

¹⁵⁰ Recital 70 of the DCD.

¹⁵¹ Article 16(3) of the DCD. The only exceptions are if the data has no use outside the context of the content or service; if the data only relates to the consumer's activity when using the content or service; if the data has been aggregated with other data by the trader and cannot be disaggregated or can only be disaggregated with disproportionate effort; or if the data has been generated jointly by the consumer and other persons who continue to use them (Article 16(3) of the DCD).

¹⁵² Article 16(3d) of the DCD.

¹⁵³ This is explicitly stated in art. 16(2) of the DCD, which provides that the trader remains bound by the obligations of the GDPR, which prevails over the DCD in case of conflict (art.3(8) of the DCD).

Using a similar language than the GDPR, the DCD provides that the data must be returned to the consumer in a **commonly used and machine-readable format**. Regarding the deadline to reply to the request, the DCD only provides that the data should be given to the consumer within a **reasonable time** after the termination of the contract. While the DCD does not provide any further information as to how these terms must be interpreted, the deadline of one month provided for in the GDPR could arguably be used to assess this reasonable character. Finally, similar to the GDPR, the DCD provides that the consumer shall be entitled to retrieve the data **free of charge**.¹⁵⁴

4.1.1.3 Free Flow of Data Regulation

The Free-Flow of Data Regulation (FFDR) of November 2018 applies for the porting of non-personal data in B2B relationships. The Regulation instructs the Commission to contribute to the development of **EU Codes of conduct to facilitate the porting of (non-personal) data in a structured, commonly used, and machine-readable format including open standard formats**.

On that basis, SWIPO (Switching cloud service providers and Porting Data), which is one of the Digital Single Market (DSM) Cloud Stakeholders Working Groups gathering more than 100 stakeholders, adopted in November 2019 two drafts Code of conduct: one on the **Infrastructure as a Service** (IaaS) market, and another on the **Software as a Service** (SaaS) market.¹⁵⁵

Those codes of conduct will be assessed by the Commission by the end of 2022.¹⁵⁶ In particular, the Commission will focus on: "(i) the impact on the free flow of data in Europe; (ii) the application of the Free Flow of Data Regulation, especially to mixed datasets; (iii) the extent to which the Member States have effectively repealed existing unjustified data localisation restrictions; and (iv) the market effectiveness of codes of conduct in the area of porting of data and switching between cloud service providers."¹⁵⁷

The Commission also expects that the codes of conduct will be complemented by model contractual clauses allowing "sufficient technical and legal specificity in the practical implementation and application of the codes of conduct, which will be of particular importance for SMEs."¹⁵⁸

4.1.2 EU rules on data sharing

4.1.2.1 Public data: Open Data Directive

The Open Data Directive (ODD) of June 2019 imposes extensive data sharing obligations to public sector bodies and public undertakings and provides the most comprehensive data governance framework in EU law. The Directive imposes to **Public sector bodies¹⁵⁹ and public undertakings the obligation to share their documents¹⁶⁰ for re-use for commercial or non-commercial purposes**.¹⁶¹

Documents should be **available in any pre-existing format and, where possible and appropriate, by electronic means, in formats that are open, machine-readable, accessible, findable, and re-usable, together with their metadata**. Where possible, the format and the metadata should comply with formal open standards.¹⁶² Public sector bodies and public undertakings

¹⁵⁴ Article 16.4 of the DCD.

¹⁵⁵ <https://ec.europa.eu/digital-single-market/en/news/presentation-codes-conduct-cloud-switching-and-data-portability>

¹⁵⁶ FFDR, art. 8.

¹⁵⁷ Commission Guidance of 29 May 2019 on the Regulation on a framework for the free flow of non-personal data in the European Union, COM (2019) 250, p. 18.

¹⁵⁸ Ibidem, p. 17 and FFDR, recital 30.

¹⁵⁹ Art.2(1) defines public sector body as: 'the State, regional or local authorities, bodies governed by public law or associations formed by one or more such authorities or one or more such bodies governed by public law.'

¹⁶⁰ Art. 2(6) of the ODD defines document as '(a) any content whatever its medium (paper or electronic form or as a sound, visual or audiovisual recording); or (b) any part of such content'

¹⁶¹ Art.3 of the ODD.

¹⁶² Art.5(1) ODD.

are also encouraged to apply an 'open by design and by default' policy for their documents. Besides, public sector bodies should make dynamic data available for re-use immediately after collection, via suitable APIs and, where relevant, as a bulk download.¹⁶³

Public sector bodies and public undertakings should also make practical arrangements facilitating the search for documents available for re-use, such as asset lists of main documents with relevant metadata, accessible where possible, and appropriate online and in a machine-readable format, and portal sites that are linked to the asset lists.¹⁶⁴

The ODD also imposes strict **non-discrimination** requirements. According to this principle, any applicable conditions for the re-use of documents should be non-discriminatory for comparable categories of re-use. Moreover, if documents are re-used by a public sector body as input for its commercial activities that fall outside the scope of its public tasks, the same charges, and other conditions should apply to the supply of the documents for those activities as apply to other users.¹⁶⁵

Regarding **pricing**, the recovery of the marginal costs incurred for the reproduction, provision, and dissemination of documents as well as for anonymisation of personal data and measures taken to protect commercially confidential information may be allowed.¹⁶⁶ In some specific cases foreseen by the ODD, a price could be charged provided it is calculated following objective, transparent and verifiable criteria.¹⁶⁷ The conditions and the actual amount of those charges, including the calculation basis for such charges, shall be pre-established and published, through electronic means where possible and appropriate.¹⁶⁸

4.1.2.2 *Financial sector: Access to payment account data*

To stimulate competition and innovation in financial services, the Second Payment Service Directive of November 2015 (PSD2) establishes a framework for new FinTech services to access the **payment account data** – in particular the payment initiative services and the account information services – **securely and after having obtained the consent** of their customers.¹⁶⁹

This sector-specific legislation complements the B2B portability right of the GDPR as it compels the banks (original controllers) to allow direct transmission of the data subjects' personal banking information to recipient controllers. PSD2 goes further than the GDPR because, on the one hand, it forces the banks to ensure the technical feasibility of this B2B financial account data portability and, on the other hand, it makes this portability continuous as data subjects can request personal data at each transaction, facilitated by APIs.

To facilitate and secure such data access and exchange, the Commission adopted **regulatory technical standards** based on a draft submitted by the European Banking Authority.¹⁷⁰ Those rules impose a common and secure open standard for communication between the data giver (the account

¹⁶³ Art.5(5) ODD.

¹⁶⁴ Art.9(1) ODD.

¹⁶⁵ Art.11 ODD.

¹⁶⁶ Art.6(1) ODD.

¹⁶⁷ Art.6(2-5) ODD. Those cases are: (i) public sector bodies that are required to generate revenue to cover a substantial part of their costs relating to the performance of their public tasks; (ii) libraries, including university libraries, museums and archives; and (iii) public undertakings.

¹⁶⁸ Art.7 ODD. Under the previous Directive, the Commission adopted a recommendation which contains interesting guidelines on charging access to data, with methodologies to calculate marginal costs and to recover costs : Commission Guidelines of 17 July 2014 on recommended standard licences, datasets and charging for the reuse of documents OJ [2014] C240/1, section 4.

¹⁶⁹ PSD2, art.66(4) for payment initiation services and art.67(3) for account information services. See Vezzoso (2018).

¹⁷⁰ Commission Delegated Regulation 2018/389 of 27 November 2017 supplementing Directive 2015/2366 of the European Parliament and of the Council with regard to regulatory technical standards for strong customer authentication and common and secure open standards of communication, OJ [2018] L 69/23, arts.28-36.

servicing payment service providers) and the data seekers (the payment initiation service provider or the account information service providers).

The UK went further than the PSD2 with the **Open Banking Programme** which led to a **common and open API** to access the account information of the customers of the nine biggest banks of the country.¹⁷¹ This obligation was imposed by the UK antitrust and consumer protection authority, the Competition, and Market Authority, in the context of its Retail Banking market investigation to increase competition and innovation in the sector.¹⁷²

In practice, the CMA forced those nine biggest banks to fund and cooperate with an independent trustee approved by the CMA. This trustee developed, within a fixed (and short) timeframe, read-only open and common technical and product data standards, and read-and-write open and common banking standards for the sharing of transaction data. Those standards ensure that any communication is secure and based on the consent of the customers. Their establishment has been coordinated with the EU standards developed by the EBA and made compulsory by the European Commission.

As underlined in the Furman Report (2019, p.70), 'one positive example from Open Banking is the effectiveness of requiring at least a subset of firms to implement and deliver the solution. Without such powers, progress is likely to be slow, disjointed, and in some cases non-existent. The issue is not just the complexity of agreeing on unified standards but, potentially important, misaligned incentives between the largest platforms and consumers. Another lesson is that just requiring common standards is not sufficient and that an active effort is needed to make this work in practice.

4.1.2.3 Automotive sector: Access to vehicle diagnostic, repair, and maintenance information

The Regulation on Motor Vehicles of May 2018 imposes to **vehicles manufacturers the obligation to share vehicle On-Board Diagnostic (OBD) and vehicle repair and maintenance data with independent repairers**.¹⁷³ The data should be provided in a **standardised and non-discriminatory** manner and presented in an easily accessible manner in the form of **machine-readable** and electronically processable datasets. Manufacturers should provide a **standardised, secure and remote facility** to enable independent repairers to complete operations that involve access to the vehicle security system.

The manufacturer should make available vehicle repair and maintenance information, including transactional services such as reprogramming or technical assistance, on an **hourly, daily, monthly, and yearly basis**, with fees for access to such information varying in accordance with the respective periods of time for which access is granted. Regarding prices, the manufacturer may charge **reasonable and proportionate fees** for data sharing but those fees could not discourage access to the information by failing to take into account the extent to which the independent operator uses it.

Thus this Regulation on Motor Vehicle complements the GRPR and gives sector-specific data access right for relevant car data to independent repairs to stimulate competition and innovation on this aftermarket.

4.1.2.4 Energy sector: Access to consumer data

To stimulate competition and innovation among electricity suppliers, the new Electricity Directive of June 2019 imposes the **sharing of consumer data**, including metering and consumption data as

¹⁷¹ See Borgogno and Colangelo (2020b).

¹⁷² See CMA Final Report of 9 August 2016 on the Retail Banking Investigation and CMA, pp. 441-460 and CMA Order of 2 February 2017 on the Retail Banking Investigation, Sect. 10 to 14 and the Associated Explanatory Note, paras.28-39. All documents are available at: <https://www.gov.uk/cma-cases/review-of-banking-for-small-and-medium-sized-businesses-smes-in-the-uk>

¹⁷³ MVR, art.61, 63 and Annex X.

well as data required for customer switching, demand response, and other services in a non-discriminatory manner.¹⁷⁴ Each Member State should organise the **management of data** in order to ensure efficient and secure data access and exchange, as well as data protection and data security and should set the **prices for data sharing which should be reasonable and justified**.

Here again, the Electricity Directive complements the GDPR by requiring the Member States to set up a specific regime for consumer data sharing and exchange between electricity suppliers.

4.1.2.5 *Electronic Communications sector: Access to directory data*

To ensure access to comprehensive publicly available directory enquiry services and directories, **providers of number-based interpersonal communications services**, which attribute numbers from a numbering plan should **meet all reasonable requests to give the relevant information for the provision of those directory enquiry services and directories**.¹⁷⁵ The providers cannot discriminate according to the place of establishment of the information seeker and should give the information also to undertakings which are established in a different Member State.¹⁷⁶

The information must be given in an **agreed format and on terms fair, objective, cost-oriented, and non-discriminatory**.¹⁷⁷ The Court of Justice clarified that the data owner can only charge the costs associated with the transmission of the information to the provider of directories. It may not charge the costs of obtaining such information which must in any event be borne by the provider of number-based interpersonal communications services and is already included in the costs and revenue of such services.¹⁷⁸

4.1.2.6 *Postal sector: Access to address database*

Whenever necessary to promote effective competition, the Postal Services Directive (as amended in 2008) provides that Member States shall ensure that **transparent, non-discriminatory access** conditions are available to postal infrastructure including the **postal address database** and information on change of address.¹⁷⁹

4.2 Existing legislations limiting data sharing

The GDPR principles of purpose limitation and data minimisation limit data sharing.¹⁸⁰ In practice, this means that these two principles have to be considered when implementing data portability and data sharing.

The EDBP recommends that the data seeker should inform the data subjects about the purposes for which the ported data will be processed and about the categories of personal data that are adequate, relevant, and necessary for these purposes, to prevent a breach of these purpose limitation and data minimisation principles.¹⁸¹ Moreover, if the data seeker realises that more than necessary data were ported for the purpose pursued, he will have to delete this excessive data as soon as possible, to avoid any liability issue.

¹⁷⁴ Electricity Directive, art.23.

¹⁷⁵ EECC, art.112(1). The relevant information concerns solely the data relating to the subscribers of the undertakings concerned and not the subscribers of other operators: Case C-109/03, *KPN Telecom v. Onafhankelijke Post en Telecommunicatie Autoriteit (OPTA)*, EU:C:2004:749; Case C-543/09 *Deutsche Telekom v Bundesrepublik Deutschland* EU:C:2011:279, para.37.

¹⁷⁶ Case C-536/15, *Tele2 Netherlands, Ziggo, Vodafone Libertel v Autoriteit Consument en Markt (ACM)*, EU:C:2017:214, para.30.

¹⁷⁷ EECC, art.112(1).

¹⁷⁸ Case C-109/03, *KPN Telecom v. Onafhankelijke Post en Telecommunicatie Autoriteit (OPTA)* EU:C:2004:749, paras 39 and 40.

¹⁷⁹ Postal Services Directive, art.11a.

¹⁸⁰ Resp. Article 5(1b) and (1c) of the GDPR.

¹⁸¹ Guidelines of 13 April 2017 of Working Party 29 on the right to data portability, p. 13.

This clarifies one of the uncertainties regarding the **liability faced by the data givers**, namely whether there is a risk that they might be found liable for the unlawful processing of the ported data made by the data seeker because of a breach of these purpose limitation and data minimisation principles. Such a concern has been raised, among others, by Facebook in its White Paper on Data Portability and Privacy.¹⁸² This uncertainty stems from the fact that the GDPR does not tackle this issue. The EDBP has indicated that insofar as the data giver responds to the request for portability, it should not be held liable as a result of the processing carried out on the data by the data seeker.¹⁸³ Indeed, the data giver acts on behalf of the data subject and should not be responsible for any later infringement potentially committed by the data seeker. Nevertheless, according to the EDBP, the data giver should still set up certain safeguards, such as internal procedures to ensure that the data that is transmitted matches the data whose portability is requested, in light of the purpose limitation and data minimisation principles.¹⁸⁴

These two principles will also have to be considered to limit the **porting of personal data from other data subjects** than the one exercising his data portability right. Article 20(4) of the GDPR provides that portability right needs to be articulated with the rights and freedoms of others, that it shall not affect. Accordingly, when a data subject exercises his right to data portability, it is necessary to ensure that the personal data of other data subjects, who have not given their consent to such portability, are not transmitted, at the same time, to a data seeker likely to process the personal data of such third parties.¹⁸⁵ Indeed, while the data subject at the origin of the portability request has given his consent to the data seeker or has concluded a contract with him, this is not the case for the other data subjects whose data could be ported as a result of the exercise of this right.¹⁸⁶

Given that the third parties in question have not consented to the transfer of their data to the data seeker, this transfer can only take place if the purpose for which the transfer is made is compatible with the data giver's initial purpose of processing.¹⁸⁷ If this is not the case, the data seeker has to rely on a new lawful basis for the processing of these third parties' data, such as the basis of legitimate interests.¹⁸⁸

To avoid such an issue, the EDBP suggests that the processing of these other data subjects' data should be authorised only insofar as these data remain under the sole control of the data subject requesting the portability and that they should only be processed for the purposes determined by this data subject.¹⁸⁹ The data seeker could therefore not process these third parties' data for purposes that he has defined himself, such as prospecting purposes. Moreover, the data seeker could not process these data for purposes that are not compatible with the purposes of the data giver. While being appealing in theory, this suggestion is nevertheless extremely restrictive and provides little interest for the data seeker, whose margin of manoeuvre will be severely limited.

However, the EDBP makes another more interesting suggestion. It invites both the data giver and data seeker to implement technical tools allowing the data subject to select the personal data he wishes to port while excluding, where possible, the personal data of other data subjects.¹⁹⁰ This makes it possible to avoid, upstream, potential infringement of the rights of these third parties. However, this is not sufficient in itself, as some personal data of third parties might necessarily have

¹⁸² See Question 5 in Facebook (2019). Charting a Way Forward: Data Portability and Privacy. White Paper. Available at: <https://about.fb.com/wp-content/uploads/2020/02/data-portability-privacy-white-paper.pdf>

¹⁸³ Guidelines of 13 April 2017 of Working Party 29 on the right to data portability, p. 6.

¹⁸⁴ *Ibidem*.

¹⁸⁵ *Ibidem*, p. 11.

¹⁸⁶ See also Question 3 in Facebook (2019). Charting a Way Forward: Data Portability and Privacy. White Paper. Available at: <https://about.fb.com/wp-content/uploads/2020/02/data-portability-privacy-white-paper.pdf>

¹⁸⁷ Article 5 (1b) of the GDPR.

¹⁸⁸ Article 6(1.f) of the GDPR.

¹⁸⁹ Guidelines of 13 April 2017 of Working Party 29 on the right to data portability, p. 12.

¹⁹⁰ *Ibidem*.

to be ported. Accordingly, in addition to these technical tools, it must also be reflected in the implementation of consent mechanisms for these other data subjects, to facilitate data portability.¹⁹¹ Once again, the difficulty is the practical implementing of such a mechanism. For example, in the banking sector, it would be nearly impossible to obtain the consent of all the persons appearing in a list of banking transactions that a data subject would like to port to another bank.

Moreover, **several IP legislations impose some limitations of data sharing to protect the investment incentives in data collection, storage, and analysis.**¹⁹² This is mostly the case of the copyright rules¹⁹³ which apply to the author's own intellectual creation. This includes for computer programs.¹⁹⁴ Interestingly, to allow the scientific research institutions to seize the potential of big data, a new exception for text and data mining for the purposes of scientific research has been introduced.¹⁹⁵

The Database Directive protects the copyright of **databases** which, by reason of the selection or arrangement of their contents, constitute the author's own intellectual creation.¹⁹⁶ In addition, this Directive creates a sui generis rights for 15 years for the maker of a database which shows that there has been qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents to prevent extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database.¹⁹⁷

The **Trade secret** Directive protects trade secret defined as information (hence data) which meets all of the following requirements: (i) it is secret in the sense that it is not, as a body or in the precise configuration and assembly of its components, generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question; (ii) it has commercial value because it is secret; (iii) it has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret.¹⁹⁸

¹⁹¹ Ibidem.

¹⁹² For an overview of those rules, see Drexler (2017).

¹⁹³ Directive 2001/29 of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, OJ [2001] L 167/10 as amended; Directive 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9 and 2001/29, OJ [2019] L 130/92.

¹⁹⁴ Directive 2009/24 of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs, OJ [2009] L 111/16.

¹⁹⁵ Directive Copyright in DSM, art.3.

¹⁹⁶ Directive 96/9 of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, OJ [1996] L 77/20 as amended by Directive 2019/790, art.3.

¹⁹⁷ Database Directive, art.7.

¹⁹⁸ Directive 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, OJ [2016] L 157/1, art.2(1).

05



**REGULATORY
GOVERNANCE OF
DATA SHARING**

5 Regulatory governance of data sharing

In this section, we consider how data sharing obligations might be imposed upon a digital platform by a regulatory body to ensure that markets remain contestable and contested. We do not consider circumstances where a digital platform might volunteer to share data in a way that would promote competition and so make regulatory intervention unnecessary. We recognise that digital platforms do share data and that initiatives such as the Digital Transfer Project, to which we referred earlier, involve voluntary efforts to address some aspects of data sharing which we highlight below.

Various proposals have been made as to the precise institutional form such a regulator might take¹⁹⁹. This may differ between the Member States or might involve the creation of a new pan-European body, either to co-ordinate the activities of national regulatory bodies or even to act in their place²⁰⁰. We consider that form should follow function and that the institutional arrangements should be informed by, and aim to address, the issues that any institution seeking to implement data sharing arrangements is likely to face. We, therefore, discuss institutional arrangements later in this report, after we have introduced the issues that we expect they will need to address. These issues relate to:

- Defining which organisations should be obliged to provide access to data;
- Defining which organisations are entitled to obtain access to data (whether at the initiative of the user or on their initiative) and the conditions to meet before they can do so;
- Defining the types of data to be shared, the geographic scope from which it is drawn, and the conditions under which sharing can occur;
- Defining the nature and scope of the technical standards to be adopted, particularly for the common use of data models and APIs (including whether such standards should align with those being developed to share data in other parts of the European economy), determining how they are developed and which organisations should adopt them;
- Defining other aspects of the data transfer process, including how individual users authorise access and what measures incumbent platforms can adopt to protect or win back users who may be considering switching;
- Defining how disputes and harms arising from failures in the data transfer process will be resolved;
- Defining how the financial terms for access to data will be determined, and what those prices will be.²⁰¹

5.1 The scope of data sharing

5.1.1 Which digital platforms should be obliged to provide access to data?

Various proposals have already been made to identify digital platforms which, as envisaged by the authors of the Furman Report, have 'strategic market status'²⁰² or, in the terms of Stigler Centre authors' 'bottleneck power'²⁰³. The Commission's documents also refer to such platforms as

¹⁹⁹ As noted previously Furman et al (2019) propose the creation of a 'Digital Markets Unit' which might sit within the UK Competition and Markets Authority or reside elsewhere. The UK Government has since established a Digital Taskforce that is chaired by a senior official from the CMA. Scott-Morton et al (2019) propose the creation of a new Digital Authority.

²⁰⁰ The Inception Impact Assessment of the European Commission services for ex ante regulation of gatekeeper digital platforms (EC 2020d) states that the Commission is considering: 'a new ex ante regulatory framework, which would apply to large online platforms that benefit from significant network effects and act as gatekeepers supervised and enforced through an enabled regulatory function at EU level.' p.4..

²⁰¹ We note that the UK Government has consulted on a similar list of issues in their Midata for energy consultation, Department of Business, Energy and Industrial Strategy (2018)

²⁰² Furman et al (2019).

²⁰³ Scott-Morton et al (2019). See also Alexiadis and de Stree (2020).

performing a 'gatekeeper' function. We do not adopt or endorse any particular term or concept in this report but use the term 'gatekeeper status' to represent some (objective) threshold which would allow us to distinguish between those digital platforms which could be subjected to more intrusive behavioural rules, including data sharing obligations, and those which would not.

Although there are differences between the proposals in these reports, the basic assumptions are broadly similar and would involve:

- **The application of a threshold ('gatekeeper') to determine which digital platforms would then be subject to, amongst other things, obligations to comply with a Code of Conduct** or other rules which would govern how they conduct themselves concerning both users of the platform and competitors.
- **An obligation on a wider set of platforms than those to which Codes of Conduct or other measures might apply ('gatekeeper minus') to share data that relates to individual users (portability of individual user data).** This might be required to promote switching between platforms or to facilitate complementary innovation by third parties in adjacent markets. The arrangements for Open Banking, discussed earlier, are an example of this. We discuss the type of data which might be shared in section 5.1.3 below.

We noted earlier that the GDPR, which applies to any organisation and not just to digital platforms or digital platforms with particular characteristics, already gives individual users the right to require the sharing or porting of data relating to them where it is 'technically feasible' to do so. However, the GDPR's provisions are likely to be insufficient if the purpose of the data sharing is to enable competition for several reasons:

- o The scope of the data to be shared, which we discuss further below, may be too limited. The Furman report (which is the study which addresses data sharing most explicitly) proposes that it might need to include inferred personal data, which not currently considered to be within the scope of the porting requirements in the Article 20 of the GDPR²⁰⁴ (and we later conclude that this would be too expansive).
- o To be effective, it is likely to be essential that the data is transferred, at the users' request, directly from one platform to another rather than being first downloaded to the user and then uploaded again.

It appears, therefore that the form of data portability or 'mobility' envisaged by Furman would involve regulatory obligations for digital platforms which go some way beyond those that are currently provided by the GDPR. If direct transfers are required between firms, specific regulation will likely be required to ensure the development of standards and common data models to enable the transfer of data specifically between certain digital platforms. Digital platforms meeting a certain threshold would be obliged to implement such standards and models to enable access to data to third parties in the appropriate format²⁰⁵.

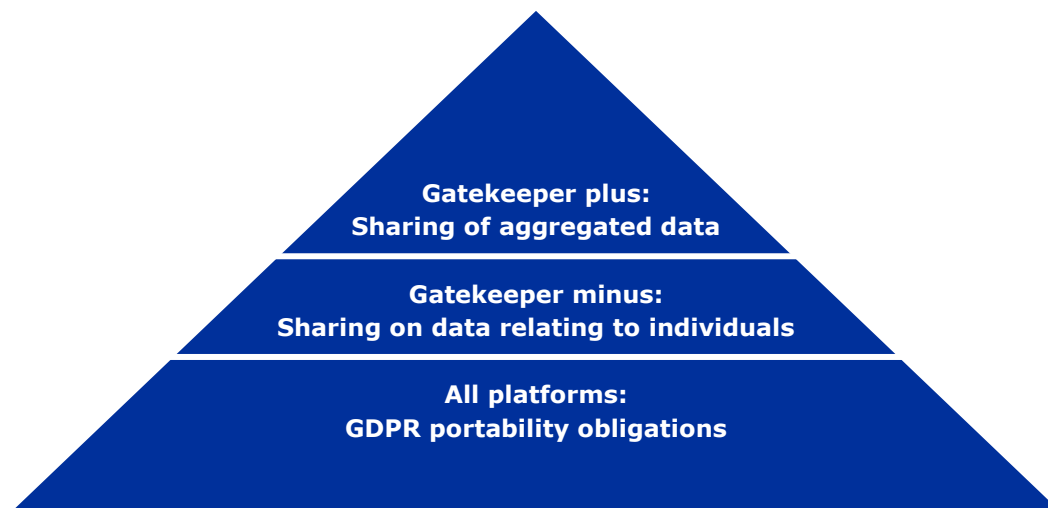
²⁰⁴ The Guidelines of 13 April 2017 of Working Party 29 on the right to data portability, WP242 rev.01 states that the GDPR does not provide individual users with rights to port such data, in contrast with the position on observed data: 'In contrast, inferred data and derived data are created by the data controller on the basis of the data "provided by the data subject". For example, the outcome of an assessment regarding the health of a user or the profile created in the context of risk management and financial regulations (e.g. to assign a credit score or comply with anti-money laundering rules) cannot in themselves be considered as "provided by" the data subject. Even though such data may be part of a profile kept by a data controller and are inferred or derived from the analysis of data provided by the data subject (through his actions for example), these data will typically not be considered as "provided by the data subject" and thus will not be within scope of this new right'. We note that the consent of the individual user would still be required if inferred personal data were to be included within the scope of a data sharing remedy (which, for the reasons explained below, we do not propose).

²⁰⁵ The Furman Report also cites Open Banking as a potential model, in which a sub-set of the largest UK banks (the so-called CMA9) were initially required to develop and implement a common set of APIs, data and security standards so as to enable access to data before including many smaller 'challenger' banks subsequently. It is important to note that the nine UK banks that were

- **An obligation on a narrower set of platforms (all of whom would have 'gatekeeper status' as a necessary, but not a sufficient, condition) to bulk sharing of aggregated data ('gatekeeper plus').** The authors of the Furman Report do not specify which types of organisations would be obliged to share data (nor under what circumstances) under what they refer to as an 'open data' initiative, preferring to leave that to consideration by their proposed new regulatory body, the Digital Markets Unit. Again, we discuss the type of data to be shared in Section 5.1.3 below.

In our view, it would be reasonable to presume that any platform that has been designated as having 'gatekeeper status' could be obliged to share aggregated data in bulk to ensure that markets are both contestable and contested. It would also be reasonable to presume that firms lacking such a designation could not be required to share data in bulk. This does not mean that data sharing would always be the remedy that is selected or that every platform designated as having 'gatekeeper status' would be obliged to share aggregated data²⁰⁶. As noted earlier, the precise conditions which a platform would need to meet to be designated as a 'gatekeeper' are beyond the scope of this report. So, also, are the additional conditions that would be required for a platform to meet the proposed 'gatekeeper plus' threshold, or the condition which need not be present for a platform to meet the 'gatekeeper minus' designation. The relevant conditions are likely to depend on the particular market and context that is being considered (as would be the case if for example, the Commission were to adopt data-sharing remedies following the application of the New Competition Tool). The terms 'plus' and 'minus' are intended to suggest a hierarchy under which different data sharing remedies might be applied, but each would need to be shown to be proportionate and non-discriminatory when applied in a specific context.

The hierarchy of obligations and thresholds we envisage can be illustrated as follows:²⁰⁷



selected by the Competition and Market Authority in that case were the largest banks, but were not considered to hold a dominant position or to meet any other clearly defined threshold.

²⁰⁶ Prufer argues that (aggregate) data sharing obligations are appropriate in what he characterizes as 'data driven markets' which can be easily and quickly foreclosed. He characterizes the problem as follows: 'This tendency of data-driven markets to tip is that the smaller firms, even if they are equipped with a superior idea/production technology, face higher marginal costs of innovation because they lack access to the large pile of user information that the dominant firm has access to due to its significantly larger user base. Consequently, if a smaller firm were to heavily invest in innovation and roll out its high-quality product, the dominant firm could imitate it quickly ---at lower cost of innovation ---and regain its quality lead. The smaller firm would find itself once again in the runners-up spot, which entails few users and low revenues, but it would still have to pay the large costs involved in attempting a leap in innovation', p.6. We find this helpful and consider that digital platforms that are found to have strategic market status are very likely to operate in markets which have these characteristics.

²⁰⁷ Mayer-Schönberger and Ramge (2018) also advocate for differentiated data sharing obligation according to the market power of the digital platforms.

Two important conclusions follow from this:

- First, the number of digital platforms that might be **obliged to share data ought to depend upon (a) the type of the data to be shared and mode of sharing and (b) the intended competitive purposes to which the data will be applied.**

Obligations might be extended to a **wider group of digital platforms** concerning the sharing of data which:

- o Has been provided to the platform by individual users and/or is observed personal data relating to a specific individual user;
- o The intervention is intended to promote complementary innovation in newly emerging adjacent markets for the benefit of individual users;
- o Requires the consent of individual users to initiate the sharing of the data.

This would imply that a wider group of platforms could be required to share individual user data continuously.

Conversely, obligations might be applied to a much **narrower group of digital platforms**, is a sub-set of those with the strategic market status, when the data:

- o Is aggregated observed data resulting from interactions between large numbers of users and the platform;
- o Is critical to the delivery of services by the digital platform in its core market and has been acquired for this purpose, amongst others;
- o The intervention is intended to promote competition into newly emerging adjacent markets (and potentially into the core market), for the benefit of users in general;
- o Does not require the consent of individual users to share the data.

This would imply that a narrower group of platforms could be required to share aggregate user data in bulk.

- **Second, the number of platforms that might be obliged to adopt common standards that would enable data sharing obligations to be met should be greater than the number of platforms that might be obliged to implement data sharing arrangements.** We discuss this issue further below. We consider it highly desirable that the same technical standards to enable data sharing should be applied both when a wider range of digital platforms share data relating to individual users and when a smaller sub-set of those platforms share data that has been aggregated. This should reduce duplication and costs and would mean that a digital platform which was already sharing data relating to individuals would be in a position to comply quickly with an obligation to share aggregated data if that were required at a later stage.

We think it is important to avoid a situation in which a regulator wishes to oblige a digital platform to share data but then finds that the practical implementation takes months or years to accomplish whilst the platform in question adopts the technical standards and makes the changes necessary to enable it to comply with the direction. Since a primary aim of the data sharing measures we consider in this report is to avoid the rapid foreclosure of adjacent markets by very large digital platforms that can otherwise leverage their data advantages from one market to another, the speed at which remedies can be implemented is both a critical consideration and a key rationale for establishing a regulatory regime rather than relying upon competition law.

Accordingly, we propose that the **obligation to adopt common standards to facilitate data sharing should apply independently (and before) consideration of whether the digital platform in question has a particular obligation to share a particular type of data at a particular time.** In other words, an obligation on the part of digital platform to adopt a set of common standards should not pre-judge the determination as to whether, or when, it should be obliged to share data to remedy competition concerns. We recognise that this could mean that some digital platforms incur costs in migrating from proprietary standards to common standards, both in terms of direct implementation costs and potentially in terms of foregone innovation or other benefits which it derived from its proprietary arrangements. The scope of the obligation to adopt common standards should therefore be confined to those platforms where there is a reasonable prospect of their being obliged to share data, rather than being a universal requirement. On the other hand, we would expect that most significant digital platforms will at least have obligations to share data relating to individual users in the way we propose (i.e. under the 'gatekeeper minus' threshold) and that other digital platforms that were seeking to obtain access to such data could be expected to adopt the standards voluntarily to obtain the benefits of doing so. Having adopted common standards for the sharing of data relating to individuals, the same arrangements would also allow for more restricted forms of sharing of aggregated data to be implemented by a subset of the very largest digital platforms.

5.1.2 Which organisations are entitled to obtain access to data (and the conditions to meet before they can do so)?

Another issue in determining the impact of data sharing on contestability relates to the question of which organisations are entitled to receive the data. If the objective of sharing data is to promote contestability and entry, whether in the form of direct replication of services in the core market or, more likely, through complementary innovation in adjacent markets, then we might not expect to find regulators restricting or predetermining which firms should benefit from having access to the data. Regulators themselves are unlikely to be well placed to predict which firms are best placed to generate benefits for users as a result of obtaining access to data.

Thus, it might be argued that no restrictions in terms of access to data are required. For example, although firms seeking access to data would need to adopt the same common technical standards discussed above, there is no obvious need for a regulator to oblige these firms to do so since, unlike the platforms that are required to provide access, those seeking access are likely to expect to benefit from doing so. The more relevant question, in this case, is whether the standards themselves are developed in such a way that they represent an unnecessary cost or barrier to entry for new entrants. That concern ought to be addressed by ensuring adequate representation of their interests in the standards-setting process.

It might also be argued that concerns, which are likely to be held both by policymakers and by those platforms being required to share data, about how third party recipients will store and use the data they obtain are already addressed by the provisions of the GDPR (at least as regards personal data), or might otherwise be addressed via contractual arrangements (as regards other forms of data) between the parties. **However, we conclude from the evidence presented earlier that both organisations and users are likely to require significant additional safeguards and assurances if they trust the arrangements that are in place.** Without such trust, individual users will not initiate transfers and the objectives of data sharing arrangements will not be achieved. Numerous studies have concluded that the absence of well-developed mechanisms to build trust around data sharing - alongside the very visible cases in which organisations have betrayed such

trust in recent years - is an important reason why the sharing of data in Europe and elsewhere remains so limited, relative to its potential²⁰⁸.

Open Banking provides an example of what might be required. In this case, organisations which wish to obtain access to data (generally new fintech companies) must first be accredited by Open Banking Implementation Entity. The accreditation process is intended to ensure that the organisation is viable and has adopted the relevant standards and protocols, and to allow the regulator to understand the purposes for and manner in which the organisation intends to use the data. Besides, the Furman Report proposed that 'sandboxes' be employed to allow new firms to conduct trials of data sharing within a controlled environment that is overseen by the regulatory body²⁰⁹. The UK Financial Conduct Authority has already begun to use sandboxes before approving new products which, in this case, are often developed by organisations which are already subject to close supervision under existing UK financial service regulation²¹⁰.

We consider that any regulatory regime for data access is likely to have to devote as much attention to the regulation of those firms that obtain access to the data (through a licensing or accreditation regime) as to regulating the firms that are obliged to provide access. This is another important reason why regulation rather than competition law remedies is preferred for data sharing since the latter tend to be suited to the imposition of obligations on the dominant firm but not for the imposition of obligations on other parties, such as the recipients of data. Financial service providers are already accustomed to operating under a very detailed and extensive set of regulatory obligations (including Know Your Customer, fraud reporting, risk management, and other obligations) that are administered by a large and well-established regulatory body (in the UK, the Financial Conduct Authority employs a staff of around 4000). In the case of digital markets that we are considering, no such regulatory or supervisory regime exists today - either in Europe or elsewhere.

The number of organisations seeking access to data which may require to be regulated is likely to be many times greater than the number of platforms that are obliged to provide access to their data. They may include existing digital platforms, publishers, or other organisations seeking to enter markets. Some of these entrants may have limited resources, in terms of staff that might be dedicated to ensuring compliance with regulatory obligations, but may nonetheless have the technical capacity to retain and exploit very large volumes of data and may present very significant risks if such data is mismanaged. As explained above, we envisage two main forms of data sharing: the sharing of data relating to specific individuals, for whom the risks are likely to be of great significance to those individuals but not to others, and the sharing of large volumes of aggregated data, where the risks are likely to be more diffuse but no less significant.

Concerns about the integrity and trustworthiness of the organisations that might obtain access to data provide one rationale for restricting access to those that are appropriately authorised and regulated for that purpose²¹¹. However, another concern is that **some access seekers may obtain access to data and use it for legitimate purposes which do not contribute to the promotion of competition in the markets concerning which competition concerns arise.** The nature of data makes this more difficult to police than, for example, the provision of access to assets such as

²⁰⁸ 'The social and economic risks associated with the possible revelation of confidential information (e.g. personal data and trade secrets) are often the main rationale for individuals and organisations not sharing their data', OECD (2019), p.17;

²⁰⁹ Data sandboxes may support the on-boarding of new organisations seeking access to data, or the testing of new functionality. In addition, as the OECD explains, data sandboxes can be a permanent arrangement which ensures the security of very sensitive data by ensuring that it remains behind the firewalls of the host organisation, see OECD (2019), p.34

²¹⁰ Financial Conduct Authority (2017)

²¹¹ We recall that the Furman Report proposes that platforms with 'strategic market status' be required to comply with a Code of Conduct or otherwise regulated. This would exclude new entrants and other platforms who we would expect to be the main access seekers under the arrangements we are contemplating. In addition, the provisions of such a Code are intended to address other competition concerns, not to safeguard the interests of users and organisations when access to data is being provided.

telecommunications or electricity networks which can only be used for a narrow set of purposes. For example, if an incumbent search engine were required to provide access to data in the expectation that this would remedy concerns about competition in the provision of general or specialised search services, then the remedy would be ineffective if the data were instead used to enter markets in which the same incumbent was neither present nor ever expected to be (e.g. was used to compete in markets credit checking services which were already judged to be functioning well). This appears to have been a concern of the Commission's advisers as regards the application of data sharing obligations under European competition law.²¹²

The counterpart CERRE report by Krämer, Schnurr, and Broughton Micova contains a more extensive discussion of the prospects that data sharing remedies might enable entry into the core markets of the large digital platforms, such as general search or e-commerce²¹³. They consider it unlikely as well as potentially undesirable. They conclude that the primary purpose and objective of data sharing measures should instead be to ensure that 'complementary' or niche markets, into which the incumbent digital platforms might otherwise be able to leverage their data-driven advantages, should remain contestable to others. If this policy were successful, they speculate that it may be that the niche competitors who have been enabled by these measures may eventually extend the scope and scale of their activities to represent a genuine competitive threat in the core market of the regulated platform.

The potential benefits, in terms of the scope of competition and contestability, of any data sharing measures are an important matter. Krämer et al explain that if data sharing measures are to be intended to preserve the contestability of niche markets then competition law remedies, which focus on 'essential data', are unlikely to be appropriate or effective²¹⁴. This is one reason why they favour ex ante regulation of data sharing of the kind we are considering in this report. We would add, however, that if the benefits are viewed as being relatively narrow, it may be more difficult for regulators to justify costly interventions on proportionality grounds. These are matters to assess in the light of the specific facts, and so we do not anticipate them here. We do not read Krämer et al to be saying that their conclusions as to which data might be shared, or with whom, would vary substantially different if the objective were to enable entry into core rather than niche markets. We explain later in this section why the types of data which we propose should be shared would, by their nature, also be more likely to enable competition in a niche than in core markets.

We find it **difficult to see how, as a practical matter, a regulator could predefine the commercial purposes to which any data, once transferred, should be applied by the recipient** (as opposed to the type of data to be shared). Nor do we see how a regulator could police any such limitations if they were attempted. This is particularly so with new digital markets, where the boundaries between one market and another are often unstable and unpredictable. As noted above, this is in stark contrast with more conventional access arrangements that involve tangible assets such utility networks, where the purposes for which those assets are used are (a) inherently limited to the same or very similar purposes as for the supplier and (b) comparatively straightforward to observe and specify²¹⁵. The value of data arises, in part, because the purposes to which it can be applied are both many and unpredictable. As we discuss below, a regulator is likely to be required to specify the type of data that is to be shared and the terms under which that is to occur. In so doing,

²¹² Cremer et al (2019), p.101 "We have already expressed some hesitation to bring data requests by claimants under Article 102 TFEU who pursue business purposes that are essentially unrelated to the market served by the dominant firm (see above). The main focus, under Article 102 TFEU, should rather be on data requests with the purpose of serving complementary markets or aftermarkets – i.e. markets that are part of the broader ecosystem that the data controller serves."

²¹³ Krämer, Schnurr, Broughton Micova (2020), Section 4.2.1

²¹⁴ Ibid p.75

²¹⁵ Indeed, some regulators have considered 'end user pricing' or 'retail minus' regimes in which the wholesale price of the regulated input is determined by reference to the use to which it is then put.

the regulator can certainly influence the purposes to which it can then be applied. But we do not think it feasible to suggest that a regulator could dictate the purposes to which data should be put.

5.1.3 The types of data to be shared, the geographic scope from which it is drawn, and the conditions under which sharing can occur, including on whose initiative?

Having determined which platforms are obliged to share data and which to receive it, the next set of issues relating to the type of data to be shared. As we explained in our introduction to this report, this is an area of some complexity and the potential for confusion, partly arising from ambiguities as to the legal treatment of some types of data and partly due to the lack of a generally-accepted taxonomy to distinguish between different types of data²¹⁶. Following Krämer, Schnurr, and Broughton Micova, we have found it useful to focus on two broad categories of data when it comes to thinking about data sharing to promote contestability in digital markets which, like them, we refer to as the **sharing or porting of individual user data** (relating to specific users) on the one hand and the **bulk sharing of aggregate user data** on the other²¹⁷.

5.1.3.1 Data relating to specific individual users

The first category is what we refer to as data relating to specific individuals, which will generally consist of **personal data that has been provided by the user and observed data that is then generated by the interactions between the user and the platform** in question²¹⁸. Inferred data, which we also discussed in the introduction, is also data that relates to specific individuals but for which, in our view, there is less likely to be strong arguments on competition grounds to require to be shared. This is because:

- We consider that it is this capacity to derive new insights from data that may have been acquired for other purposes is which drives the complementary innovation in adjacent markets which we think the sharing of data relating to specific individuals ought to be seeking to promote. Allowing access seekers to simply obtain existing insights from the incumbent platform by requiring it to share inferred personal data derived from the core market would, therefore, undermine the incentives of both parties to the arrangement²¹⁹.
- Insights that are derived for another purpose or from interaction in another market may not be particularly relevant to the complementary activities which we are seeking to promote (but may instead be more relevant to entry and direct competition in the core market from which, and for which, those insights have been derived).
- Restricting the sharing of data relating to specific individuals to provided and observed data is consistent with our proposal that the number of platforms that would be obliged to share such data should be quite large (i.e. the 'gatekeeper minus' group). The more extensive the scope of the data to be shared, the narrower the scope of platforms that might reasonably

²¹⁶ "A new taxonomy of data is badly needed. Industry, government and citizens are too frequently in disagreement as to what exactly constitutes personal data and what does not – and without an understanding of how data get positioned in each category, or flow between them, it is impossible to have a discussion about how to govern and regulate those flows", OECD 2019, p.28

²¹⁷ Krämer, Schnurr and Broughton Micova, 'we therefore suggest using two types of data access and sharing remedies in concert. The first type of remedy is to facilitate access to broad user raw data for third parties. This can only be achieved by bulk sharing of sufficiently anonymised raw data. Such data will therefore generally lack depth but is in terms of breadth representative of the raw user data that the original data controller has access to. The second type of remedy... is to facilitate access to deep user data. This data contains personally identifiable information, or at least allows traceability of an individual. Such data cannot be shared in bulk but requires the consent of each individual data subject anytime it is shared with a third party. Thus, the sum of data that is shared in this way generally lacks breadth, because it is unlikely that a sufficiently representative sample of users will consent to data sharing for a given third party.' p.88

²¹⁸ For the reasons given in the introduction, we consider that restricting access to provided data, which the individual user may be able to replicate themselves, is unlikely to ever be sufficient as a data sharing measure to promote competition. It is access to the observed data, which the individual user cannot replicate, which is of critical importance.

²¹⁹ We therefore agree with Prufer (2020), who says: 'If such data would be required to be shared, it might facilitate free-riding of smaller competitors and crowd out the dominant firm's incentives to invest into analytics in the first place. If only raw data are shared, it also incentivizes competitors to develop own analytics techniques, which can lead to a plurality of approaches, differentiated products, and, hence, more choice for consumers.', p.11

be obliged to do so. In this case, the scope of the data is limited both because it consists only of provided and observed data and because each transfer will relate only to a specified individual.

However, even if the scope of the data to be shared is confined to provided and observed personal data, as we propose, various legal and practical challenges remain²²⁰.

(i) Individual user consents

As explained earlier, this data may be 'controlled' by the organisation concerned but it is not considered under current European law to have acquired the data from the user in the way that it might acquire other assets and the user is not considered to have relinquished their rights over the data simply by sharing it²²¹. Rather, the user is said to have granted the organisation 'usage rights' to exploit the data for certain purposes, generally on the presumption that the user will benefit from it doing so. In this view, **if data is now to be shared with another platform, the individual user must give their prior consent to the transfer**. As explained in section 2, they are only likely to do this if they consider that the potential benefits from doing so outweigh the costs.

The need for individual user consents may be relatively clear concerning data that refers only to the individual user in question but soon becomes more challenging in the case of data which relates to several identifiable individuals, as with records of banking transactions between different persons, address files containing the contacts of many people, or photographs in which several people have been tagged. In the first two instances, European privacy regulators, as represented by guidance issued by the European Data Protection Board, appear to consider that the consent of the individual making or receiving the payment or holding the contacts is sufficient to initiate the transfer. The position as regards the latter scenario (and many others) remains less clear and appears to depend on whether the interests of the other individuals might be harmed by the transfer²²². That might be the case in some instances and not in others, but it will be difficult for a regulator, or the platform that is providing access to data, to assess these risks in advance (and impractical to do so on a case by case basis). 'Harm' in this case is also far from straightforward and could, presumably refer to losses in privacy but also loss of economic rights that might otherwise be conferred by intellectual property law (e.g. in the creation of a picture posted on another user's account).

It will therefore be very important for **any arrangements involving the sharing of data relating to individuals to seek to minimise the potential harms to other individuals which might otherwise arise from the sharing of data which is personal to more than one individual** since it will be impossible to seek or co-ordinate multiple consents before the data is shared²²³. Some of the other issues we discuss in this section – such as rights to redress and security standards – are intended to contribute to the minimisation of harm. It will be particularly important that individual users who consent to the sharing of their data do not feel exposed to risks that might otherwise arise from claims made by other individuals under, for example, the GDPR. Otherwise, there is a significant risk that individual users will be reluctant to consent to the sharing their provided and observed data

²²⁰ We are aware that Krämer, Schnurr and Broughton Micova also suggest that 'Only data that was created as a *by-product* of consumers' usage of a dominant service may have to be shared (e.g., search queries, likes, clicks, or location); but not (volunteered) user data that represents the essence of the service itself (e.g., documents uploaded to a cloud storage provider, posts on a social media site, customer reviews on a reviews' site, or GPS data from a geo-tracking app'. p.89. We recognize the point but prefer to adopt the conventional distinction between volunteered and observed data in our discussion. It seems to us that their concerns might be addressed by restricting access to certain types of data, but might also be addressed when determining which platforms are obliged to share data, or when determining the prices at which access to data is to be provided.

²²¹ OECD 2019, p.100-3

²²² Egan (2019), p.12

²²³ Guidelines of 13 April 2017 of Working Party 29 on the right to data portability, WP242 rev.01 states that: 'the data controllers should implement consent mechanisms for other data subjects involved, to ease data transmission for those cases where such parties are willing to consent, e.g. if they also want to move their data to some other data controller. Such a situation might arise, for example, with social networks, but it is up to data controllers to decide on the leading practice to follow.'

with other organisations and the measures will be rendered ineffective. The experience of personal data sharing in Europe to date (or the lack thereof) suggests this could be a significant barrier, although we have not seen evidence or research to date which would allow us to fully understand precisely what those barriers might be²²⁴.

(ii) Other approaches to consent: opt-out arrangements

Thus far, we have assumed that the sharing of personal data about a specific individual would require the prior consent of that individual, as is required under the current GDPR. This has led us to focus on potential barriers to the provision of consents by individuals, such as concerns about liability to third parties or other risks that an individual may feel they will be exposed to. Some of these are considered further later in this section.

We should consider the possibility, however, that **even if the various other measures that we propose to reduce perceptions of risk are taken, individual users will still not be prepared to consent to the sharing of their data, even if we could be confident that they would stand to benefit from doing so**. This may be because individual users will find it very difficult to assess and value the potential benefits which they might obtain from sharing data which relates to them or may be susceptible to well-known biases such as loss aversion. The innovations which might be enabled by data sharing may not be familiar to them or may depend upon their first making their data available. Individual users may become more confident once they see other individuals sharing their data or once they see the benefits that others are obtaining from doing so. However, there is an obvious hold up problem which may mean that any data sharing remedy which relies on prior consents from individual users could never achieve sufficient critical mass to justify the costs of implementation, or to have meaningful consequences for competition in digital markets. We take this possibility seriously.

This has led us to investigate other approaches to consent for data that relates to individuals, perhaps to kick start the data sharing process. One approach, which we discuss below, would involve the adoption of an **'opt out' mechanism as a basis of obtaining the consent of an individual user, or a group of users, for the transfer of data that relates to them**. We note that the opt out approach could be adopted either in relation to transfers of data which relates only to a specific individual, or could be employed for the bulk transfer of large volumes of aggregated data, thereby avoiding the need to otherwise anonymise the data. On the other hand, it may also require amendment of the GDPR. We discuss the transfer of aggregated data later in this section.

There are already **instances where personal data has been transferred between organisations without every individual user having to provide their consent**. For example, such arrangements have sometimes been employed by firms that are to be **acquired** by another firm, with the result that very large volumes of personal data are then be transferred in bulk from one corporate entity to another²²⁵. This is relevant since, in theory, a digital platform can be thought of as having several different means by which it could acquire data relating to individuals:

- It could develop and provide services to individual users to acquire both their provided and observed data;

²²⁴ On whether GDPR promotes or inhibits competition in digital markets more generally (to which they answer in the affirmative), see Gal and Aviv (2020). The focus of our report is different, since we presuppose a shortfall in competition and are then consider the impact of the GDPR on the type of data access measures which might be adopted to remedy such a shortfall.

²²⁵ Ilan et al (2016) note: 'In 2001, the French DPA declared (in the context of a merger of three companies) that personal data files may only be assigned or made available to a third party on the condition that data subjects be given advance notice as well as the right to object to such transfer. In Germany, it is necessary to provide notice of the transfer in the context of the transaction with a deadline to object where the transferred data goes beyond so-called "list data" (name and postal address). The Bavaria DPA issued fines to a buyer and target in an asset deal in 2015 where customer data was transferred without the parties providing the customers with a deadline to object to the transfer prior to the transaction' footnote 10

- It could acquire the same data by acquiring another firm which already controls it, or by acquiring the data assets of that firm or
- Under certain conditions, it could acquire the same data using regulated data sharing arrangements of the kind we are considering in this report.

We can see a strong case for requiring 'opt in' consents when a user is first asked to provide data and to engage with a new service, as in the first case listed above. However, we also see a case for applying a different principle when an existing service (including privacy terms) will continue to be provided to the individual user, but the shares of the company providing them are to be acquired by another entity. In the case of such mergers, European privacy regulators appear to have generally accepted either that the transaction can proceed without any prior communication with or consent from individual users – presumably on the basis that there is a reasonable expectation that the acquiring firm will continue to supply services on the terms to which they have already consented and would require further consents if those terms were subsequently to be varied²²⁶. Instead, the focus of regulators (and the legal advisers to the parties) appears to have been directed towards identifying any residual liabilities for past breaches of privacy laws by the acquired firm and we have been unable to identify any cases since the adoption of the GDPR in which parties to a merger have been prosecuted for failing to obtain appropriate consents from users before the conclusion of the transaction²²⁷.

Of greater relevance to our case are those where a **firm's data assets have been separated from the rest of the business and sold independently**. This is the closest analogy to the regulated data sharing arrangements which we are considering in this report and which can be considered as attempting to mimic outcomes which might otherwise arise in a competitive market in which data assets were being sold by a willing seller (although in practice, it appears that data assets are generally sold under a distressed sale or bankruptcy proceeding).

The Federal Trade Commission has considered several such cases:

'The FTC often cites a settlement it reached with internet retailer *Toysmart* in 2000 (the "Toysmart Settlement") which allowed Toysmart, after it ceased operations, to transfer customer personal data to a third party despite its privacy policy stating that such personal data would "never be shared with a third party." The FTC had sued to block Toysmart's sale of its customer database, alleging a violation of Section 5 of the FTC Act. Under the Toysmart Settlement, Toysmart was able to sell the customer data, but: (i) not as a stand-alone asset; (ii) only to a purchaser engaged in substantially the same lines of business as Toysmart; and (iii) only to a purchaser who agreed to be bound by and adhere to the terms of Toysmart's privacy policy and to obtain affirmative (opt-in) consent from consumers for any material changes to the policy that affect information collected under the Toysmart policy (hereinafter, the "Toysmart Principles"). As an alternative to the Toysmart Principles, the FTC proposed (in the *RadioShack* and *Borders* cases, discussed below) requiring the target to obtain affirmative (opt-in) consent of the data subjects to the transfer of the data to the purchaser and to purge the data of those who did not consent.'²²⁸

²²⁶ As Ilan et al note, this became a source of controversy after Facebook acquired Whatsapp in 2014 but subsequently sought to change that latter's privacy policy in 2016. The US FTC subsequently issued guidance that acquiring companies are expected to honor the privacy promises that have previously been made to users, see FTC (2015)

²²⁷ The only merger-related fine we have identified in Europe was proposed by the UK Information Commissioner on Marriott International in 2019 in relation to pre-merger security breaches by Starwood Group, a hotel chain which it had acquired in 2016. Marriott discovered the breach in 2018 and was fined £99 million, see UK Information Commissioner (2019).

²²⁸ Ilan et al (2016), op cit

Although US cases, the principle that data might be transferred from one organisation to another without seeking user consent provided that the acquiring party 'engaged in substantially the same lines of business' and 'agreed to be bound by and adhere to the terms of Toysmart's [the acquired party] privacy policy' is potentially a useful one for our purposes. For example, **it might be possible to consider arrangements for the sharing of personal data relating to a specific individual or group of individuals without requiring their consent provided the access seeker concerned was 'engaged in substantially the same lines of business' and was required to adhere to similar terms concerning privacy.** The regulator could presumably require that platforms seeking access to data comply with certain common privacy principles or policies to obtain the benefit of any such 'opt out' provisions.²²⁹ On the other hand, we also recognise that there may be significant practical difficulties with such arrangements, including the challenges of defining and restricting the uses to which data is to be put (which we discussed above) and issues with the reconfiguration of data sharing arrangements if individual users were subsequently to exercise their right to withdraw their implied consent.

We conclude that **European policymakers should give serious consideration to changes to the GDPR that would create a new type of 'opt out' regime for the sharing of personal data to promote competition under certain (restrictive) conditions.** This could have several features:

- 'Opt out' consents could be limited to the provision of data about specific individuals for a limited period whilst competition in a new adjacent market is 'seeded'. After this period had expired, further transfers of personal data of specific individuals could require their consent in each case.
- 'Opt out' consents could be conditional upon the recipients of the data complying with certain conditions, both concerning the uses to which they would put the data (whilst noting the challenges we identify above) and concerning privacy and other matters. Firms that were unwilling to meet such conditions would instead be required to obtain their data after first having acquired the consent of each individual. This could provide a strong incentive for compliance.
- 'Opt out' arrangements might also be used to facilitate the transfer of large volumes of aggregated personal data without the need for individual user consents in those circumstances where anonymisation was found to be impractical, risky or would serve to inhibit the effective promotion of competition. We consider this alongside other ways in which the sharing of aggregated data might be undertaken in the next section

In its European Data Strategy, the Commission recognises that certain aspects of the GDPR may require revision to accommodate the pursuit of other objectives, potentially including the more extensive sharing of data to promote competitive markets, something which may not have been fully anticipated when the GDPR was being developed before its adoption in 2016. We consider that if data sharing measures to promote competition are to realise their potential, policymakers will indeed need to revisit some of their assumptions about the circumstances under which opt out arrangements can be applied²³⁰.

²²⁹ We note that 'opt in' was also the approach adopted by the UK Competition and Markets Authority (after consultation with the UK privacy regulator) in transferring data about vulnerable customers of energy suppliers to Ofgem, the energy regulator, who would then be expected to transfer the data to rival suppliers at their request, Competition and Markets Authority (2016) para 11.64.

²³⁰ However, we note that this does not appear on the list of Commission actions arising from the first review of June 2020 of the implementation of the GDPR, COM (2020) 264

Incumbent platforms might be expected to emphasise the potential risks to users' privacy being unreasonably compromised if opt out arrangements were to be adopted. This to be a legitimate concern,²³¹ but it has to be balanced against the economic and other risks which users may also face if digital markets were otherwise to remain uncontested.

5.1.3.2 Aggregated user data

The second category of data which we have identified as being of greatest relevance for the promotion of competition in digital markets is aggregated data that is derived from observed interactions between very large numbers of users and the digital platforms in question. This data is very important for the development of new insights about how individual users might respond to new services in adjacent markets or to changes to existing services. In general, the capacity of a platform to generate such insights will improve as the volume and variety of the data set increases. Such data facilitates improvements in product quality and innovations which are likely to benefit all users of the platform irrespective of whether data which relates specifically to them has been included in the data set or not.

If aggregated observed is to be shared, several scenarios might be considered:

- The first would involve the **sharing of aggregated personal data, in which the identities of individual users are retained in the data**. This can be thought of as sharing the same data as the incumbent platform itself has access to, and which is, therefore, likely to yield the maximum competitive benefit for the other access seekers. However, it will be clear that the sharing of such data is problematic, and likely impossible if the consents of each user contributing data to the data set were to be required before anything could be shared. A potentially more feasible alternative in these circumstances is to rely on 'opt out' arrangements of the kind described above, although this may still involve significant transaction costs if large numbers of users decide to exercise their option to opt out and will involve delays whilst users are notified of their rights to do so and given a reasonable opportunity to respond (although under the continuous data sharing arrangements that we envisage, we assume there would also be a right to opt out at any time thereafter).
- The second scenario would involve the **prior anonymisation of the data before it is shared**. This also presents some challenges, since it is not always clear that robust legal boundaries could be drawn between data that is personal and that which is anonymised²³² (i.e. cannot be reconfigured to reveal the identities of individual persons)²³³. More importantly, it is likely that something of competitive significance will be lost in the

²³¹ See Egan/Facebook (2019), op cit, or Google (2020) argues, in its submission to the CMA on the digital advertising interim report, 'aside from the risk of a data breach, the very fact of us sharing query data with third-parties could do irreparable harm to our reputation. Users trust us to treat their queries appropriately. Handing over those queries to third parties - especially if this is done for money - may cause users to lose confidence in their ability to search privately with us.', p.20

²³² Cremer et al (2019) note: 'From a legal perspective, the precise requirements for data use to be qualified as anonymous for the purposes of the GDPR have not yet been fully clarified by the EU courts.', p.87. The Article 29 Working Party published an opinion on anonymisation techniques in 2014, concluding: '... anonymisation techniques can provide privacy guarantees and may be used to generate efficient anonymisation processes, but only if their application is engineered appropriately - which means that the prerequisites (context) and the objective(s) of the anonymisation process must be clearly set out in order to achieve the targeted anonymisation while producing some useful data. The optimal solution should be decided on a case-by-case basis, possibly by using a combination of different techniques, while taking into account the practical recommendations developed in this Opinion. Finally, data controllers should consider that an anonymised data set can still present residual risks to data subjects. Indeed, on the one hand, anonymisation and re-identification are active fields of research and new discoveries are regularly published, and on the other hand even anonymised data, like statistics, may be used to enrich existing profiles of individuals, thus creating new data protection issues. Thus, anonymisation should not be regarded as a one-off exercise and the attending risks should be reassessed regularly by data controllers', EDPB (2014), p.3

²³³ The OECD (2019) warns: 'However, developments in data analytics (and AI) combined with the increasing volume and variety of available data sets, and the capacity to link these different data sets, have made it easier to infer and relate seemingly non-personal or anonymised data to an identified or identifiable entity, even if the entity never directly shared this information with anyone' p.85, see also Cremer et al (2019), p.78

process²³⁴. Indeed, one of the primary motivations behind the technological arms race that exists between those that seek to anonymise data and those that seek to frustrate such efforts arises from the fact that de-anonymised data is likely to have a greater commercial (i.e. competitive) value.

- The third scenario would not involve the transfer of data from one platform to another, but would **involve the platform instead using the data they retain to answer queries from other firms or allowing other firms to train their algorithms on the data set (without their obtaining access to the underlying data itself)**. In earlier sections, we provided examples of cases where firms such as Telefonica and Microsoft retain aggregated data sets within their organisations, likely for both commercial and privacy related reasons, but allow third parties to interrogate that data and derive insights from it. The basis on which such access is provided is currently determined by the holder of the data, rather than by a regulator.

There are challenges with each of these approaches. The consequence of an anonymisation requirement would be that access seekers obtain access to data that is less competitively valuable/more anonymised than the data that is retained by the platform that supplies it, even if both parties have access to similar volumes of data. Whether the resulting asymmetry means the sharing of anonymous data would prove to be an ineffective remedy to promote competition is difficult to predict without examining the specifics of the case. It may be that regulators would start by requiring the sharing of anonymous data, but consider the other approaches if that proved to be ineffective (although the regulator will also need to be mindful of the risks of strategic behaviour by the parties involved)

We should also consider the possibility that obligations to share anonymous data may introduce perverse incentives for the holders of personal data, who may become more reluctant to undertake the process of anonymisation even when it might otherwise be appropriate to do so and when this would contribute to the privacy of users of the platform. Thus, it may be necessary for regulators to require digital platforms to create and retain both original and anonymised versions of the same data, with the former being required to fulfil the needs of its users and the latter being required solely to fulfil requests for data access from other organisations (although in practice, we would also expect that digital platforms may generate multiple data sets with common inputs for different purposes of its own)²³⁵.

We discussed the challenges associated with 'opt out' arrangements earlier in this section²³⁶. **The third approach, which involves the retention of the data by the incumbent platform but the obligation for them to facilitate its interrogation by potential competitors seems in many ways the most attractive in theory.** However, the ability of the incumbent platform to control the interfaces through which the data is interrogated (and to itself have access to the insights which its competitors might derive from the interrogation of the data) means that these arrangements involve a much greater degree of interaction (and potential for conflict) between the

²³⁴ Note that this does not imply that non-personal data is of lesser value than personal data. Many forms of non-personal data, such as intellectual property, have very great commercial value. The point is rather than data which relates to identifiable persons is likely to be more valuable if the identifiers are present in the data, than if they have been removed.

²³⁵ Krämer, Schnurr and Broughton Micova raise a separate concern, namely that de-anonymisation of data that has been properly anonymized would then compromise the privacy of individuals. They propose legal sanctions against those who attempt to de-anonymised shared data sets (without saying whether those sanctions would be applied by the regulator that facilitates the sharing of the data, by the data protection authority or by some other body), p.92. This would complement an obligation on the part of regulated platform to anonymise aggregated data before it is shared, an obligation which we would expect to be enforced by the regulator.

²³⁶ The OECD also note that 'opt out' arrangements allow user to retain a degree of control over their data which anonymization does not: 'A bigger concern might be that mandating FRAND access without an opt-out option for users would remove the ability of a consumer to exclude other's from accessing their data and hence reduce the value of a consumer's data, in effect transferring that value from consumers to firms (with consequences for income distribution).', OECD (2020)

access seekers and the provider of access to data. Amongst other things, this is likely to mean that the regulatory overheads associated with overseeing such arrangements would also be considerably higher. One potential way to reduce this would involve transferring the data to a 'neutral' or independent entity, with which both the access provider and the access seekers would then interact on equal terms²³⁷. However, this such 'structural separation' of the data assets is likely to be a very complex and difficult arrangement to implement and enforce and may encounter the same requirements for user consents and other barriers that we have identified with the other approaches²³⁸.

5.1.3.3 Other aspects relating to the scope of data to be shared

(i) Heterogeneity of requirements from potential entrants

It seems likely that those seeking access to data to support their entry into adjacent markets and/or to facilitate innovation are likely to have a wide range of different requirements when it comes to the volume and nature of the data to which they seek to access. These requirements may also change over time. The arrangements for sharing data will need to reflect this.

In this report, we have proposed two broad categories of data that might be shared. The first is provided and observed **data that relates to specific individuals**. This data is likely to be relatively limited in volume and straightforward to define, although the volume and variety of observed data acquired by large digital platforms can still be very considerable indeed. Our initial view is that **all user data falling within this category would be required to be shared from the outset** unless the incumbent platform could provide good reasons as to why some data might be withheld or provided separately²³⁹. We would also anticipate those data sharing arrangements for data relating to individuals would likely be implemented before arrangements were fully in place to support the sharing of aggregated data (although, as we explain elsewhere, we would expect there to be synergies between the two sets of arrangements, particularly as regards common technical interfaces, standards, and the physical infrastructure to support sharing).

The position with **aggregated provided and particularly observed data** is more complex since the volume of data that could potentially share could be overwhelming. Many new entrants would likely be unable to store, let alone process, the data sets that have been acquired by the world's largest digital platforms (which may also be a reason to favour the 'data interrogation' model for the sharing of aggregated data, in which the underlying data is retained on the infrastructure of the incumbent platform). **Some degree of disaggregation may therefore be required to make the data usable for competitors**. It is common in other contexts for the access provider to be required to develop a menu of options or a 'reference offer', often in collaboration with access seekers and under the oversight of the regulator, to define the different types of assets that might be shared, and the conditions under which each will occur. Although regulators may be tempted to allow platforms to themselves come to commercial agreements about the scope and other characteristics of the data to which access is to be provided, experience from access arrangements in other sectors suggests that agreement will not always be forthcoming²⁴⁰. Data holders and potential competitors will have conflicting interests, which are likely to lead to disputes as to what and how assets should be shared. But potential competitors are also likely to disagree amongst themselves. It seems likely to us that the nature of digital data, which can be easily manipulated and reconfigured, should mean

²³⁷ This approach is also proposed by Prufer (2020)

²³⁸ For some of the challenges, see Krämer, Schnurr and Broughton Micova, p.91

²³⁹ As noted previously, Krämer, Schnurr and Bourghton Micova argue that data that is volunteered by users such as content or GPS data should be withheld.

²⁴⁰ The fact that the participants in the Data Transfer Project are co-operating to share data suggest to us that the question of the scope of data to be shared has already been resolved by the GDPR in that case and/or that the participants do not see the arrangements arising from the DTP as having great competitive significance. That said, we have no insights into the extent to which there have or are in fact differences between the participants on such matters.

that it should be much easier for access providers to serve a wide range of different requirements than is often the case when access is being provided to physical assets such as cables or pipes.

(ii) Geographic scope

Two further questions arise in this context. First, it would appear unsatisfactory for an incumbent digital platform to be obliged by different regulators to supply different types of data in different Member States if the same competition concerns were to have been identified in each instance. That suggests to us either that the European Commission should either ensure co-ordination amongst national regulatory authorities or that the determination of the data to be shared and the menu to be offered should rest with the Commission itself, or a designated agency, and should **apply on a pan-European basis**.

Second, a question arises concerning the sharing of aggregate data by digital platforms which may operate both within Europe and beyond, as many do. The data which a platform may be required to share at the request of an individual user will be limited to and determined by the identity of that individual user. The request of an Italian user of Google to transfer their data under the GDPR cannot result in the transfer of data that Google has acquired from its operations in the United States, or likely even in another Member State. However, the same considerations may not apply in the case of aggregated data, where the geographic scope of the data which may be required to be transferred could be:

- Data acquired from users, or relating to users, to whom services are provided within a specific Member State;
- Data acquired from users, or relating to users, throughout the entirety of the European Union;
- Data acquired from users, or relating to users, to whom services are provided both inside and outside of the European Union.

The broader the geographic scope of the data to be transferred, the greater the volume of data and, potentially, the greater the variation in it. We noted earlier that both volume and variation may be sources of competitive advantage, and markets may be more contestable if data to which access is obtained has been acquired across a wider geographic market or larger population. On the other hand, user preferences and behaviour may also vary significantly between the Member States or between regions, and more localised data may have a higher value than data derived from elsewhere. We have no view on the appropriate geographic scope of the data that a platform might be obliged to transfer, but a regulator would need to determine what it is. If this were to include data that was acquired from the provision of services to users outside of the European Union (but which was necessary to facilitate competition within the European Union), then European policymakers may need to consider how such obligations would sit alongside privacy or other regulatory provisions in other regions.

5.2 The conditions of sharing

There are several other matters where the intervention or the co-ordinating role of a regulator will be required. Ctrl-Shift (2019) presents the findings of a series of personal data sharing trials which were undertaken in the UK in 2019 provides a useful overview and includes, in addition to the issues already discussed:

- The need for a clear definition of liability when data is transferred from one organisation to another;
- The need for the transfer process to be easy for the user so that the cost – in the time taken and cognitive or physical effort – is low enough to ensure that adoption reaches its full potential;

- The need for common standards and processes for revoking data sharing so that users can be confident data sharing has ceased;
- The need for education so that users are confident about both their rights and the benefits they might obtain from authorising the sharing of data;
- User-friendly verification, authentication and access control processes – not passwords;
- The possibility of extending data access between many different sectors of the economy, to realise benefits from complementary innovation and synergies.

Many of these issues are of greater relevance to data relating to individual users than to the sharing of aggregated data, but some apply to both. We consider some of them in more detail below.

5.2.1 Technical standardisation

As noted earlier, the evidence as to what makes for effective data sharing and what has inhibited sharing between organisations in the past²⁴¹ suggests that collaboration is required amongst a significant number of the parties that are likely to be required to provide data access, or which might stand to benefit from it, to first develop a common set of standards to govern how it should work. 'Standards' in this context relate both to the external interfaces or APIs which organisations develop to transfer data between them, and to the models which govern how data is structured and labelled ('the metadata') so that it can be understood but also so that it can be retrieved and analysed.²⁴²

²⁴¹ Deloitte (2016), Ctrl-Shift (2019)

²⁴² Benson (2009) explains: 'Data is defined as the "symbolic representation of something that depends, in part, on its metadata for its meaning.'" It follows therefore that the quality of the metadata must play an important part in determining data quality. Metadata gives data meaning. For example "50-02-01" is a meaningless string of characters but apply the metadata "Date of Birth" and it becomes meaningful data. To make it unambiguous we need to have syntax such as CCYY-MM-DD and the associated value becomes 1950-02-01

As the OECD explains:

'Even when commonly used machine-readable formats are used for accessibility, interoperability is sometimes not guaranteed. These common formats may enable "syntactic" interoperability, i.e. the transfer of "data from a source system to a target system using data formats that can be decoded on the target system". But they do not guarantee "semantic" interoperability, "defined as transferring data to a target such that the meaning of the data model is understood". Both, syntactic and semantic interoperability are needed. Besides being accessible and interoperable, data need to be findable. This may require that data be catalogued and/or searchable.'²⁴³

Without giving thought to these issues, there is a risk that the first digital platform to be obliged to transfer data will develop the APIs and other standards to do so without consideration of the interests or views of others who may be required to do so subsequently. Faced with an obligation to share data, a platform is likely to seek to minimise both its costs and disruption to existing business practices by proposing that others adopt the proprietary standards in which it has already invested. The incumbent platform would, in doing so, become the standards setter for the rest of the industry, both at the time at which data access is provided and likely for subsequent changes to the standards. Alternatively, different standards may emerge in different contexts, which could create significant entry barriers for firms that seek to obtain access to data from several platforms and raise costs for all parties.

There may be benefits from ensuring that standards develop in a way that will allow for their application across a wide range of markets beyond those where there are currently concerns about large digital platforms. This is the model being adopted by the Australian Government in their implementation of Customer Data Rights, where it is envisaged that technical standards which are initially being developed to support access to data held by deposit takers and other financial institutions will subsequently be used to support access to data held by utility companies and other retail organisations. Such '**anticipatory' standards-setting** seems particularly relevant to large digital platforms, whose breadth of activities may raise concerns that require data access remedies is potentially a very wide range of different markets – including those involving data acquired by the Internet of Things devices, but also healthcare, financial services and many other parts of the economy in ways which are difficult to anticipate today. Digital platforms (and regulators) should not find themselves in the position of having to develop a new set of technical standards each time data access is required to promote competition in a new market.

Experience suggests that the **development of technical standards is best regarded as a process rather than being a discrete event. We would expect any regulator to play an important role in convening the technical forum in which common standards for APIs and data models would be developed in a manner that fairly balances the interests of all parties**, and ensuring that there is an appropriate representation of interests without the process becoming unmanageable. This process could seek to build upon work that has already been undertaken by the Data Transfer Project since this already involves a number of the global digital platforms who might be expected to be subject to obligations to share data in the future (although the regulator would need to ensure that all interests are properly represented and that the resulting outputs do not enable incumbent firms to impose unreasonable costs on others). Standards developed for Open Banking or other data access arrangements may be of some relevance, but we consider that the variety and volume of the data that could potentially be within scope for data sharing by large digital platforms are likely to be orders of magnitude greater than anything that has applied to banks or other service providers to date. There are other potentially relevant initiatives

²⁴³ OECD (2019), p.93

to consider, such as ODPI, supported by the Linux Foundation, which promotes open-source standards to facilitate data sharing and 'is committed to simplification & standardization of the big data ecosystem with common reference specifications and test suites'²⁴⁴, or OAUTH, which develops standardised authorisation protocols²⁴⁵. Regulatory oversight will be required to ensure that standards facilitate market entry and contestability and can be readily implemented by new entrants as well as incumbent providers.

Experience also suggests that standards and processes will need to evolve as the scope of data to be transferred, and the purposes to which it is applied, develop over time. Participants, particularly the large incumbent platforms, will require time to reconfigure their existing proprietary models and adapt to and converge upon new industry standard data models and processes. Deadlines will need to be set and sanctions applied if they are not met. Regulators will need to assess claims that the costs of doing so are prohibitive or should be recovered from access seekers rather than being met by the incumbent platform themselves (we consider pricing issues later in this report). Interim arrangements may be required, such as extracting data in open formats such as JSON or CSV (i.e. without the metadata) until the proprietary data models that are employed by today's digital platforms can be adapted to the new standards, or by using 'adaptors' or 'translators', as the DTP proposes to do²⁴⁶. Not everything may be possible at the outset, and so the conflicting needs of different access seekers may have to be prioritised and adjudicated upon.

We explained earlier why we think the issues of participation in the development and adoption of common standards is held separate from the question of whether a particular digital platform is subject to obligations to share data. We would expect that all of the digital platforms that participate in the development of standards would be required to provide access to data relating to specific individuals, and would be required to adopt common technical standards to do so. But the question of whether they would be obliged, for example, to share aggregate data, and in what circumstances, is one that should be considered separately from the implementation of the standards themselves.

5.2.2 The data transfer process

In addition to ensuring that the recipient of data is certified and authorised to do so, it will be important to **ensure that any transfer of data is properly managed**. In the case of transfers of aggregate data that are initiated by another digital platform, this should be a relatively straightforward matter and, as noted earlier, many platforms will already have robust commercial arrangements to ensure that data is shared with authentic third parties securely. However, the evidence suggests that in the case of data transfers that require the prior consent of an individual user, the **process of authentication can be more challenging**.

The case of 'open banking' illustrates the point. When Open Banking was initially launched in the UK, users were redirected from the organisation seeking access to data to the relevant bank or banks, to confirm that they consented to the transfer. The bank or banks required users to authenticate themselves in the normal way, by using passwords and other authentication methods then employed. If a user wanted to ensure access to multiple accounts or sources of data, then they would often be required to re-authenticate themselves on each occasion. Joint accounts might require authentication from both signatories.

In this case, the UK banks appear to have had legitimate concerns to protect their users' interests and ensure that account details were not transferred to third parties without their consent or

²⁴⁴ <https://www.odpi.org/>

²⁴⁵ [Oauth.net/about/intro](https://oauth.net/about/intro)

²⁴⁶ For a detailed discussion, see Gal and Rubinfeld (2019).

knowledge, but they also had incentives to frustrate access to user data²⁴⁷. The processes they adopted proved so cumbersome that only very determined users were prepared to complete the process. Recognising this, the Open Banking Implementation Entity then directed the banks to implement 'app to app' authentication, which allowed users to consent to transfers by using fingerprint or facial recognition technologies on their smartphones. This had a material impact on the volume of data transfers that were authorised by users, and hence on the potential impact of the data access measures for competition²⁴⁸. However, we also note there are concerns that the requirement that users of Open Banking re-consent to the services every 90 days, as required by the Second Payment Services Directive, continues to represent a barrier to adoption.

Digital platforms such as Facebook and Google already offer their authentication services to third party platforms (Google Sign In) which allow their users to connect to those platforms without the need to re-authenticate²⁴⁹. Two large digital platforms, Google and Apple, supply the operating systems for the majority of the world's smartphones (and Microsoft and Apple for most PCs), and so the adoption of fingerprint, eye or facial recognition as a means of authenticating consents for data transfers ought, in our view, to be feasible provided these firms are involved in the process. Regulatory oversight may be required to ensure that it is implemented in a manner which both safeguards the interests of users and achieves the objective of promoting competition. Some commentators go further and argue that the '**digital identities**' of users ought to be administered by an independent third party (which, as noted, is a feature of some PDS providers), or even by Government bodies, rather than by the platforms themselves²⁵⁰.

User authentication is one aspect of the process involved in transferring data from one party to another, and one which we have seen presents an opportunity for an incumbent digital platform to frustrate the implementation of measures to promote competition²⁵¹. Experience of other measures, such as the implementation of number portability between telecommunications operators or switching between energy providers suggests that there will be many other opportunities for anti-competitive conduct. As noted earlier, a prerequisite for the effective implementation of data sharing measures, irrespective of the purpose, will be trusted on the part of the users and organisations who stand to benefit from it. Actions that create doubt or uncertainty about the reliability of the process, or the risks involved, will tend to favour the incumbent platform and reduce the volume of transfers that occur.

We note that several studies of data sharing arrangements that require the consents of individual users place emphasis not only on the ease of using the data transfer process itself but also on the need for policymakers or regulators to **educate and inform users about the benefits** of their

²⁴⁷ As has otherwise occurred in the past in cases where user authentication was much weaker, as was the case with the implementation of carrier selection in the early years of telecommunications liberalisation, particularly in the US. This process of transferring users without their consent was known as 'slamming', see <https://www.fcc.gov/general/slamming-policy>

²⁴⁸ Fingleton/Open Data Institute, op cit p.23: 'One company we spoke to, Account Technologies, experienced an approximate 60% increase in customer conversions from one bank's customers after it implemented the app-to-app standard, telling us: "the results were absolutely astounding." This approach is now mandatory for the CMA9.'

²⁴⁹ www.developers.google.com/identity

²⁵⁰ Experience of digital identities varies significantly between Member States. The UK has tried (and failed) to promote market-based solutions through the Verify.gov programme, initially assuming that banks would take the lead, see 'Implementing a 21st century approach to e-identity', Computer Weekly, January 2020. In contrast, Estonia has issued every citizen with an e-identity, see <https://e-estonia.com/solutions/e-identity/id-card/>. To date, 13 Member States have pre-notified at least one national e-identity scheme, in accordance with the requirements of the Electronic Identification Regulation (eIDAS), 2014/910. The argument for a 'business line restriction' for such 'ancillary services' is made by Krämer, Schnurr and Broughton Micova, see p.84-5

²⁵¹ Fingleton/Open Data Institute note that under the Second Payment Systems Directive, users are required to fully re-authorise their permissions every 90 days. Although ostensibly to reaffirm customer consents and retain customer control, this provides an incumbent platform with a periodic win back opportunity: 'The current PSD2 legislation requires a full reauthorisation every 90 days, which can make Open Banking products cumbersome for users and lead to user attrition for TPPs, increasing costs for them'. They suggest a cost benefit review is undertaken to assess the merits of this obligation.

doing so²⁵². Even if the benefits to an individual user, in terms of being able to switch between platforms or obtain access to complementary services, seem self-evident to regulators, many users may not be aware of them (as we discussed earlier when considering the case for 'opt outs'). Regulation may be required to ensure that digital platforms inform users of their rights or even to inform potential entrants of the opportunities that are available to them.

Even if this is done, the requirement that individual users must themselves initiate the transfer of certain data provides a potential opportunity for the incumbent platform to introduce incentives that would be intended to deter the user from completing the process. It is difficult to anticipate what form these inducements might take. Users might, for example, be targeted with offers of premium services at no charge or other inducements if they were to cancel the transfer. Or they might be presented with pop up notifications by the incumbent platform drawing attention to the risks of authorising third party access to their data²⁵³. Balancing the interests of users, who may stand to benefit from receiving such targeted offers and may consider the threat of initiating a data transfer as a means of restoring some power in the other unequal bargaining relationship they have with large digital platforms, to facilitate competition more generally will be a challenge. In the case of pop ups, the incumbent platform may argue that it has a legitimate responsibility to warn users of the risks when they authorise the transfer of data to a third party²⁵⁴. Given the competing interests, this may be another instance where a **regulator might convene an industry working group, composed of representatives of both those platforms with obligations to share data and those that hope to acquire access to it, to develop a set of rules and processes** to which all will then be required to adhere.

5.2.3 *Rights of redress and other contractual matters*

Even if clear rules and processes are governing the transfer of data from one party to another, users and organisations will still have concerns about accountability and redress if the process does fail in some way. This could involve a failure to comply, either at all or in part, with a request to share data, a delay in doing so, or a failure on the part of the data recipient to then manage the data appropriately or use it for the purpose for which it was intended. There is a wide range of potential scenarios and harms that might result. In complex transactions involving multiple parties, users or organisations will worry that failures result in each party assigning responsibility for the error or failure to the other. The individual user may find themselves in the middle of a lengthy and complex dispute, in which the large platforms will enjoy significant information advantages. Without some assurance that their interests will be protected, many will conclude that any benefits from data sharing are outweighed by the risks.

Similar considerations may apply for organisations, particularly new entrants. They may fear that they would be unable to obtain appropriate redress for non-performance or they may be uncertain about the potential risks and costs they might be exposed to from their users in the event of their own inability to provide services if they cannot maintain access to data (e.g. if access is disrupted or

²⁵² Ctrl-Shift (2018), p.12: 'Consumers have a lack of know-how and understanding of the digital market, and limited knowledge about their data, how it is used, and how they could use it. This makes the individuals vulnerable to abuse and lacking in the skills to access the opportunity'.

²⁵³ There have been many disputes about the use of pop ups with 'security warnings' being displayed to users seeking to change their default browser settings, for example. Similar pop ups could be displayed by incumbent platforms when receiving an authentication request from a user to provide data access to a third party, even if the third party had been authorized to obtain access by a regulator. Regulators may find themselves having to intervene to address such practices.

²⁵⁴ Egan/Facebook (2019) argues: 'Service providers might explore tools to help users understand security risks and protocols for their downloaded data. Providers could also consider giving users guidance on how to inspect recipient organizations for potential abuse or insufficient security safeguards. For instance, providers could teach users ways to confirm the authenticity of the recipient organization (that it is what it says it is); check the website security for recipient organizations (e.g., the difference between HTTP and HTTPS); secure their devices when they download data (e.g., not using public Wi-Fi when downloading data); and identify whether the recipient organization has appropriate policies in place (e.g., checking privacy policies to determine whether an entity will sell user data that it receives).', p.18

unexpectedly withdrawn)²⁵⁵. Under conditions in which a large digital platform is obliged to share data with potential competitors, there is likely to be an acute information asymmetry and bargaining imbalance between the organisation providing access to data and the organisation acquiring it. These need to be addressed by a **regulatory regime which clarifies the rights which access seekers have, and the processes they can pursue, in the event of non-compliance by the incumbent platform, but also one which clarifies the risks and potential costs to which the access seekers themselves are exposed if they fail to comply with their obligations.**

It seems clear that where a large digital platform is obliged to provide access to data to promote competition, the platform itself should not be in a position to dictate the uses to which the data is to be put, how it is to be managed or the penalties that might be payable in the event of non-compliance with such provisions. These are all issues that would normally be addressed by contract involuntary data access arrangements.²⁵⁶ Otherwise, the large digital platform is likely to be able to exploit its market power to impose conditions that would be intended to inhibit entry or otherwise restrict competition. At the same time, however, the incumbent digital platform may have legitimate concerns that data which it has acquired from users may be put to illegitimate or illegal uses for which it may incur significant reputational costs²⁵⁷. Some argue that a dominant platform, for competition law purposes, might have more onerous obligations concerning safeguarding the rights of its users than other firms²⁵⁸.

Regulators that are independent of the parties to the data transaction and who have strong investigative powers can be expected to play an important role in providing the dispute resolution processes and means of redress for both individual users and organisations, thereby building trust in the measures which are being taken. This will be easier in some cases than others. In the Open Banking case, for example, there were already well-developed processes under existing financial service regulation for dealing with instances where 'no fault' payments are made to fraudsters, or where funds need to be recovered. Similar arrangements will need to be devised from scratch for cases where data is incorrectly transferred between digital platforms, needs to be recovered or deleted, or where there is otherwise a security breach with harmful consequences for other parties. These obligations should extend beyond those provided for in the GDPR guidance if users and organisations are to be confident in sharing data. At the same time, risks and costs need to be appropriately allocated, and should not represent a significant barrier to entry. Having the regulator play a central role in the resolution of disputes and the redress of harms also allows the regulator to determine where there are improvements to be made in the transfer processes and to direct the parties to make improvements to reduce problems in the future.

A further issue may arise concerning any liabilities arising from certain types of data that may be hosted by digital platforms and that are currently subject to the provisions of Articles 12 to 15 of the eCommerce Directive²⁵⁹. These provisions are intended to protect digital platforms from claims

²⁵⁵ Deloitte (2016) identify this issue as a major barrier to (voluntary) data sharing between commercial organisations: 'Existing liability laws are based on the concept of tangible products. Companies cannot be sure whether they can have recourse to this legislation for databased products, so prefer to fall back on contractual liability on a case-by-case basis', although the primary focus of their study is data derived from IoT, robotics and autonomous systems.

²⁵⁶ See the model contracts described in Support Centre for data sharing (2019).

²⁵⁷ Egan/Facebook (2019) says: 'In our conversations with stakeholders so far, the general view about these questions has been that a transferring entity may—and should—impose some baseline privacy and data protection restrictions around transfers even when carrying out the transfer to comply with a portability request. But, as discussed below, questions remain about what kinds of conditions are appropriate. Restrictions along the lines of those we impose through Platform strike some as too restrictive to be consistent with portability. Our recent settlement with the FTC suggests that some regulators may view Platform-style transfers as distinct from portability transfers. Where the line is between these two categories will likely be the line between portability and other data transfers', p.10

²⁵⁸ Cremer et al say: 'Dominant firms may be subject to a particularly stringent data protection standard under both tests. In protecting consumer choice vis-à-vis dominant firms, competition law and data protection law can thus complement each other', op cit, p.80

²⁵⁹ Directive 2000/31 of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce), OJ [2000] L 178/1.

concerning data that is provided by users and which the platform may store or transmit (e.g. photos) and would be expected to apply to both the donor of data and any recipient. The European Commission proposes to revisit these provisions as part of its proposals for a new Digital Services Act. Any amendments would then need to be reflected in the liability and redress provisions for any data access arrangements between digital platforms.

5.2.4 Prices for data access

5.2.4.1 Prices and incentives

Any regulatory regime for the provision of access to data is likely to require rules governing the monetary terms on which access to data is being provided. The incentives upon an access seeker to acquire data through the regulatory process rather than engage in its efforts to acquire data by competing in a relevant market (or by acquiring the data assets of another firm) will depend, in part, on the respective costs of each activity. Experience suggests that one of the challenges in implementing regulatory arrangements that provide for access to valuable assets is that competition between firms in the commercial arena can quickly be displaced to competition between firms in the regulatory arena and the courts²⁶⁰.

The question of incentives is a complex one when it comes to data. As noted in the introduction, unlike many tangible assets to which regulatory access is required, data is non-rivalrous, meaning that its consumption by one party does not diminish the opportunities for others. The marginal costs of sharing data are also likely to be very low. Some argue that these considerations would support the sharing of large volumes of data at no cost to the recipient since the direct costs of implementing sharing arrangements may be very low whilst the potential benefits of sharing data are very high. This is why 'open data' initiatives, which involve the sharing of data held by Governments and other public authorities, are invariably undertaken without their being any charge levied for the data that is being shared.

Under the GDPR, no charges apply (except in exceptional circumstances) when an individual seeks access to data which relates to them and which is held by either a commercial or a public organisation because the data in question is not considered to be the exclusive intellectual property of the organisation which holds it or something which it can sell, but rather something which remains at all times under the control of the user²⁶¹.

We think the charging arrangements for data sharing to promote competition and innovation remain under-researched but are likely to prove quite complex. The primary concern, as noted above, is to ensure that both the incumbent digital platform and potential access seekers maintain incentives to engage in beneficial economic activities (for users of platforms) whilst data is being shared between them. The acquisition of data may be a by-product of these beneficial activities, as when banks that provide current account services acquire data about the spending patterns of users which can be used to sell them other financial products. In this case, the banks do not supply current account services with the primary aim of acquiring spending data and do not rely upon it to provide current account banking services (although of course, they use it when cross-selling other financial products and services). In contrast, digital advertising platforms such as Google and Facebook acquire user

²⁶⁰ Aside from acquiring data directly from users themselves, firms may also acquire data by acquiring other firms who have acquired data. The OECD (2019) report that 'Some of the largest M&As motivated by access to big data in the last five years include: Monsanto's acquisition of the Climate Corporation, an agriculture analytic firm, for USD 1.1 billion in 2013; IBM's acquisition of a majority share of the Weather Company, a weather forecasting and analytic company, for over USD 2 billion in 2015 (Waters, 2015[4]); and Alibaba's total investment of USD 4 billion between 2016 and 2018 to acquire Lazada, a leading e-commerce platform founded in 2012 in Singapore. Start-ups specialised in big data are also increasingly the target of acquisitions. The annual number of these acquisitions increased from more than 100 acquisitions in 2013 to more than 400 acquisitions in 2017, with the average price paid exceeding USD 1 billion in some quarters', p.16

²⁶¹ Scassa (2018) notes: 'Personal information is generally not capable of ownership — at least not by the persons to whom it pertains — although in recent privacy discourse, it is increasingly common to hear references to individuals "owning" their personal information', p.13

data primarily to support and improve the provision of targeted digital advertising services to their advertising customers and do so by offering a wide range of valuable digital services from which they otherwise derive no revenues. If they were unable to acquire the data, it is not clear why they would offer services to users for free, or at all.

In the banking example, an individual user's spending data is acquired by the bank at negligible marginal cost, assuming the bank would provide the current account services in any event and earn its revenues from sources other than the data. Thus, requiring the bank to share such data with potential competitors in other markets for no charge is unlikely to significantly affect its incentives to invest in the provision of current account services. There would appear, in other words, the little cost to the sharing of data (beyond the direct costs of implementing the arrangements) in terms of investments in the market from which the data is being derived. We understand that no charges are levied by the banks for account data under the Open Banking arrangements.

Conversely, allowing the bank to impose significant charges upon access seekers for the provision of data which it has itself acquired at minimal cost will distort competition in the adjacent markets for which the spending data is an important input by raising the costs of rivals. The competition that is thereby lost may deprive users of significant benefits in terms of innovation, quality, or lower prices. In these circumstances, there would seem to be a strong case for requiring the sharing of data at little or no charge to the access seekers, since the potential costs of doing so are low and the potential benefits high. A similar argument appeared to apply in the case concerning terms of which KPN, the Dutch telecoms operator, was required to share name, address, and telephone number data, access to which was required for competitors to offer directory enquiry services. In that case, the Court of Justice decided that KN could only recover the costs associated with the transmission of the information to the provider of directories but not for the costs of acquiring such information, which could be considered a by-product of providing telephony services for which KPN was already remunerated²⁶².

In a similar vein, Prufer argues that all data, including aggregated data, should be shared by certain digital platforms without charge, on the basis that:

'that user information is a free by-product of running a service. Some have claimed that because obtaining access to more user information helps to improve the quality of one's service, accumulating it is justified as an end in itself. However, the data gathered in this way is certain to be transformed into revenue at some point (for instance, through advertising or some other sale of access to one's user groups), and protecting those indirect revenues is not a goal of the policy proposal at hand because, in the long run, they are subject to the main market-tipping dynamics characterized by Prüfer and Schottmüller (2017). User information, therefore, has the attributes of a public good. It is efficient to share it with every party that can (potentially) use it as input into its service and that benefits users in the end'.²⁶³

We would not characterise the issue in this way, since the accumulation of user data is the primary rationale for many of the investments which digital platforms like Google and Facebook and not a 'by product' of them²⁶⁴. We consider that investments by powerful digital platforms, even if they are made in markets which they dominate or if they contribute to the revenues which the platforms earn from those markets, generally involve some degree of risk and that changing the commercial incentives which these firms face (by imposing data sharing obligations upon them) will have an impact upon outcomes in related markets which a regulator would be irresponsible to ignore. This

²⁶² See p. XX of this report.

²⁶³ Prüfer (2020), p.14

²⁶⁴ Some forms of data, such as location data, might be a 'by product' of interacting with a particular platform, see footnote 199.

does not mean that the interests or incentives of the incumbent platforms should be given priority over others or that data should not be shared without charge in some circumstances, rather it means that regulators will need to engage in a more complex balancing act than Prufer seems to recognise.

Thus, in the case of Google, requiring Google to share the data it has acquired from user interactions with the Android eco-system with potential entrants into complementary markets at no charge could be expected to have significant consequences for Google's incentives and capacity to invest in Android (or other as yet unthought of services by which Google might otherwise expect to expand the value and scope of observed data which it can acquire from its users). This is because Google invests in Android for the primary purpose of acquiring (or retaining access to) user data that is generated when users interact with mobile devices and using it to generate revenues in its digital advertising business (which then fund the investments in Android, amongst other activities). Although the competitive advantages which Google derives from its user data might take time to diminish and Google may continue to invest heavily in the Android operating system (and associated applications) in the meantime, the commercial advantages which Google might expect to obtain from those investments will have reduced to the extent that it is now, and in the future, obliged to share them with potential competitors without being compensated for doing so. In the long run, Google could either be expected to evolve its business model to rely upon other sources of revenue to support its investments in Android²⁶⁵ or to reduce its level.

The consequences of competition in adjacent markets are also different. In this case, Google faces costs in acquiring the data for itself, but potential competitors who can obtain access to the same data without charge can avoid it. Competition may again be distorted, but in this case, because prices will not reflect the economic costs associated with the supply of the asset. Competition would also be distorted if Google were instead to be required to charge prices which were above the costs which Google itself faced, and so in this case the risk of distortion is symmetric.

The purpose of this brief exposition is to show that, **as with many regulated access sharing arrangements, different incentives and objectives will need to be balanced**. Regulators may, for example, conclude that the risks (in terms of investment and innovation incentives) from setting an access charge for data that is too low may be more than offset by the risks (in terms of competition or foreclosure) from setting an access charge that is too high. The magnitude of the risks arising from a failure to provide the correct pricing signals may also depend on the scope or volume of data that is to be shared. The important point is that it is **too simplistic to assume that all data should be shared at prices which would only allow the incumbent platform to recover the direct costs of implementing the data sharing arrangements, or which would only contribute to some proportion of the costs of storing the data, without also having regard to the investments that may have been made to acquire the data**.

5.2.4.2 Pricing of data relating to specific individuals

Our initial view is that there is likely to be a **stronger case for 'zero price' arrangements for the sharing of data that are related to specific individuals than for aggregated data**, which we consider further below. We note that the Second Payment Services Directive requires access to account services to be provided on terms which are 'objective, non-discriminatory and proportionate basis'²⁶⁶ and that charges are not levied between the parties for access to data or the implementation of payment instructions. As noted earlier, transfers of personal data that might be initiated by an

²⁶⁵ We have seen an early indication of this when, in early 2019, Google introduced a fee of \$10-40 for OEMs who wished to install the Google Mobile Services suite of apps onto android devices (although the underlying Operating System remains royalty-free). This followed the Commission's decision to require Google to supply the Android operating system and Mobile Services Suite without pre-installing the Chrome browser, see 'Google app suite costs as much as \$40 per phone under new EU Android deal', The Verge, 19 October 2018, available at <https://www.theverge.com/2018/10/19/17999366/google-eu-android-licensing-terms>

²⁶⁶ Article 36, Second Payment Services Directive, 2015/2366

individual user under the GDPR would also not attract charges, and so it may create distortions if similar remedies for competition purposes were instead to require payments to be made.

It might be argued that data relating to specific individuals can be expected to involve lower costs of acquisition or to be acquired for different purposes than aggregate data. However, the aggregate data we propose to be shared would likely be derived from the same provided and observed personal data as that which would be shared at the request of specific individuals. The **justification for having no charges** is, therefore, more likely to be based on the assumptions that (a) data relating to specific individuals could only be useful for firms that are seeking to engage in complementary innovation, rather than competing in the core market of the incumbent platform (where charges might be more appropriate) (b) any risk to investment or innovation incentives is likely to be minimal given that the high transactions costs of sharing data which requires individual user consent will impose a natural limit on the volume of data that will be shared (c) other non-economic considerations, such as human or other legal rights.

5.2.4.3 Pricing of aggregated data

As discussed earlier, existing European regulations do not provide much guidance as to how charges for the sharing of aggregated data might be determined. In the case of the Regulation concerning the provision of vehicle data to independent repairers, the relevant provision states that manufacturers should charge 'reasonable and proportionate fees', without indicating how these might be assessed in the event of a dispute²⁶⁷. Advisers to the European Commission have suggested that a **FRAND framework**, similar to that adopted by participants who contribute essential patents to the development of new industry standards, might provide a suitable precedent²⁶⁸. It is easy to see why - in both cases - a firm with exclusive rights over data needs to be appropriately remunerated for providing access to data whilst retaining incentives to invest and innovate. However, the application of FRAND in determining the level of charges that can be levied has proved very difficult to implement in practice. As one commentator noted concerning a leading European 'excessive pricing' FRAND case (involving Qualcomm):

'there is no "magic formula" that would allow a competition authority or a court to determine what "fair and reasonable" royalties are since this determination is context-specific. Moreover, while determining whether the fairness and reasonability of the price of a physical product are excessive is already difficult, that task is even more complex for non-physical constructs, such as intellectual property rights. Although several benchmarks were proposed to determine whether Qualcomm's royalties were "fair and reasonable", these benchmarks suffered from major weaknesses, either because they were theoretically unsound or because they would raise complex implementation issues.'²⁶⁹

The OECD has identified several benchmarks that regulators might employ if trying to set prices for individual user data, although all of them have shortcomings²⁷⁰. Similar considerations would likely apply to the valuation of aggregated data that was transferred in bulk. The benchmarks considered by the OECD include:

- **Market prices, such as the prices paid by organisations when acquiring data from data brokers or in bi-lateral transactions with each other.** At first sight, this is an attractive approach until it is realised that regulatory interventions to oblige firms to share data are likely to arise because the data in question is not otherwise already traded on

²⁶⁷ Regulation 2018/858 on the approval and market surveillance of motor vehicles Article 63(1)

²⁶⁸ "Also, there may be a need to oversee that data access is granted on fair, reasonable and nondiscriminatory (FRAND) terms – which need to be specified case by case.", Cremer et al (2019), p.109

²⁶⁹ Geradin (2013) p.7/8

²⁷⁰ OECD (2013).

commercial terms. There are therefore unlikely to be any market prices for the data in which regulators are interested. The OECD raises other issues, such as the fact that data, as a non-rivalrous good, may be traded many times. This would also be the case if a digital platform were subject to an obligation to provide access to data to a large number of access seekers (in UK Open Banking, over 300 entities have now been authorised). Any regulated prices would need to ensure that whatever investments the incumbent platform had made and which needed to be recovered were appropriately allocated amongst the population of access seekers and that there was no over- or under-recovery in aggregate. The costs for each recipient in relation to a given set of data assets might be expected to reduce for all with the addition of each new access seeker. However, the total number of access seekers from whom any costs are to be recovered may be difficult to predict in advance, which may make it difficult for both regulators and access seekers themselves to predict the costs they might incur (leading to obvious hold up issues if some potential competitors delay entry until it becomes clearer whether others will enter too).

- The **revenues of the digital platforms themselves (or revenue per data record), or their market capitalisations**. This approach also has many difficulties, including the risk that such revenues and valuations may include a significant proportion of supernormal profits which the digital platforms in question can earn by virtue of the very lack of contestability which the regulator is seeking to address. Besides, aggregate measures such as revenues or market value capture a large number of other factors, including non-data related activities and other sources of legitimate competitive advantage.
- **Experiments with users to determine the value, in price terms, they ascribe to different types of data**. This can involve, for example, asking users what they would have to be paid to disclose certain data about themselves or how much they might be prepared to spend to keep it private. These yields interesting results but have many of the challenges associated with consumer surveys of this kind and regulators are unlikely to be prepared to rely upon them without other evidence.
- Evidence from the **prices which organisations pay to insure themselves against the loss of data, or the costs they report as having incurred when such breaches happen**.

The OECD shows that very large variations in data values can be produced depending on the valuation methodology adopted and when the valuation is obtained. We do not consider it likely that any particular methodology will emerge to dominate in the foreseeable future. However, we do recommend that regulators consider ways in which commercial markets for data might be further encouraged, particularly concerning personal data of the kind that is retained by PDSs. The benefits of such developments may not come only from the widespread adoption of such services by users (and the potential for disrupting the source of market power which data access remedies are otherwise intended to address, something we discuss further below), but also by their role in establishing the market value of different types of data which regulators could then use as a basis for setting regulated prices.

The lack of methodologies for setting regulated prices for data reflects the current dearth of examples of such arrangements in the world. However, one potentially interesting area concerns how **financial exchanges, such as stock exchanges, sell market data to the participants** of those exchanges. Such data includes bid/offer prices, volumes of transactions, and other data which is generated by, and required for, the trading of securities. This data is 'observed data' in the sense that it is generated through the interactions between users of the trading platform. It is customarily sold, on commercial terms, by the exchanges to those who wish to trade.

In recent years, both the US Securities and Exchange Commission and the UK Financial Conduct Authority have expressed concerns about the rising costs of trading and, in the latter case, have now initiated an investigation into such practices. Several proposals have been made to regulate the

prices charged to traders, not all of which may be feasible²⁷¹. But, so far as we are aware, no regulations have been adopted and no prices have been set. If that happens, we would expect them to offer useful insights for any regulator engaged in the setting of prices for data to be transferred between digital platforms.

5.2.5 Promoting disruptive business models

The preceding discussion has focussed on the issues regulators and policymakers will need to address in implementing arrangements to ensure that access is provided to data that is held by large digital platforms to promote contestable markets. It proceeds on the assumption that such platforms derive significant competitive advantages from their accumulation and management of large volumes of valuable data. This appears to have been a reasonable presumption during a period in digital data that has come to be stored, organised and manipulated by large commercial organisations running the massive centralised computing systems that have historically been required to facilitate billions of transactions securely and efficiently. Measures to promote the sharing of data in such circumstances then tend to envisage regulators requiring or overseeing the transfer of data from one centralised source to another firm employing a similar business and technological model to promote competition between them.

Although not the primary focus of this report, it is important to note that advances in technology, particularly concerning cryptography, and changes in the purposes for which data is acquired or used, particularly with the growth in the Internet of Things, may mean that data comes to be acquired and controlled under different organisational arrangements in the future. As the New York Times noted:

'The true believers behind blockchain platforms like Ethereum argue that a network of distributed trust is one of those advances in software architecture that will prove, in the long run, to have historic significance. That promise has helped fuel the huge jump in cryptocurrency valuations. But in a way, the Bitcoin bubble may ultimately turn out to be a distraction from the true significance of the blockchain. The real promise of these new technologies, many of their evangelists believe, lies not in displacing our currencies but in replacing much of what we now think of as the internet, while at the same time returning the online world to a more decentralized and egalitarian system'²⁷²

The European Commission, in its 2020 data strategy communication, has made a related point:

'The volume of data produced in the world is growing rapidly, from 33 zettabytes in 2018 to an expected 175 zettabytes in 2025. Each new wave of data represents major opportunities for the EU to become a world leader in this area. Furthermore, how data is stored and processed will change dramatically over the coming 5 years. Today 80% of the processing and analysis of data takes place in data centres and centralised computing facilities, and 20% in smart connected objects, such as cars, home appliances or manufacturing robots, and computing facilities close to the user ('edge computing'). By 2025 these proportions are likely to be inverted'²⁷³

We highlighted earlier the potential role of new organisational models for data management, such as PDSs like *Solid*, a venture established by Sir Tim Berners-Lee which could significantly disrupt the business models of today's large digital platforms²⁷⁴. The *Decode* project is a European body that is

²⁷¹ Copenhagen Economics (2018). See also FCA (2020) 'FCA begins review on data in wholesale markets', available at <https://www.fca.org.uk/news/press-releases/fca-begins-review-data-wholesale-markets>

²⁷² S Johnson 'Beyond the Bitcoin Bubble', 16 January 2018, New York Times Magazine

²⁷³ Communication from the Commission of 19 February 2020, A European strategy for data, COM(2020) 66, p.2

²⁷⁴ <https://solid.mit.edu>

promoting the trials of such models with organisations drawn from across the European Union²⁷⁵. Although it is difficult - and well beyond the scope of this report - to assess how these and other initiatives might affect the functioning of digital markets in the future, policymakers will need to consider them and may wish to consider **measures which would be intended to accelerate or otherwise promote disruptive innovation of this kind, either in addition to or instead of the measures to promote the transfer of data between digital platforms** which are the focus of this report. For example, the Commission's latest European Data Strategy Communication referred to earlier, indicates that European funding may be allocated to promote such activities.²⁷⁶

5.2.6 *Exiting from data sharing arrangements*

This report has thus far focussed on the issues and obstacles to be addressed by a regulator seeking to implement data sharing arrangements to ensure the contestability of digital markets. However, we end this section by highlighting challenges that can arise after a regulator has successfully introduced access obligations into the market. There is no reason to expect that such arrangements, once established, should remain in place in perpetuity. There are also risks that innovation and investment will be inhibited or deferred if firms assume a position of 'regulatory dependency' from which there is no prospect of escape. As noted earlier, there is a tendency in these circumstances for the field of competition to shift from competing for the attention of users to lobbying the regulator.


Equally, however, it is not clear that they should be subject to an arbitrary 'sunset' date at which point any obligations to share data would cease. Such arrangements might continue voluntarily at that point, but if they did not then their withdrawal would be likely to inhibit competition and disrupt the provision of services to users in a way which we think would be difficult to justify. The regulator cannot predict when the measures are being implemented how markets will develop in the future, particularly in fast-moving and unpredictable digital markets, including new markets that may not even be discernible at the time. A more flexible approach will therefore be required.

One approach is for **the regulator to undertake a periodic re-assessment of the data sharing arrangements and their justification**. This could involve, as occurs under the European Electronic Communication Code every 5 years, a periodic review of the 'gatekeeper platform' designations from which obligations to share different types of data will have been derived and which we discussed in Section 5.1.1. It may be that the measures that have been taken to promote entry into niche markets have been so successful that the regulated platform no longer enjoys a 'gatekeeper' position in its core market. More likely, the entrants into complementary markets may have acquired sufficient scale (in terms of users from whom they can obtain volunteered and observed data independently of any sharing arrangements) to mean that they no longer require access to data held by the regulated platform. On the other hand, since the main purpose of the data sharing measures is likely to be to prevent foreclosure in new but as yet unknown niche markets into which new, but as yet unknown competitors might seek to enter, and since the possibility of such markets arising is unlikely to have obvious limits, it may be that the data sharing measures we envisage in this report would persist for much longer than the needs of any individual access seeker.

This raises the question of **whether individual access seekers should be incentivised to reduce their dependency on the data sharing arrangements over time**, even if they remain generally available to newer entrants. A version of this approach was adopted in the telecommunications sector and known as the 'ladder of investment'. In this view, regulators should, over time, inflate the charges payable by those who relied on regulated access to networks, thereby increasing their costs relative to the option of investing in their facilities. The analogy for our purposes might involve the regulator gradually raising the cost of access to shared data (potentially, in the case of data relating

²⁷⁵ <https://www.decodeproject.eu>

²⁷⁶ Communication from the Commission of 19 February 2020, A European strategy for data, COM(2020) 66, p.16



to specific individuals, from a price of zero) for those platforms that had already been relying upon shared data for some time. As the cost of shared data increased, it might be expected to incentivise access seekers to acquire data directly from users instead now that they had the scale to do so. Alternatively (or also), the regulator might reduce over time the scope (or quantity) of the data to which access is provided.

There are, however, several significant objections to this concerning data sharing. First, it is not clear that such incentives are required with the data sharing measures we propose. New platforms will need to acquire as many users as possible if they are to succeed and, by doing so, will acquire data. Conversely, if the competing platform already has powerful incentives to acquire its users and hence data (given network effects and other features of digital markets²⁷⁷), it is unlikely to be able to avoid the additional charges that are applied by the regulator by acquiring more users (and hence data) than is already the case. There is a danger in these circumstances that raising the costs of data access will simply represent a transfer of rents from one platform to another without having any impact on their conduct, or that it might prove counterproductive. Second, it will be clear that applying such a policy would be very difficult in practice. We have already noted that the issue of determining appropriate charges for access to data is likely to be very challenging, with a high possibility of error (in either direction). Adjusting these charges to incentivise some forms of competition without excluding others would be no easier. Regulators in the telecommunications sector sometimes proved unable, in the face of lobbying from firms who were dependent on access to regulated services, to inflate charges in the way that had been originally anticipated and the application of the policy produced mixed results.²⁷⁸

This does not mean that regulators should simply ignore the risk that data sharing measures might, over time, distort competition for users and their data. Indeed, this may be even more important if, as we anticipate, data sharing measures could, if implemented, be expected to persist for far into the future.

²⁷⁷ See, e.g. Martens (2020), p.10-

²⁷⁸ Feasey R and Cave M (2017), p.16-17

06

The background features a dark blue field with various geometric shapes in lighter shades of blue. In the top left, there is a white triangular shape. On the left side, there are several overlapping triangles pointing right. The bottom half of the page is filled with a pattern of overlapping triangles of various sizes and orientations, some pointing right and some pointing left.

**POLICY
RECOMMENDATIONS**

6 Policy recommendations

The sharing of data is a collaborative process, involving an organisation which currently controls data, and the intended recipient and, in many cases, a person to whom the data relates and whose consent may be required before a transfer can be made. Aligning the interests of these different parties and overcoming other barriers to sharing can prove to be a complex and difficult task, as evidenced by the low levels of voluntary data sharing that we often observe today.

Some form of regulatory intervention is likely to be required when control over data is a source of market power. In this study, we identify several issues that the regulator will need to address if data sharing is to be an effective remedy to concerns about market contestability and the capacity of those controlling data to leverage the advantages it confers into new or adjacent markets. Some of these issues have already arisen and been addressed in existing regulated data sharing arrangements, such as Open Banking, or voluntary arrangements such as the Data Transfer Project. Others have yet to be fully considered.

Regulating recipients as well as donors

The first conclusion is that **regulation for data sharing should not be viewed as being limited to the oversight of a small number of large platforms that might be obliged to share data. Instead, it will require strict oversight of potentially a very large number of smaller firms seeking access to such data.** This is particularly important because **we conclude that it will not be practical or desirable for regulators to seek to restrict the use to which the data is subsequently to be put** (despite access to the data being required to preserve the contestability of adjacent markets into which the incumbent platform might otherwise leverage its data advantages). Given the potentially wide range of applications to which data could be used, and the wide range of organisations which may require access to such data, individual users will not consent to the sharing of data unless they can be confident that any recipient of the data will keep it secure, adhere to other conditions of sharing and so preserve trust in, and the integrity of the overall data sharing process. The controllers of commercial data will also not comply with sharing obligations if misuse by others puts their reputation or commercial position at risk, whilst potential recipients of data may be putting themselves in a position of acute dependency (since they may rely upon uninterrupted data sharing to sustain their services for users) and will not do so unless they consider that they have adequate protections and rights of redress.

It follows that if regulated data sharing is to be adopted at a significant scale, regulators will need to establish an effective regime for overseeing those in receipt of data and for enforcing the rules effectively on an ongoing basis. This will need to include rules governing the resolution of disputes and determining how liabilities fall if consumers or other firms are harmed. Since those who receive data are unlikely to hold market power or otherwise to be guilty of any abuse, we consider that oversight of such arrangements is unlikely to be an appropriate task for a competition authority and will instead require a dedicated regulatory body.

Extensive obligations to adopt common technical standards

Secondly, all forms of data sharing will require the **adoption of common technical standards by both those sharing data and those in receipt of it.** The same standards should be adopted for all the different forms of data sharing that we propose. We consider that potential recipients of data have sufficient incentives to adopt the standards since they would not otherwise obtain access to the data they require. Those platforms that have been directed to share data will need to be obliged to adopt the relevant standards, such that data can be shared in a form and manner which supports the regulatory objectives. In the early stages of regulation, this may impose additional costs on the newly regulated entities as they have to restructure the way they manage their existing data assets or adopt new external interfaces. This may also contribute to delay in the implementation of new data sharing obligations, which will be a particular concern if the objective of data sharing is to prevent leveraging into emerging digital markets. In the longer term, we conclude that data sharing

regulation should promote the very extensive adoption of common technical standards, both by organisations which may not currently have obligations to share data (but which might be required to in the future), those who may not currently request access to data (but will want to preserve the option to do in the future), and concerning forms of data which may not currently be shared (but which may be required to be shared in future). **This 'anticipatory' approach to technical standards means that regulators should consider the application of common technical standards to data sharing in sectors well beyond the existing scope of large digital platforms,** as has been proposed in Australia. In short, **we recommend regulators should decouple requirements to adopt common technical standards from obligations to share data in the expectation that the former will be much more extensive than the latter.**

Anticipating and policing of anti-competitive conduct

Thirdly, **regulators should anticipate that any data sharing process is likely to vulnerable to anti-competitive conduct which is intended to inhibit the effective sharing or use of data. Several aspects should receive particular attention.** First, those obliged to share data may seek to use their influence over the development of common technical standards to limit the scope, delay the implementation, or otherwise raise the costs of data sharing for other parties. These firms have a legitimate interest in shaping their development - but so too will potential recipients of data. We recognise that several existing data standardisation initiatives could provide useful contributions and regulators will want to ensure that industry participants collaborate in further work rather than seeking to impose their requirements from the outset. However, regulators will need to be prepared to intervene in the development process or if there are disputes that need to be resolved. Second, some types of anti-competitive conduct may be presented as being necessary to protect the interests of users. For example, it might be argued that 'pop ups' that warn a user that they are about to consent to the transfer of their data to a third party, and which ask them whether they wish to do so, are a legitimate safeguard against unintended or unauthorised transfers. Our first recommendation - that recipients of data be strictly regulated - is partly intended to remove the justification for such practices. Finally, digital platforms may seek to impose restrictive commercial terms in their data sharing agreements, intended to limit the capacity of potential recipients to compete with the platform. Experience of applying access arrangements in other contexts suggests that different forms of anti-competitive conduct can emerge over time (as some practices are stopped but others replace them) and will require constant scrutiny.

The most important and difficult role for regulators will lie in determining the type and scope of data that is to be shared and which organisations should be obliged to share it. **We conclude that two forms of regulated sharing are likely to dominate.**

Recommendations on sharing of data about individual users

The first form of sharing - and **the one which is likely to be capable of being implemented first** - will be the **sharing or porting of data about individual users.** This mode of sharing is likely to be appropriate when the individual concerned will benefit directly from the sharing process, likely through the provision by the recipients of complimentary services in adjacent markets. The value of the data, in this case, lies in its depth and personalised nature, rather than in its volume. The process to enable the sharing of the data will generally require that the user consent to the transfer and the process by which these user consents are obtained and authenticated will have a significant impact on the effectiveness of this remedy. Technologies such as biometric ids will have a significant role to play.

The data to be transferred would be data provided by the user to the platform and data derived from observations of that individual's interactions with the platform. It would exclude 'inferred data' that is created by the platform itself (as well as excluding third party data that is purchased from other sources). The presumption should be that all relevant data about an individual would be shared.

The overall competitive impact of these data sharing arrangements will necessarily be limited, given the relatively high transaction costs associated with first obtaining individual consents from every user and the relatively small volumes of data that will be transferred each time consent is obtained. Over time, however, data that is obtained in this way could accumulate and be used for other purposes. For this reason, **we recommend that obligations to share data about individual users in the way we propose should be quite extensive and apply to digital platforms which we would describe as meeting the 'gatekeeper minus' threshold.** This would mean a strong presumption that the obligation to share would apply to all platforms which the regulator had determined as having 'gatekeeper' or equivalent status and to some others as well. However, **this obligation would not apply to every platform or firm**, and so would be less extensive than, for example, **the 'data portability' obligations which apply under the GDPR (which are narrower in scope).** We do not recommend that the European Commission seek to expand the existing GDPR data portability requirements to address the competition concerns we consider in this report and conclude that a separate regime, specifically designed for this purpose, is the better approach.

We consider that there is a case for a regulator to require the sharing of individual user data without any form of payment passing between the donor and the recipient. Each party would be expected to bear its costs concerning the transfer.

Consideration of 'opt out' arrangements

It is unclear at this stage how effective the arrangements for the sharing of individual data outlined above would prove to be. However, there is a risk that the high transaction costs and uncertain benefits continue to deter users and would render this approach relatively ineffective in preserving the contestability of the markets we are concerned with. **In such circumstances, we recommend the European Commission should consider more radical approaches, including changes to the GDPR which would allow for individual users to 'opt out' their personal data** (rather than requiring them to 'opt in') when transfers of their data are initiated - provided always that the recipients of the data comply with the relevant regulatory conditions.

We recognise that this may represent some loss of consumer sovereignty over their data, but consider that such a trade-off may need to be made if data sharing arrangements are to achieve their aim of ensuring contestability in digital markets. It is far from clear that the interests of European consumers are better served by preserving rights to consent whilst allowing new digital markets to be dominated by existing 'gatekeeper' platforms. Indeed, in the long run, the privacy rights of European consumers may be better served by measures that more effectively promote competition. **We, therefore, recommend the European Commission consider provisions in the forthcoming Data Act to enable the use of 'opt out' arrangements for the sharing of personal data to preserve market contestability under certain prescribed conditions.** There is certainly no precedent for such arrangements since control of personal data sets has often changed without individual user 'opt ins' when one firm acquires another firm or when one firm acquires another's data assets.

Recommendations on the bulk sharing of user data

The second form of sharing will be the bulk transfer of aggregate user data. As with the first category, this would involve sharing data provided by individual users or arising from their interactions with the platform but would exclude inferences that are generated by the platform itself. This mode of sharing is likely to support entry into adjacent or emerging markets, with such entry being supported by insights derived from large data sets, or even to support competition in some or all of the core market activities from which or by means of which the data has been derived.

The overall competitive impact of these data sharing arrangements could be significant – likely more significant than for individual user data - since the volume of data to be shared is likely to be very substantial and may represent a significant proportion of the donor platform's data assets. In some circumstances, it may be necessary for the data to be shared without first anonymising it to allow recipients to effectively rival the incumbent platform. **Since obtaining individual consent from**

every user would not be feasible in these circumstances, we recommend that regulators and policymakers consider other mechanisms to enable the bulk sharing of non-anonymised user data.

Alternatively, regulators should consider requiring the platform that controls the data to allow third party access to the full data set for training algorithms or otherwise deriving the same sorts of insights from the data that are available to the incumbent. The terms under which such access is provided would also need to be carefully regulated since those seeking access to the data sets would remain dependent upon the owner of the assets providing full and unrestricted access. Similar challenges arise even when the data is held by a 'neutral' intermediary. Such arrangements are therefore likely to require a high degree of regulatory oversight (and associated cost), although they also have considerable advantages if non-anonymised data is important to preserve contestability or if very large data sets are involved.

Although we would expect all the relevant data about an individual user to be shared with every recipient, **there is likely to be much greater heterogeneity of demand amongst potential recipients of bulk transfers of aggregate data.** Some potential recipients may require (or may only be able to handle) relatively small volumes of data, representing only a fraction of that held by the donor. Others may require the sharing of much larger data sets. There may also be questions about the geographic scope of the data to be shared. This will present two challenges. First, the regulator will need to ensure that a **suitable menu of data options** is developed, preferably collaboratively and inclusively, to ensure that the needs of as wide a range of potential recipients as possible will be met as far as possible. This is likely to involve a degree of compromise on the part of some parties, with the regulator adjudicating between conflicting demands.

Second, **we consider there is a strong prima facie case for assuming that recipients of aggregated data should be required to pay the data, with the payment varying by the volume and value of the data being shared (and not simply the costs of implementing the data sharing arrangements or storing the data).** The primary concern here is to preserve incentives for both parties in the sharing arrangement to innovate and invest in existing or new digital services to acquire additional data for themselves. We do not want data sharing arrangements to crowd out other forms of commercial activity from which users derive significant benefits, particularly in many digital markets.

However, we do not make firm recommendations as to how these prices should be derived because we have yet to find a well-developed methodology for doing so. Requiring firms to agree with terms on a 'FRAND' basis may be inadequate. **We recommend that a study be undertaken by the Commission to consider how regulators would establish wholesale prices for data that was to be shared.** The methodologies and the practices to calculate the marginal costs and to recover costs developed for public data may feed this study. We also consider that setting appropriate wholesale prices for the receipt of aggregate data will also be necessary to ensure that recipients have appropriate incentives to reassess their data requirements as they grow and develop their businesses, allowing for the possibility that they would terminate existing data sharing arrangements once they have acquired, or are in a position to acquire, sufficient data for themselves from their users. Otherwise, extensive data sharing arrangements could likely become a permanent feature of European digital markets in the years to come.

The final question in this context concerns the identity of the platforms that would be obliged to share aggregated personal data on a bulk basis. **We conclude that this should be a much-limited set of entities than we recommend for the porting of individual data** and would not necessarily be a requirement of every platform that was found to hold 'gatekeeper' status under the European Commission's latest proposals, although we think a designation of 'gatekeeper' status should establish a rebuttable presumption. **We, therefore, characterise this sub-set of entities as being those that meet a (more demanding) 'gatekeeper plus' threshold.** The analysis required to demonstrate this would need to be undertaken on a case by case basis.



The challenge ahead

The recommendations in this report, if adopted, would represent an extensive programme of regulatory activity that would need to be undertaken by bodies with responsibilities for implementing data sharing which have yet to be assigned in Europe. Establishing the institutional and regulatory framework to deliver data sharing at scale will require legislation and, besides, we recommend that the European policymakers consider further legislative changes in the forthcoming Data Act to enable the sharing of personal data on an opt out basis under certain narrowly prescribed circumstances and to ensure contestability in digital markets.

Finally, we are mindful that data sharing remedies that we have considered in this report arise from the assumption that digital platforms will continue to derive significant market power from their centralised control of big data sets which they have accumulated by enabling diffuse groups of users to transact with each other through the platform. This may be the case, but regulators and policymakers should also keep an eye on (and potentially take steps to promote) new technologies and architectures which might in the future enable a much greater degree of decentralisation and wider distribution of data, thereby removing the very sources of market power which this report has sought to address.

The background is a solid dark blue. It features several overlapping triangles in various shades of blue, ranging from light to dark. Some triangles are pointing right, some left, and some are partially cut off by the edges of the page. The word "REFERENCES" is centered in the upper half of the page in a bold, white, sans-serif font.

REFERENCES

References

- Abrahamson Z. (2014), 'Essential Data', *Yale Law Journal* 124, 867-881.
- Alexiadis P. and A. de Streel (2020), *Designing an EU Intervention Standard for Digital Platforms*, EUI Working Paper-RSCAS 2020/14.
- Argenton C. and J. Prüfer (2012), "Search engine competition with network externalities", *Journal of Competition Law & Economics*, 2012, 8(1), 73-105.
- Autorité de la concurrence and Bundeskartellamt (2016), *Competition Law and Data*.
- Balto D.A. and M.C. Lane (2016), "Monopolizing water in a tsunami: Finding sensible antitrust rules for Big Data", *Competition Policy International*.
- Baumol W.J. , J. Panzar and R. Willing (1982), *Contestable Markets and the Theory of Industry Structure*, Saunders College Publishing/Harcourt Brace.
- Benson (2009) 'Data portability: the antidote to data lock in', ECCMA White Paper.
- Borgogno O. and G. Colangelo (2019), 'Data Sharing and Interoperability Through APIs: Insights from European Regulatory Strategy', *Computer Law & Security Review* 35, 1-17.
- Borgogno O. and G. Colangelo (2020a), 'Data, Innovation and Transatlantic Competition in Finance: The Case of the Access to Account Rule', *European Business Law Review*.
- Borgogno O. and G. Colangelo (2020b), 'Consumer inertia and competition-sensitive data governance: the case of Open Banking', available at SSRN.
- Bourreau M. and A. de Streel (2019), *Digital Conglomerates and EU Competition Policy*, Mimeo.
- Bresnahan T.F. and M. Trajtenberg (1995), 'General purpose technologies: Engines of growth?', *Jour. of Econometrics* 65(1), 83-108.
- Crémer, J., de Montjoye, Y.-A. and H. Schweitzer (2019). *Competition policy for the digital era*, Report to the European Commission.
- Colangelo G. and M Maggolino (2018) Big data as misleading facilities, *European Competition Journal*, 249-281.
- Copenhagen Economics (2018), 'Pricing of market data'.
- Costa-Cabral F. and Lynskey O. (2017), 'Family ties: the intersection between data protection and competition in EU Law', *Common Market Law Review* 54(1), 11-50.
- Ctrl-Shift (2018), *Data Mobility: The personal data portability growth opportunity for the UK economy*, Report for the UK Department for Digital, Culture, Media & Sport.
- Ctrl-Shift (2019) 'Data Mobility Infrastructure Sandbox'.
- Deloitte, et al. (2018), *Emerging issues of data ownership, interoperability, (re-)usability and access to data and liability*, Study for the European Commission.
- de Streel A. and T. Tombal (2020), 'The Fifty Shades of Data Sharing and EU Law', available at SSRN.

Deloitte, Open Evidence, WIK and Timelex (2017), *Emerging issues of data ownership, interoperability, (re-)usability and access to data, and liability*, Study for the European Commission.

Department for Business, Energy & Industrial Strategy (2018), 'Implementing Midata in the domestic energy sector: Government response to call for evidence'.

Drexl J. (2017) "Designing Competitive Markets for Industrial Data - Between Propertisation and Access", *JIPITEC*, p. 257-292.

Drexl J. (2018), *Data access and control in the era of connected devices*, Report for BEUC.

Egan (2019), '*Charting a way forward: data portability and privacy*', Facebook.

Everis (2018), *Data sharing between companies in Europe*, Study for the European Commission.

Expert Group for the Observatory on the Online Platform Economy (2020) *Work stream on Data*.

Ezrachi A. and M.E. Stucke (2016), *Virtual Competition: The Promise and Perils of the Algorithm-Driven Economy*, Harvard University Press.

Facebook (2020), *Comments to the Federal Trade Commission on Data Portability*.

Feasey R and Cave M. (2017), *Policy towards competition in high speed broadband in Europe, in an age of vertical and horizontal integration and oligopolies*, CERRE Report.

Federal Trade Commission (2014), *Data brokers: a call for transparency and accountability*.

Federal Trade Commission (2015), *Mergers and privacy promises*.

Financial Conduct Authority (2017), *Regulatory sandbox lessons learned report*.

Fingelton/ODI (2019), *Open Banking: preparing for lift off*.

Fraunhofer (2017), 'German Government and Fraunhofer drive forward plans to implement Industrie 4.0 on an international scale', *press release*, 27 July 2017.

Furman, J., D. Coyle, A. Fletcher, D. McAuley and P. Marsden (2019). *Unlocking digital competition*. Report of the Digital Competition Expert Panel.

Gal M.S. and D.L. Rubinfeld (2019), 'Data Standardization', *NYU Law Review* 94.

Gal and Aviv (2020), 'The Competitive effects of the GDPR', forthcoming, *Journal of Competition Law and Economics*.

Geradin D. (2013), "Ten Years of DG Competition Effort to Provide Guidance on the Application of Competition Rules to the Licensing of Standard-Essential Patents: Where Do We Stand?", *Northwestern University School of Law*.

Google (2020), '*Online Platforms and Digital Advertising: Comments on the Market Study Interim Report*'.

Graef I. (2016), *EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility*, Kluwer Law International.

Graef, I., Husovec, M. and Purtova, N. (2019), 'Data Portability and Data Control: Lessons for an Emerging Concept in EU Law', *German Law Journal*, 1359-1398.

Graef I., M. Husovec and J. van den Boom (2019), 'Spill-overs in data governance: Uncovering the uneasy relationship between the GDPR's right to data portability and EU sector-specific data access regimes', *Journal of European Consumer and Market Law*.

Graef I., R. Gellert, and M. Husovec (2019), 'Towards a Holistic Regulatory Approach for the European Data Economy: Why the Illusive Notion of Non-Personal Data is Counterproductive to Data Innovation', 44(5), *European Law Review*, 605-621.

Graef, I., T. Tombal and A. de Streel (2019) "Limits and Enablers of Data Sharing. An Analytical Framework for EU Competition, Data Protection and Consumer Law", *TILEC Discussion Paper 2019-024*.

House of Lords (2016), *Online Platforms and the Digital Single Market*, 10th Report of Session 2015–16, HL Paper 129.

IDC and Lisbon Council (2020), *European Data Monitor*, Study for the European Commission, available at: <http://datalandscape.eu/>

Ilan, Rosen, Ronco and Gerlach (2016), 'Privacy in M&A Transactions: personal data transfer and post-closing liabilities', *Harvard Law School Forum on Corporate Governance*.

Judge Business School (2015), 'Personal data stores', Study for the European Commission.

Kerber, W. (2016) "Governance of Data: Exclusive Property vs. Access", *IIC*, Volume 47, p. 759-762.

Kerber W. (2018), "Data Governance in Connected Cars: The Problem of Access to In-Vehicle Data", *JIPITEC* 9, 310-331.

Kerber W. (2020), 'From (Horizontal and Sectoral) Data Access Solutions towards Data Governance Systems', *MAGKS Joint Discussion Paper* 40.

Kerber W. and H. Schweitzer (2017), 'Interoperability in the Digital Economy', *JIPITEC*, 39.

Krämer J., D. Schnurr and S. Broughton Micova (2020), *The Role of Data for Digital Markets Contestability: Case Studies and Data Access Remedies*, CERRE Report.

Krämer J., P. Senellart and A. de Streel (2020), *Making Data Portability More Effective for the Digital Economy*, CERRE Report.

Lambrecht A. and C. Tucker (2015), "Can Big Data Protect a Firm from Competition?", available on SSRN.

Lerner A. (2014), "The Role of 'Big Data' in Online Platform Competition", available on SSRN.

Lundqvist, B (2018) Competition and Data Pools, *Journal of European Consumer and Market Law*, 146-154.

Lynskey, O (2017), 'Aligning Data Protection Rights with Competition Law Remedies? The GDPR Right to Data Portability', *European Law Review* 42(6), 793-814.

Maier N. (2019), 'Closeness of substitutions for big data in merger control', *Journal of Competition Law and Practice*.

Martens B., A. de Streel, I. Graef and T. Tombal, *Business-to-Business data sharing: An economic and legal analysis*, JRC Digital Economy Working Paper 2020-05, July 2020.

Mayer-Schönberger V. and T. Ramge (2018), *Reinventing Capitalism in the Age of Big Data*, John Murray.

Meadows M. (2015), 'The Essential Facilities Doctrine in Information Economies: Illustrating Why the Antitrust Duty to Deal is Still Necessary in the New Economy', *Fordham Intellectual Property, Media and Entertainment Law Journal* 25(3), 795-830.

Newman N. (2014), "Search, Antitrust and the Economics of the Control of User Data", *Yale Journal of Regulation* 30(3).

OECD (2013), *Exploring the economics of personal data: a survey of methodologies for measuring monetary value*, DSTI/ICCP/REG(2011)2.

OECD (2015), *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing.

OECD (2019), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing.

OECD (2020), 'Working Party No. 2 on Competition and Regulation Lines of Business Restrictions – Background note', OECD Publishing.

Osborne Clarke (2016), *Legal study on ownership and access to data*, Study for the European Commission.

Prüfer J. and C. Schottmuller (2017), *Competing with Big Data*, CentER Discussion Paper 2017-007.

Prüfer (2020), 'Competition Policy and Data Sharing in Data Driven Markets, Friedrich Ebert Stiftung.

Ram and Murgia, 'Data brokers: regulators try to rein in the 'privacy deathstars', *Financial Times*, 8 January 2019.

Rubinfeld D.L. and M.S. Gal (2017), "Access Barriers to Big Data", *Arizona Law Review* 59, 339-381.

Scassa T. (2018), *Data Ownership*, Centre for International Governance Innovation Papers, No. 187.

Schweitzer H., Haucap J., Kerber W. and Welker R. (2018), *Modernising the law on abuse of market power*, Report for the German Federal Ministry for Economic Affairs and Energy.

Scott Morton, F., Bouvier, P., Ezrachi, A., Jullien, A., Katz, R., Kimmelman, G., Melamed, D. and J. Morgenstern (2019). *Committee for the Study of Digital Platforms, Market Structure and Antitrust Subcommittee*, Stigler Center for the Study of the Economy and the State.

Shapiro C. and Varian H. (1999), *Information Rules – A Strategic Guide in the Information Society*, Harvard Business School Press.

Sokol D. and R. Comerford (2016), "Antitrust and Regulating Big Data", *Georges Mason Law Review*, 1129-1161.

Support Centre for data sharing (2019), *Report on collected model contract terms*, available at: <https://eudatasharing.eu/legal-aspects/report-collected-model-contract-terms>.

Support Centre for Data Sharing (2020), *Analytical report on EU law applicable, to sharing of non-personal data*.

Stucke M. and J. Grunes (2016), *Big Data and Competition Policy*, Oxford University Press.

Taddy M. (2019), 'The Technological Elements of Artificial Intelligence', in A.K. Agrawal, J. Gans and A. Goldfarb (ed), *The Economics of Artificial Intelligence*, University of Chicago Press.

Tombal T. (2020), Economic dependence and data access, *IIC* 51(1), 70-98.

Tractebel (2019), '*European smart metering benchmark*'.

Tucker D. and H. Wellford (2014), "Big Mistakes Regarding Big Data", *Antitrust Source*.

UK Competition and Markets Authority (2015), *The commercial use of consumer data*, Report CMA 38.

UK Competition and Markets Authority (2016), *Energy Market Investigation final report*, CMA.

UK Competition and Markets Authority (2019), *Online platforms and digital advertising*, Market study interim report.

UK Information Commissioner (2019), 'Statement on intent to fine Marriott International Inc more than £99 million for GDPR data breach', press release, 9 July 2019.

US Senate Committee on Commerce, Science and Transportation (2013), *A Review of the Data Broker Industry: Collection, Use and Sale of Consumer Data for Marketing Purposes*, Staff Report.

Varian H. (2019), 'Artificial Intelligence, Economics, and Industrial Organization', in A.K. Agrawal, J. Gans and A. Goldfarb (ed), *The Economics of Artificial Intelligence*, University of Chicago Press.

Vezzoso S. (2018), "Fintech, Access to Data, and the Role of Competition Policy", available at SSRN.

VVA (2017), *Data in platform-to-business relations*, Study for the European Commission.



cerre

Centre on Regulation in Europe

 Avenue Louise, 475 (box 10)
1050 Brussels, Belgium

 +32 2 230 83 60

 info@cerre.eu

 cerre.eu

 [@CERRE_ThinkTank](https://twitter.com/CERRE_ThinkTank)