

Chapter 5

ENHANCING DATA ACCESS, SHARING AND RE-USE

KEY FINDINGS

- Overall use of data has increased over time, but varies significantly across sectors, countries and – most significantly – by firm size.
- More than 25% of all information and communication technology (ICT) firms in the European Union used big data in 2018 compared to 10% of all firms.
- Besides the ICT sector, utilities, transportation and logistics are also highly intensive users of big data, with around 20% of firms in these sectors using big data in 2018.
- Adoption has also increased in other sectors in a number of countries. In Germany, 12% of all manufacturing firms used big data in 2018 compared to 9% in the European Union. This uneven picture of diffusion has important implications for productivity performance.
- While access and sharing can help increase the value of data to data holders, they can create 10 to 20 times more value to data users, and 20 to 50 times more value for the wider economy. In some cases, however, data access and sharing can also reduce the potential income of data holders. This underscores the incentive challenge for governments.
- Among the 205 policy initiatives on data access and sharing (across 37 countries) that were analysed, 126 (61%) aimed at enhancing access to public sector data, 44 (21%) aimed at help share private-sector data across the economy and 24 (12%) improved data analytic capacities. Innovative mechanisms for the controlled sharing of sensitive data have aided in the response to COVID-19 and merit further analysis.

Introduction

Data, and data access and sharing, have become fundamental for social and economic activities. In the context of the COVID-19 pandemic, leveraging data has been centre-stage in establishing effective frontline responses to the crisis. It will also be an essential part of the recovery and resilience-building phase.

This chapter underlines how fundamental data, and data access and sharing, have become for social and economic activities. It presents trends in data use across the economy, as well as recent empirical studies of their effects on productivity. It focuses on industries unrelated to information and communication technologies (ICTs).

While overall use of data increased between 2016-18, it still varies significantly across sectors, countries and – most significantly – by firm size. The ICT sector, and in particular large firms, remained by far the dominant users of data in 2018. More than 25% of all ICT firms in the European Union, for instance, used big data in 2018, compared to 10% of all firms. Besides the ICT sector, utilities, transportation and logistics are also highly intensive users of big data, engaging around 20% of these firms in 2018.

Adoption has also increased in other sectors in a number of countries. In Germany, for example, 12% of all manufacturing firms used big data in 2018 compared to 9% in the European Union. This uneven picture of diffusion has important implications for productivity performance.

The chapter also focuses on how data access and sharing can facilitate the use of data across societies, including across borders. It highlights promising venues to overcome the challenges to such outcomes such as risks to privacy, intellectual property rights and of losing control over data.

Studies show data access and sharing can increase the value of data for the wider economy. While they can help increase the value of data to data holders, they can create 10 to 20 times more value to data users, and 20 to 50 times more value for the wider economy. In some cases, however, data access and sharing can also reduce the potential income of data holders, which underscores the incentive challenge facing governments.

More differentiated and balanced data governance approaches are needed. These should better leverage technological solutions for privacy protection and enhanced control over data and information, such as cryptography and data sandboxes. This, in turn, would protect data confidentiality, give stakeholders more control over their data and incentivise data access and sharing. Further research into the concept of “data ownership” and its relationship to different types of data will also be critical to formulating effective policy.

The chapter concludes with an overview of government initiatives to facilitate data access and sharing, including across borders. Many of these initiatives also aim to address the challenges associated with protection of privacy, intellectual property rights and data control. All surveyed countries had initiatives that foster and enhance access to and sharing of public sector data in 2018. However, significantly fewer countries targeted private-sector data. Even fewer governments had initiatives to improve the capacity to analyse data in their countries. Of 205 policy initiatives across 37 countries, 61% aimed at enhancing access to public sector data, while 21% aimed to help share private-sector data.

Governments have recognised the availability of data-related skills and competences can be a critical bottleneck for the effective re-use as well as provision of data in both the private and public sectors. However, only 12% focused on improving data analytic capabilities across society. Innovative mechanisms for the controlled sharing of sensitive data have aided the response to COVID-19 and merit further analysis.

Trends in the use of data and data analytics

Data and (big) data analytics can boost productivity and innovation

The effective use of data can help boost productivity and improve or foster new products, processes, organisational methods and markets. There is still little reliable quantification of the economic effects of data use. However, firms that use data exhibit faster labour productivity growth than those that do not by approximately 5% to 10% (OECD, 2015^[1]). In addition, findings from McKinsey & Company (2017^[2]) suggest that data monetisation is an increasingly important driver of revenue growth. The monetisation of data reportedly contributes to 10% or more of the total revenue for 32% of high-performing businesses and 9% of all other businesses.¹

In manufacturing, data are typically obtained through sensors that are increasingly used to monitor and analyse the efficiency of machines, optimise their operations and provide after sale services, including preventive maintenance. The data are sometimes also used to work with suppliers. In some cases, they are even commercialised through new services such as optimising production control (OECD, 2017^[3]). Increasingly, manufacturing activities rely on data flows that connect geographically dispersed stages of production across global value chains (see section below). This has significant impact on the productivity and innovation capacity of manufacturing firms.

In the United States, for instance, Brynjolfsson and McElheran (2019^[4]) estimate that being at the frontier of data-driven decision in manufacturing is linked with improvements in revenue-based productivity of 4% to 8%. The authors show that timing, however, is essential. Leading adopters of data analytics are receiving the biggest gains, while laggards that reach the frontier later tend to have lower net benefits or none at all.

Based on German firm-level data, Niebel, Rasel and Viete (2018^[5]) find evidence that use of data and analytics increases the likelihood of a firm becoming a product innovator, as well as for the market success of product innovations. These results hold for both manufacturing and service sectors, but are contingent on firms’ investment in IT-specific skills (Niebel, Rasel and Viete, 2018^[5]). Others have documented similar findings (Bajari et al., 2019^[6]; Wamba et al., 2017^[7]; Brynjolfsson and McElheran, 2016^[8]; Bakhshi, Bravo-Biosca and Mateos-García, 2014^[9]).

In agriculture, data captured by sensors on farm equipment are combined with weather, climate and soil data, to provide information about production processes. This often involves transfers of different types of data, including personal or commercially sensitive information, from and to other countries. The use of all this data together with data analytics (i.e. precision agriculture) provides productivity gains by optimising the use of agriculture-related resources. These include, but are not limited to,

savings on seed, fertiliser and irrigation, as well as farmers' savings in time (OECD, 2017_[3]). By some estimates the economic benefits from precision agriculture can be around USD 12 billion annually for the United States. This represents about 7% of the total value added of USD 177 billion contributed by farms to the gross domestic product (GDP) of the United States in 2014 (Schimmelpfennig and Ebel, 2016_[10]).

Data and data analytics enable the business models of online platforms

Online platforms have become a key element of the digital economy as they support many economic and social activities on line. Most of them are large ICT companies such as Apple and Google. However, increasingly traditional (non-ICT) companies such as Nike and TomTom have established online platforms. These firms generate data as a by-product of their actual business activity to support the sales of goods and services. Companies such as John Deere and DuPont Pioneer, for example, take advantage of the “industrial Internet”. They integrate sensors with their latest equipment to build online platforms that help farmers manage their fleet and decrease downtime of their tractors, as well as save on fuel (OECD, 2017_[3]).

As a common and major characteristic, all online platforms benefit from data-enabling multi-sided markets. Activities on one side of the market go hand in hand with the collection of data, which is exploited and used on the other side of the market (OECD, 2015_[1]). These online platforms also take advantage of network effects emerging on at least one side (OECD, 2019_[14]).

The business model of online platforms therefore relies heavily on the combination of the use of data and these network effects that typically affect all sides of the market. As the utility for users on all sides of the market increases with the increase in their numbers, users are more willing to pay for access to a bigger network and/or to contribute with their own data. Combined with the increasing returns to scale and scope the data enable, these network effects can lead to huge profit margins for platform providers (OECD, 2015_[1]; OECD, 2019_[14]).

Online platform providers can combine various revenue models and data-enabled services across all sides of the markets of their platforms. Li et al. (2019_[12]), for instance, show that the online platform Amazon Marketplace generates revenue through a wide number of data-enabled services. These include the following:

- a buyer-seller matching service
- a service to sellers to promote their products to some individuals²
- the licensing of access to its internally collected customer behaviour data
- the use of data to improve its own algorithms.

In addition, based on its data-driven understanding of customer needs, Amazon is also offering its own products that compete directly with independent sellers on its platform. Based on the vast amounts of data it can access, these products can be customised and priced to meet specific groups of consumers.³

Private equity investments in data and (big) data analytics continue to increase

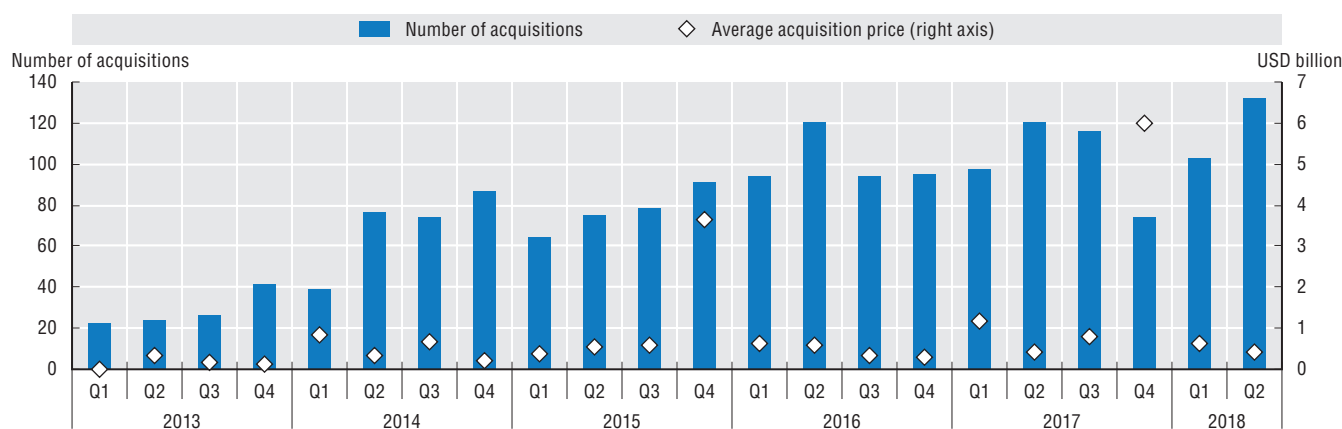
The importance of data and data analytics is reflected in the growing number of mergers and acquisitions of data-intensive firms. In 2013, for example, Monsanto acquired the Climate Corporation, an agriculture analytic firm, for USD 1.1 billion. In 2015, IBM acquired a majority share of the Weather Company, a weather forecasting and analytic company, for over USD 2 billion (Waters, 2015_[13]). Meanwhile, Alibaba invested USD 4 billion between 2016 and 2018 to acquire Lazada, a leading e-commerce platform. The annual number of acquisitions increased from more than 100 in 2013 to more than 400 in 2017, with the average price paid exceeding USD 1 billion in some quarters (Figure 5.1).

Business adoption of big data is growing but varies by sector and type of data

The ICT sector remains the most intensive user of big data, with social media data playing the most important role. More than half of all ICT firms in the European Union used social media data in 2018 (Figure 5.2). Besides the ICT sector, utilities (including electricity, gas, steam, air conditioning and water supply businesses) and transportation & logistics are also highly intensive users of big data. Around 20%

of these firms used big data in 2018, focusing on geolocation data of portable devices. Utility businesses, in addition, also use data originating from smart devices or sensors intensively. These two sectors also had the biggest increase in big data adoption between 2016 and 2018 with around 25% more of their businesses adopting the technologies in the European Union.

Figure 5.1. Trends in the acquisition of big-data and analytics firms, Q1 2013-Q2 2018

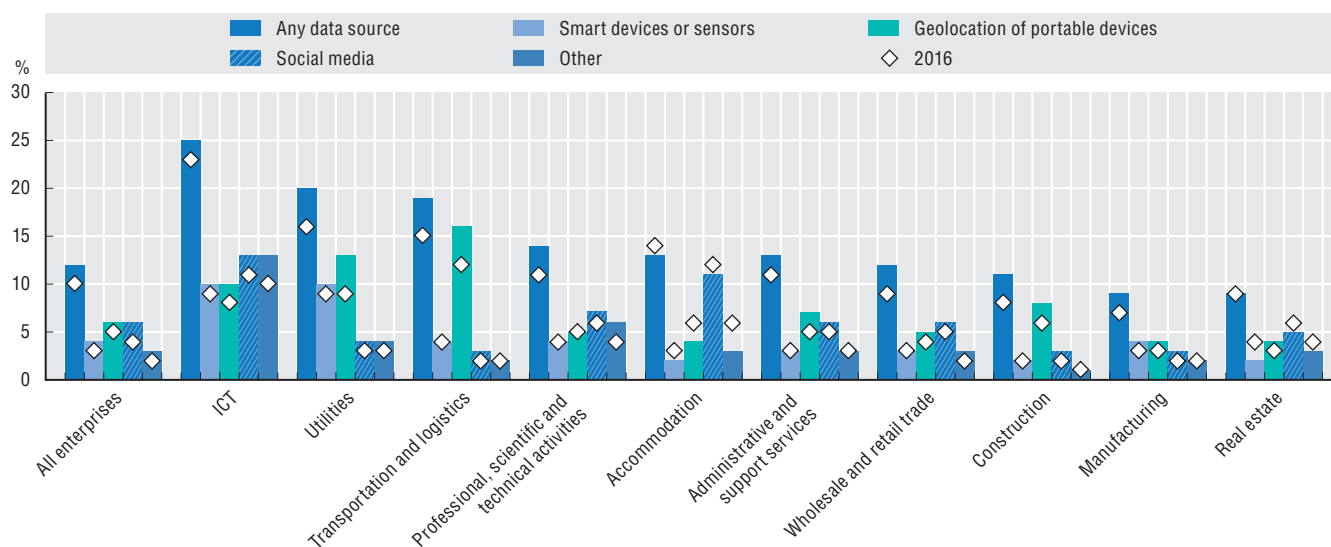


Notes: In the figures presented here, the sample is restricted to firms founded after 2010 (i.e. no more than five years-old in 2016) that attracted equity funding over 2011-16. Equity funding includes venture capital and other risk finance such as business angel investments or debt financing. Digital-related sectors are identified on the basis of correspondence between the sectors available in the database with the ISIC Rev.4 industry list.

Source: OECD (2019^[14]), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, <https://doi.org/10.1787/276aaca8-en>.
StatLink <https://doi.org/10.1787/888934192110>

Figure 5.2. Business use of big data by data source and industry in the European Union, 2018

As a percentage of enterprises



Source: OECD based on Eurostat, *Digital Economy and Society Statistics* (database), <http://ec.europa.eu/eurostat/web/digital-economy-and-society/data/comprehensive-database> (accessed in December 2019).

StatLink <https://doi.org/10.1787/888934192129>

Other sectors also saw a significant increase in big data adoption. However, adoption of big data in the real estate and accommodation sector either stagnated or even fell slightly in 2018. That said, the adoption of big data can vary significantly across countries. As highlighted in Chapter 4, adoption of data analytics by businesses has increased, particularly among large firms in Germany, France, Finland, Korea and Portugal.

The geographic difference is particularly noticeable in the manufacturing sector. The adoption of big data analytics has increased by around 30% on average in manufacturing across the European Union between 2016 and 2018. However, Germany's manufacturing sector experienced an increase of 140% within this same period. In 2018, 12% of all manufacturing firms in Germany used big data (compared to 9% in the European Union). In the United States, the share of manufacturing plants that adopted data-driven decision making nearly tripled between 2005 and 2010, jumping from 11% to 30% (Brynjolfsson and McElheran, 2019^[4]). The authors note this rapid diffusion was uneven and that economies of scale (firm size), as well as complementarities between investments in skills and competences, can explain to a significant extent the variation.

Data and data analytics can help enhance well-being and address global challenges, including the COVID-19 crisis

The full impact of data and data analytics goes beyond its positive effects on productivity growth and innovation. The use of data can also contribute directly to the well-being of citizens. Quantification remains challenging, however, because market transactions do not capture many if not most of the benefits related to use of data (OECD, 2015^[1]).⁴ For example, data access and sharing are needed to enhance public service delivery; tackle longstanding issues that require new ways and tools to leverage data; and identify and address emerging governmental and societal needs and emergencies. In science and technology, data access and sharing provide a range of benefits to society such as reproducibility of scientific results, facilitating cross-disciplinary co-operation (OECD, 2020^[15]). Data have also been critical during emergency response such as during the 2011 Fukushima nuclear incident, the 2014-16 Ebola outbreak in West Africa and, more recently, during the COVID-19 crisis.

At the early stage of the COVID-19 pandemic, the collection and sharing of data became essential to understand and respond to the scale of the public health challenge. Of particular importance to an effective frontline response are data concerning the spread of the virus. These include the location and number of new confirmed cases, rates of recoveries and deaths, and the source of new cases (international arrivals or community transmission). Access to and sharing of data are also crucial to assess and improve the capacity of the health care system to address the crisis and the effectiveness of containment and mitigation policies that restrict the movement of individuals. Transborder co-operation in the collection, processing and sharing of these data (subject to necessary and proportionate safeguards) may expedite effective and united global frontline responses.

Governments are turning to a wide array of digital technologies and advanced analytics to collect, analyse and share data for frontline response to the COVID-19 crisis (Dunant, 2020^[16]). Above all, most countries are leveraging the widespread use of mobile phones given the more than 7.85 billion subscriptions worldwide as of 2018 (ITU, 2020^[17]). This includes in particular the collection and sharing of geolocation and proximity data. These data are generated in two ways. On the one hand, they can be derived from mobile call data records, i.e. data produced by telecommunication service providers on telephone call or other telecommunications transactions. On the other, they can be collected from mobile applications (apps) made for COVID-19 response. Government initiatives to improve the effectiveness of frontline responses to COVID-19 are explored later in this chapter.

In addition, symptom tracking apps are being deployed to help slow the outbreak. They help researchers better understand symptoms linked to underlying health conditions. This, in turn, helps identify i) how fast the virus is spreading in different areas; ii) high-risk areas; and iii) who is most at risk. According to researchers, the C-19 COVID Symptom Tracker app, developed in the United Kingdom, can help collect data to reveal essential information about the symptoms and progress of COVID-19 infection in different people. It can also help researchers understand why some individuals develop more severe or fatal disease, while others have only mild symptoms due to COVID-19 (King's College London et al., 2020^[18]). Data and data analytics can provide valuable indicators on population movements and infections over time, especially when mobility and contact tracing data are poor. However, their mass collection and analysis raise data governance and privacy concerns. These issues are discussed in more detail in Chapter 6.

Data sharing and re-use beyond borders

A significant share of the global volume of data and its processing will rarely be located within just one organisation or even a single country. They will instead be distributed around the globe, reflecting the global distribution of economic and social online activities. Data flows, including across borders, are critical for two reasons. On the one hand, they are a condition for information and knowledge exchange. On the other, they are also vital for the functioning of a globally distributed digital economy. In addition, data flows can facilitate collaboration between governments to improve their policy making at international level. Finally, they can help address global challenges such as the Sustainable Development Goals or the management of pandemics such as COVID-19.⁵

Different approaches can enhance data access and sharing

Three approaches to enhancing data access and sharing have been most prominently discussed in the literature and by policy makers: **open data**, and more recently **data markets** and **data portability**. Besides these three, a wide range of other approaches exist with different degrees of data openness. The level of access responds to the various interests of stakeholders and their risks in data sharing such as (bilateral or multilateral) engagements in **data partnerships**. Many approaches are based on voluntary and mutually agreed terms between organisations. Others are mandatory, such as the Right to Data Portability under the European Union (European Union, 2016_[19]) General Data Protection Regulation (GDPR) (Art. 20) or Australia's recently proposed Consumer Data Right (see OECD (2019_[14]) for more examples).

Contractual agreements and data markets

Increasingly, businesses are recognising the opportunities of commercialising their proprietary data (OECD, 2015_[1]). Some organisations offer their data for free (via open access), especially non-governmental organisations and governments as highlighted below. However, many businesses engage in bilateral arrangements to sell or license their data. For example, the French mobile ISP Orange acts as a data provider. Its Floating Mobile Data technology collects mobile telephone traffic data, which determine speeds and traffic density at a given point in the road network. The anonymised mobile telephone traffic data are sold to third parties to identify “hot spots” for public interventions or to provide traffic information services.

Data commercialisation remains below its potential, even among data-intensive firms, despite the increasing interest of organisations to commercialise their data and meet the growing demand for data. In a Forrester Research survey of almost 1 300 data and analytics businesses across the globe, only one-third of respondents reported commercialising their data. High tech, utilities and financial services rank among the top industries commercialising their data, while pharmaceuticals, government and health care were at the bottom of the list (Belissent, 8 March 2017_[20]).

With the emergence of data intermediaries, the commercialisation could become more mainstream. Data intermediaries provide potential sellers and buyers with services such as standard licence schemes, and a payment and data exchange infrastructure. With more such intermediaries, even less data-savvy firms may find it easier to commercialise their data.

Data portability

Data portability is often regarded as a promising means for promoting cross-sectoral re-use of data and for strengthening control rates of data for both individuals and businesses. For individuals, data portability could help strengthen control rights over personal data. It could do the same for businesses for their data, especially small and medium-sized enterprises (SMEs) (Productivity Commission, 2017_[21]). Data portability provides restricted access through which data holders can provide customer data in a commonly used, machine-readable structured format. These data are delivered either to the customer or to a third party chosen by the customer.

Several countries have prominent data portability initiatives. In 2010, the United States initiated My Data, which includes the Green Button (US Department of Energy, n.d._[22]). In 2011, the United Kingdom launched the Midata data portability initiative (BIS, 2011_[23]). In 2016, the European Union approved the

5. ENHANCING DATA ACCESS, SHARING AND RE-USE

Right to Data Portability (Art. 20) (European Union, 2016^[19]) GDPR. Most recently, Australia proposed its Consumer Data Right (CDR).

Data portability initiatives may vary significantly in terms of their nature and scope across jurisdictions. The GDPR Right to Data Portability (Art. 20), for instance, states that

the data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance.

This GDPR differs in important ways from the “data portability” concept explored in the voluntary-based Midata initiative in the United Kingdom.

To what extent data portability may effectively empower individuals and foster competition and innovation remains to be seen. Estimates on the costs and benefits of data portability are still rare. Although not specific to data, other portability studies suggest that data portability may have overall positive economic effects, specifically by reducing switching costs. One study on limitations to move mobile apps across platforms (such as changing from the operating system of Apple to one of another smartphone) can be a barrier. Enabling app portability would help reduce switching costs, which are estimated to be between USD 122 and USD 301 per device (OECD, 2013^[24]; iClarified, 2012^[25]).

Open data

Open data is the most prominent approach to enhance access to data and the most extreme form of data openness (OECD, 2015). In the public sector, open government data has been promoted for many years by initiatives such as data.gov (United States), data.gov.uk (United Kingdom), data.gov.fr (France) or data.go.jp (Japan) (Ubaldi, 2013^[26]).

Open data should be accessed on “equal or non-discriminatory terms” (OECD, 2006^[27]), limiting the conditions under which data can be provided via open access. In most cases, for instance, confidential data such as personal data cannot be shared via open access. Furthermore, as highlighted above, open data is expected to be provided for free or at no more than the marginal cost of production and dissemination. Therefore, businesses that want to commercialise their data, either directly (by selling data) or indirectly (by providing value-added services), may find open data less attractive.

Organisations in the public and private sector increasingly recognise that non-discriminatory access is crucial for maximising the (social) value of data: it creates new business opportunities, as well as economic and social benefits. Assessing the resulting economic and social benefits of moving towards open data, however, remains challenging. As highlighted by Dan Meisner, Thomson Reuters’ Head of Capability for Open Data, indirect benefits and network effects at play “don’t really fit very well into an Excel model for calculating your internal rate of return” (ODI, 2016^[28]).

Restricted data-sharing arrangements can sometimes be more appropriate. In some cases, data are considered too confidential to be shared openly with the public. In others, there are legitimate (commercial and non-commercial) interests that oppose such sharing. Privacy, intellectual property (e.g. copyright and trade secrets), and organisational or national security concerns, legitimately prevent open sharing of data. In these cases, however, data users within a restricted community may still have a strong economic and/or social rationale for sharing data, under voluntary and mutually agreed terms.

It is common to find restricted data-sharing agreements in several areas. These include digital security, science and research, and as part of business arrangements for shared resources (e.g. within joint-ventures). These voluntary data-sharing arrangements can be based on commercial or non-commercial terms depending on the context. The following sections highlight two types of arrangements. First, **data partnerships** recognise that data sharing provides significant economic benefit both to data users and data holders. Second, **data for societal objectives** initiatives share data to support societal objectives.

Data partnerships (including data public-private partnerships)

In data partnerships, organisations agree to share and mutually enrich their data sets, including through cross-licensing agreements. One big advantage is the facilitation of joint production or co-operation with suppliers, customers (consumers) or even potential competitors. This also enables the

data holder to create additional value and insights that a single organisation would not be able to create. This provides opportunities “to join forces *without merging*” (Konsynski and McFarlan, 1990^[29]). Examples include the following:

- Nectar, a UK-based programme for loyalty cards, pooled data with firms such as Sainsbury (groceries), BP (gasoline) and Hertz (car rentals). “Sharing aggregated data allows the three companies to gain a broader, more complete perspective on consumer behaviour, while safeguarding their competitive positions” (Chui, Manyika and Kuiken, 2014^[30]).
- DuPont Pioneer and John Deere launched a joint venture in 2014. It aimed to develop a joint agricultural data tool and link Pioneer’s Field360 services, a suite of precision agronomy software, with John Deere Wireless Data Transfer architecture, JDLink and MyJohnDeere (Banham, 2014^[31]).
- Telefónica collaborated with organisations such as Facebook, Microsoft and UNICEF to exchange data of common customers (based on customers’ consent) for Telefónica’s personalised AI-enabled service Aura. Thanks to this collaboration, customers will be able to talk to Aura through Telefónica’s own channels and some third-party platforms like Facebook Messenger. In the future, they will also be able to talk through Google Assistant and Microsoft Cortana.

Similar arrangements exist in the form of public-private partnerships. For example, Transport for London (TfL), a local government body responsible for the transport system in Greater London (United Kingdom), forged new strategic partnerships with major data, software and Internet services providers such as Google, Waze, Twitter and Apple. In some cases, this partnership enabled TfL to access new data sources and crowdsource new traffic data (“bringing new data back”), to undertake new analysis. In doing so, TfL could gain access to updated navigation information (on road works and traffic incidents) and could enhance the efficiency of its planning and operation.

Data partnerships (including data public-private partnerships) raise several challenges (OECD, 2019^[14]). Ensuring a fair data-sharing agreement between partners can sometimes be challenging, particularly when they have different market power. Considerations of privacy and international property rights may also limit the potential of data partnerships. These considerations can make it harder to sustain data sharing (see for comparison barriers to knowledge sharing during pre-competitive drug discovery). Where data partnerships involve competing businesses, data sharing may increase the risk of (implicit) collusion, including the formation of cartels and price fixing. In the case of data public-private partnerships, the double role of governments as an authority and service (data) provider may also create challenges. In this case, questions have been raised about what types of rules should apply for this type of data sharing, and what should the private sector exchange in return for the data.

Data for social good initiatives

Private-sector data can also be provided (donated) to support societal objectives, ranging from science- and health care research to policy making. In an era of declining responses to national surveys, the re-use of public- and private-sector data can significantly improve the power and quality of statistics. This is true for both OECD countries and developing economies (Reimsbach-Kounatze, 2015^[32]).

The re-use of private-sector data also provides new opportunities to better inform public policy making. Close to real-time evidence, for instance, can be made available to “nowcast” policy relevant trends (Reimsbach-Kounatze, 2015^[32]). Other examples range from trends in the consumption of goods and services to flu epidemics and employment/unemployment trends. The monitoring of information systems and networks can also identify malware and cyberattack patterns (Choi and Varian, 2009^[33]; Harris, 18 April 2011^[34]; Carrière-Swallow and Labbé, 2013^[35]). Some of these arrangements have been classified as “data philanthropy” to highlight the gains from the charitable sharing of private-sector data for public benefit (United Nations Global Pulse, 2012^[36]).⁶

Data access and sharing can generate significant positive social and economic benefits

Available evidence shows that enhancing data access and sharing can generate positive social and economic benefits for data providers (direct impact), their suppliers and data users (indirect impact) and for the wider economy (induced impact). These benefits are generated thanks to the following:

- greater transparency, accountability and empowerment of users, for instance, when open data is used for (cross-subsidising) the production of public and social goods

- new business opportunities, including for the creation of start-ups and in particular for data intermediaries and mobile app developers
- competition and co-operation within and across sectors and nations, and including the integration of value chains
- crowdsourcing and user-driven innovation
- increasing efficiency due to linkage and integration of data across multiple sources (OECD, 2019_[14]).

The quantification of the overall benefits of enhancing data access and sharing remains challenging.⁷ Recent available studies by sector (public vs. private sector) provide a rough estimate of the magnitude of the relative effects of enhancing data access and sharing. Overall they suggest that enhancing data access and sharing can increase the value of data to holders (direct impact). Further, it can help create 10 to 20 times more value to data users (indirect impact) and 20 to 50 times more value for the wider economy (induced impact). In some cases, however, enhancing data access and sharing may also reduce the producer surplus of data holders.

Deloitte (2013_[37]), which was used as basis for the Shakespeare Review of the United Kingdom (BIS, 2013_[38]), assessed the economic impact of access to public sector information (PSI) in the United Kingdom.⁸ The direct economic impact (as revenues of PSI holders) is estimated at GBP 0.1 billion (USD 0.13 billion). Meanwhile, the indirect impact (on data users and suppliers of private sector in health data) is estimated to be between GBP 1.2 billion (USD 1.6 billion) to GBP 1.8 billion (USD 2.4 billion) per year.⁹ The wider indirect and induced impact of PSI was conservatively estimated at around GBP 5 billion (USD 6.5 billion) per year. This included, for instance, time saved as a result of access to real-time travel data, which is valued at GBP 15 million (USD 19.5 million) to GBP 58 million (USD 75 million). Overall, this led to an estimate of between GBP 6 billion (USD 8 billion) to GBP 7 billion (USD 9 billion), or around 0.5% of GDP.

A study by the McKinsey & Company (2013_[39]) looks at the benefits of re-using both public and private-sector data. The study examines seven areas of the global economy: education, transportation, consumer products, electricity, oil and gas, health care and consumer finance. It estimates that re-use of data across these seven areas could help create value worth USD 3 trillion a year worldwide.¹⁰ By scaling the results of this study to the G20 economies, Lateral Economics (2014_[40]) estimates that open data could increase G20 output by around USD 13 trillion over the next five years. The authors note this increase “would boost cumulative G20 GDP by around 1.1 percentage points of the 2% growth target over five years” (Lateral Economics, 2014_[40]). Similar scaling for Australia suggests that “more vigorous open data policies could add around AUD 16 billion per annum to the Australian economy” (this would represent almost 1% of GDP or USD 13 billion).

More recent studies are available at organisational level. Recent estimates based on open data provided by TfL, for instance, strongly confirm the positive net benefits of open data (Deloitte, 2017_[41]). The Deloitte study shows that re-use of TfL’s open data was generating annual economic benefits and savings of up to GBP 130 million (USD 168 million) for TfL customers, road users, London and TfL itself. This includes a gross value added of GBP 12 million to GBP 15 million (USD 15 million to USD 19 million) per year for businesses, which also directly created more than 500 jobs. However, this does not account for the significant contribution of TfL’s open data to improving societal outcomes, facilitating innovation and improving the wider environment (e.g. air quality and lower emissions).

IDC and the Lisbon Council (2018_[42]) assess the data market size and the GDP impact of the data economy in the European Union. They focus on the value added created from data re-use, including the provision of data and its exploitation in the private sector.¹¹ The direct impact is estimated by the volume of the data market as a proxy (i.e. revenues of data suppliers and adjusted through including imports and excluding exports). According to the study, the data market volume in the European Union was estimated at EUR 59 billion in 2016 and EUR 65 billion in 2017 (an increase of roughly 20% year on year). The indirect impact (i.e. the impact on data suppliers and the impact on data users through innovation and efficiency gains) was more than 50% of total impact in 2017. Overall, the study suggests an overall impact of the data economy on GDP of 2.2% (EUR 306 billion) in 2016 and 2.4% (EUR 336 billion) in 2017.

Overall, these and other similar studies suggest that enhancing data access and sharing can help generate social and economic benefits. For the public sector, these benefits are worth between 0.1%

and 1.5% of GDP. When they include private-sector data, the benefits range from 1% and 2.5% of GDP. In a few studies, they rise up to 4% of GDP (OECD, 2019^[14]).

Social and economic activities increasingly rely on transborder flows

The creation of economic and social value increasingly depends on the ability to move and aggregate data across a number of locations scattered around the globe. These data flows enable firms to co-ordinate their research and development (R&D), supply, production, sales and post-sales processes effectively (United States Department of Commerce, 2016^[43]; Casalini and López González, 2019^[44]). Many manufacturing companies, for instance, use data flows to monitor the status, performance and condition of their machines in different locations. Boeing, for example, uses data generated by its 737 models, around 20 terabytes of data for every inflight hour, to diagnose problems in real time (Pepper and Garrity, 2014^[45]). Volkswagen and Amazon Web Services, as another example, announced the co-development of the “industrial cloud” in March 2019 to connect “data from all machines, plants and systems in all factories”.¹² Real-time data are then aggregated at a global level and potentially monetised via a new service.

Transborder data flows are especially important for SMEs, enabling a new breed of “micro multinationals” that is “born global” and constantly connected (MGI, 2016^[46]). Start-ups, for example, rely on cross-border data flows to deliver their digital services as a platform. At the same time, they also collect transaction and consumer behaviour data in various locations. These data must then be transferred across borders to be stored, aggregated and analysed. Finally, insights based on aggregated global data serve as the basis for commercial services that can be delivered in multiple locations (e.g. targeted advertising, or demand forecasting, and price elasticities of consumers).

Transborder data access and sharing is also relevant for improving well-being. For example, the National Health Service in England outsourced the processing of MRI scans using the company Alliance Medical, which has around 200 imaging sites across Europe. Meanwhile, the Swedish company Hermes Medical Solutions offers cloud-based software applications to share medical images across 30 countries, though 95% of patient data are stored in Sweden.

Overall, this means that data also increasingly underpin international trade, reducing trade costs. In so doing, they support growing trade in goods and enable trade in services previously considered non-tradeable (OECD, 2017^[47]; 2018^[48]; 2019^[49]).¹³ Some estimates suggest that the value of cross-border data flows has exceeded the value of cross-border merchandise trade. MGI (2016^[46]), for instance, estimates the international flow of data added USD 2.8 trillion to the global economy (more than trade in goods); this was expected to grow to USD 11 trillion by 2025.¹⁴

Estimates based on volume (i.e. measured in bytes) can only partially help assess the real value of data and data flows given they have little connection with the information contained within each data “unit” (OECD, 2019^[50]). The transfer of a megabyte of new car design, for example, carries a different value than a megabyte of an individual’s purchase history. Cisco (2018^[51]) shows that video accounted for 75% of all Internet Protocol traffic in 2017, the greatest single category of online data flow (Chapter 3).¹⁵

Other estimates of the “value” of transborder data flows are based on costs associated with restricting them (lower costs would suggest a lower value). For example, the US International Trade Commission (2014^[52]) estimates the GDP of the United States would be 0.1% to 0.3% higher if foreign digital trade barriers were removed. Similarly, for the European Union, barriers to transborder data flows are estimated to reduce GDP by 0.4% to 1.1%, depending on the strength of data localisation requirements (van der Marel, Lee-Makiyama and Bauer, 2011^[53]). Another study suggests that data regulations lead to a reduction of real GDP in the European Union by 0.48% (Bauer, Ferracane and van der Marel, 2016^[54]).

Connectivity, understood as high-quality access to communication services at competitive prices, is the key enabler of data flows among countries. As discussed in Chapter 3, continued investment in backbone connectivity, including in submarine cables, is essential. This allows countries to keep pace with data transmission requirements and to support data flows with each other.

The installed capacity of submarine cables can provide a complementary, although only indicative, view on which global regions are most integrated in terms of cross-border data flows. Available evidence from TeleGeography (n.d.^[55]) suggests that some parts of the globe are much more connected than

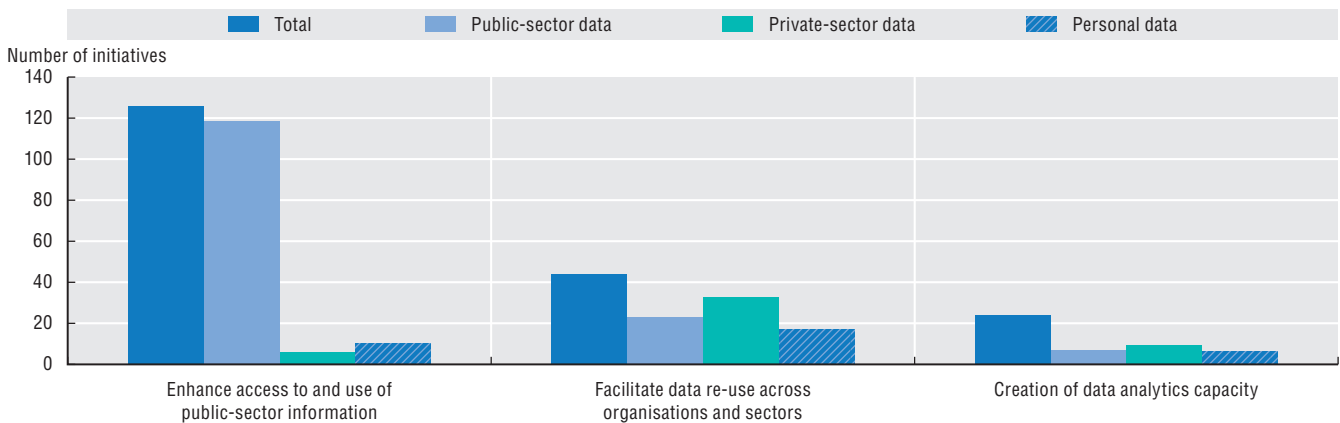
others. The trans-Atlantic route between the East Coast of the United States and Europe, and the trans-Pacific route from the West Coast of the United States to East Asia, for example, are well connected. However, there is significant and increasing capacity in backhaul and backbone connectivity in other regions as well (Chapter 3).

Facilitating data sharing and re-use: An overview of government initiatives

Governments play a major role in encouraging, facilitating and enhancing data access and sharing through policy action and governance frameworks. The leadership role of governments is also reflected in their ability to foster and enhance access to and sharing of public sector data.

All OECD countries and most partner economies have one or more initiatives to enhance access to and sharing of data in their economies. The scope of these initiatives may vary significantly across countries, however. While all these countries had initiatives that foster and enhance access to and sharing of public sector data in 2018, significantly fewer countries targeted private-sector data (Figure 5.3). Even fewer governments had initiatives to improve the data analytic capacity in their countries.¹⁶

Figure 5.3. Government policy initiatives enhancing data access and sharing, 2017-18



Notes: Based on two country surveys, the most recent of which – the EASD Policy Questionnaire – was conducted between June and September 2018 and covered 20 countries plus the European Union. This survey was complemented by the responses to the OECD Digital Economy Outlook (DEO) Policy Questionnaire. This questionnaire included additional 16 countries, many of which are partner economies. As a result, this report analysed 205 policy initiatives across 37 countries. StatLink contains more data.

Source: OECD calculations based on the 2018 OECD questionnaire on policies for enhancing access to and sharing of data (EASD Policy Questionnaire) and the 2017 OECD Digital Economy Policy Questionnaire.

StatLink  <https://doi.org/10.1787/888934192148>

Governments are leading by example in enhancing access to and sharing of public sector data

Access to open government data and public sector information

The large majority of government initiatives on data sharing and re-use focus on access to and sharing of public sector data (almost 65% of all initiatives), with most aiming at enabling open access to government data (open government data). Even before the emergence of open data initiatives in the United States, the United Kingdom, France, Japan or Singapore, governments recognised the need to provide public sector data “at the lowest possible cost, preferably at no more than the marginal cost” as stated in OECD (2008_[56]).¹⁷ This motivated the establishment of PSI initiatives.

In many countries, PSI initiatives were legally backed by freedom of information legislation, and were therefore broader in scope than open data initiatives.¹⁸ As a result, many countries have PSI initiatives, while others have open data initiatives or both. This is the case for EU member states, which are subject to Directive (EU) 2019/1024 of 20 June 2019 on Open Data and the Re-Use of Public Sector Information. This directive replaces the Public Sector Information Directive (Directive 2003/98/EC). That said, a general trend towards the establishment of open data portals can be observed across the OECD.

Facilitating data sharing within the public sector

There is a noticeable trend towards facilitating data sharing within the public sector (almost 15% of all initiatives on public sector data). This trend is motivated by governments' commitment to become more data-driven and to exploit technological trends such as big data and artificial intelligence (AI). Australia's data sharing and release legislation (DS&R legislation) (Box 5.1) is a prominent example. Estonia's Information Sharing Data Sheet (X-Road) initiative aims to facilitate data exchange and linkage by interconnecting the country's main national databases. It is motivated by the "once only" principle according to which public agencies should only collect data not previously maintained in any other public sector databases (Information System Authority [Estonia], 2019^[57]). Similarly, Singapore set up the Government Data Architecture on 1 October 2019 to improve data quality and speed of access to data and to facilitate the secure use and sharing of data across public agencies.¹⁹ Another example, but with a focus on capacity building includes the United Kingdom Government Data Ethics Framework. It aims to ensure that public servants from across disciplines understand insights from data and emerging technologies and use data-informed insight responsibly (DCMS, 2018^[58]).²⁰

Box 5.1. Balancing the benefits with the risks: Australia's data sharing and release legislation

On 1 May 2018, in response to recommendations from the Productivity Commission Data Availability and Use Inquiry, the Australian government committed to reforming its national data governance framework. To that end, it developed new data sharing and release legislation (DS&R legislation) with five goals. It sought to i) promote better sharing of public sector data; ii) build trust in use of this data; iii) establish consistent and appropriate data safeguards that dial up or down depending on sensitivity of data; iv) enhance the integrity of the data system; and v) establish institutional arrangements.

The DS&R legislation recognises that greater sharing of data can lead to a variety of benefits. These include more efficient and effective government services for citizens; better informed government programmes and policies; greater transparency around government activities and spending; economic growth from innovative data use; and research solutions to current and emerging social, environmental and economic issues.

To balance these benefits with the risks and enhance trust in data sharing and re-use, Australia established the Office of the National Data Commissioner and the Data Integration Partnership for Australia, among other entities.

Geospatial and transportation data: A highly valued public sector data

Opening geospatial data (e.g. maps) and transportation data ranked high on the agenda of public sector data initiatives (representing almost 8% of the initiatives). Geospatial (geo-) data provide information about specific geographic locations. They are typically used for geographic information systems (GIS).²¹

The most prominent examples of GIS are digital maps, but geo-data may also include data on addresses, cadastral parcels, administrative units, geology, and agri- and aqua-cultural facilities. Further, it may include transportation data to the extent that data cover geolocation information (e.g. data on traffic flows and public transportation schedules). The combination of all these data has become the foundation for many location-based services and therefore recognised as critical for the functioning of multimodal transport. This may explain why many countries have classified geospatial and/or transportation data among their high-value data sets, such as the Geocoded National Address File in Australia. In Switzerland, the Federal Office of Transport is looking to facilitate the exchange of data between the various public and private actors active in the Swiss public transport system. It is therefore focusing on geo-data, price data of transportation services and operational data.

Few countries are facilitating or regulating data access and sharing within the private sector

Few countries have initiatives to facilitate data sharing within the private sector (almost 15% of all initiatives). However, sharing and re-use of private-sector data was the most frequently cited emerging challenge (followed by public-private partnerships) among countries that responded to the 2018 OECD questionnaire on policies for enhancing access to and sharing of data (EASD Policy Questionnaire).

Voluntary and collaborative approaches

Most initiatives (around 55%) to facilitate or regulate data access and sharing within the private sector are voluntary. These initiatives tend to be used where the risks of detrimental consequences of mandatory access and sharing outweigh the expected public benefits. Data access regulation, for example, could undermine incentives to invest in data. In other cases, regulation might not be granular enough for specific issues, and would thus reduce innovation and competition. Against these risks, and to incentivise and co-ordinate actions that facilitate data access and sharing in the private sector, many governments have incentives for voluntary initiatives. Two major types of voluntary government-led initiatives are among the most cited by survey respondents: i) contract guidelines; and ii) data partnerships, including public-private partnerships.

Contract guidelines define a set of contractual clauses based on defined principles. They constitute the default position for parties when negotiating their data-sharing agreements, with a focus on potentially contentious issues. Since the guidelines are voluntary, parties can deviate from the proposed contractual clauses at their will (freedom of contract). Parties would typically do so if such deviation would better reflect their common interests and the context of their agreements.

Examples of government initiatives include the Contract Guidance on Utilisation of AI and Data, formulated by Japan's Ministry of Economy, Trade and Industry. This guidance elaborates issues and factors to be considered when drafting a contract on the utilisation of data and AI. It is intended to be used as a reference when private businesses conclude contracts related to data sharing (Data Section) or development and use of AI-based software (AI Section). The Data Section categorises data utilisation contracts into three types: i) data provision contracts; ii) data creation contracts; and iii) data sharing (platform) contracts.²² In this context, Japan also revised the Unfair Competition Prevention Act in 2018 to develop an environment where data can be exchanged with confidence. The act defines unauthorised acquisition, use and disclosure of "protected data" that meet the statutory requirements as unfair competition and provides civil measures against such misappropriation.

In the United States, as another example, the American Farm Bureau Federation (AFBF), together with commodity groups, farm organisations and agriculture technology providers, helped establish the Privacy and Security Principles for Farm Data to address controversial issues related to questions of the "ownership" of agricultural data. As of 1 April 2020, 37 organisations had signed onto the Core Principles, pledging to incorporate them into their contracts with farmers. To verify compliance with the Core Principles, AFBF and the other interested stakeholder groups formed a non-profit organisation, AG Data Transparency Evaluator. This entity audits companies' agricultural data contracts and offers a seal of approval for those that meet the criteria (AG Data Transparent, 2016_[59]).

Singapore, as another example, launched the Trusted Data Sharing Framework in June 2019. It lays out key risk-based business, legal, technical and operational considerations to guide businesses when exploring data partnerships, including involving third-party intermediaries. The framework is intended to establish a set of baseline practices by providing a common "data sharing language". It suggests a systematic approach to the broad considerations for establishing trusted data-sharing partnerships. Legal templates are provided to kick-start discussions in response to industry feedback that businesses often go into protracted legal negotiations in setting up their data-sharing partnerships.

Aside from contract guidelines, Singapore also provides a regulatory "sandbox" as a safe environment for industry to engage the regulators on novel use of data. The sandbox allows companies to engage the regulator on new ways to use data. At the same time, it allows the regulator to be in line with industry development on data use. This could manifest either through new data generated by new technology, new technology enabling new uses of data or new application(s) of existing technology. This approach also informs the regulator and new developments in industry. Finally, it assesses the need for policy review to ensure a supportive regulatory environment for growth of data ecosystem.

Data partnerships enable organisations to share and mutually enrich their data sets, including through cross-licensing agreements. A number of governments encourage the establishment of data partnerships, both within the private sector and/or between the private and public sectors. Many of these initiatives are enabled by open access to public sector data. In Chile, for instance, the government has engaged in agreements on open data with academic and research institutions for the re-use of data in open format.

Other data partnerships are incentivised through research-related funding. Industrial Data Space (IDS), for example, enables better data control and agency across all domains. Co-ordinated by the Fraunhofer Gesellschaft, IDS uses an open, vendor-independent architecture of a peer-to-peer network. IDS has been funded by the German Federal Ministry of Education and Research since 2015 with approximately EUR 13 million.

In some other initiatives, the government's role has been to incentivise and “orchestrate” data partnerships. Either it acts as (independent) trusted third party or it engages the private sector in public-private partnerships. The Data Integration Partnership for Australia, “an investment to maximise the use and value of the Government's data assets” (Australian Government, 2017^[60]), was presented earlier.

Another example is Japan's Certification System for data sharing, which allows data-sharing companies to request data provided to relevant ministries and agencies. The government then provides support, in particular through tax incentives and administrative guidance. However, it could also revoke accreditation in some cases. The Digital Hub Denmark is an example for data public-private partnerships, where both public and private-sector actors agree to mutually share their data. The partners comprise the government, the Confederation of Danish Industry, the Danish Chamber of Commerce and Finance Denmark. The partnership aims to make Denmark one of the main European tech-hubs within AI, Internet of Things and big data. The Digital Hub will improve companies' access to talent and investments, and facilitate the matchmaking between larger companies, start-ups and universities. Access to data thus constitutes just one element of the overall objective of the partnership.

Data of public interest

Among the mandatory approaches, the most common are data-sharing agreements restricted to trusted users (restricted data sharing). These include promoting data sharing between the private and public sector with a focus on “data of public interests” or within network industries such as transportation and energy for ensuring interoperability of smart services. A number of countries have adopted the concept of data of public interest, but its scope varies significantly.

In some countries, data of public interest explicitly refers to private-sector data (of public interest), while in others it refers to public sector data. Sometimes both private and public sector data, as well as personal and non-personal data, are included.

Australia is considering the establishment of a framework to identify “National Interest Datasets” or “designated datasets” (Australian Government, n.d.^[61]; 2018^[62]). These datasets would primarily include public sector data, but may also include private-sector data controlled by the public sector under certain conditions.

In France, the Law for a Digital Republic (*Loi pour une République numérique*) defines “data of general interest” as including: i) private-sector data from delegated public services such as utility or transportation services; ii) private-sector data that are essential for granting subsidies' and iii) private-sector data needed for national statistics (Government of France, 2016^[63]).

Under the concept of “private-sector data for public interest purposes”, the European Commission is examining data sharing between the private and public sector to guide policy making and the improvement of public services (European Commission, 2018^[64]).

Data of public interest are typically intended to be used mainly by governments or public sector institutions. However, in some cases, access to data is regulated based on competition and (system) efficiency considerations. This is particularly the case in network industries such as telecommunication, energy and transport. Finland's 2018 Act on Transport Services is a three stage legislative project to streamline all transport market regulations into one package. The act introduces significant changes to transport markets that have so far been strictly regulated and steered by public measures. It promotes customer-oriented, market-based transport services on the basis of sound competition. The act's goals are twofold. First, through deregulation, it gives more room to develop innovative, digitally enabled services. Second, it obliges all service providers to open certain essential data to all and to open ticketing and payments APIs (application programming interfaces) for single trip/ticket to third parties. The act makes it possible to examine transport as a whole, i.e. as one service.

The Finnish Act on Transport Services assumes that future transport will rely on open access to necessary data, the interoperability of information and information systems through APIs and the openness of these interfaces. By the end of 2018, around 5 200 companies in the Finnish transportation sector had made their data available, mostly via APIs, since the adoption of the act. Estimates suggest that this amount covers around 80% of transportation services used in Finland. These include taxi services (with more than 1 400 datasets), on-demand transportation services (around 400 datasets), timetable-bound public transportation services (around 240 datasets), rental services and commercial car-sharing services (around 20 datasets) and commercial parking services. In addition, the most important actors have opened their ticketing and payment system APIs, particularly those within the largest cities.²³

Data portability

Data portability with a focus on (consumer) personal data is another means to promote access and sharing in the private sector. These different types of initiatives are discussed in the following sections.

Data portability is often regarded as a promising means for promoting cross-sectoral re-use of data. At the same, it may strengthen control rights of individuals over their personal data and of businesses (in particular SMEs) over their business data (Productivity Commission, 2017^[21]). Prominent data portability initiatives include the Green Button in the United States (US Department of Energy, n.d.^[22]), which is part of the country's "My Data" initiative (launched in 2010). In 2011, the United Kingdom began its Midata data portability initiative (BIS, 2011^[23]). The European Union approved the Right to Data Portability (Art. 20 of the GDPR) (European Union, 2016^[19]). Most recently, Australia proposed its CDR.

The entry into force of the GDPR in May 2018 formalised the right of data portability within the European Union. Whereas the directive that preceded the GDPR gave data subjects the right to access their data,²⁴ the GDPR granted them a separate, distinct right of personal data portability. That right, in Article 20 of the GDPR, provides that the data subject "shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance...". "Data subject" only includes natural persons – corporations cannot take advantage of the right to data portability (Art. 4[1]). The right applies where processing is based on consent or another legitimate category (Art 9[2]) is necessary for the performance of a contract, or when the processing is carried out by automated means (Art. 20[1]). Recital 68 explains that data controllers "be encouraged to develop interoperable formats that enable data portability" but that the right does not oblige controllers to adopt or maintain processing systems that are technically compatible.

In August 2019, Australia introduced a CDR. This legislation enabled consumers in designated sectors of the Australian economy (a "CDR consumer") to have certain information disclosed to them or to accredited persons.²⁵ The right applies in respect of "CDR data". This is intended either to include information relating to the CDR consumer or information that is about goods or services in a particular sector that does not relate to any identifiable consumer.

The legislation defines three categories of actors. The first category is data holders, who are the original holders of CDR data. The second category is CDR consumers, who can be either individuals or small businesses²⁶ that hold rights to access data held by data holders and direct that data be shared with an accredited person. The third category is accredited data recipients are individuals or businesses that meet a series of criteria for accreditation to be further specified in the consumer data rules. In particular, accredited data recipients must comply with safeguards to protect the privacy or confidentiality of CDR data (Division 5). Those safeguards include requirements that accredited persons do not solicit CDR data or use it for direct marketing, that they manage the data openly and transparently, and that they comply with notification processes. The right, which will initially apply in the banking sector, will progressively extend to other sectors such as energy and telecommunications (Parliament of Australia, 2019^[78]).

Singapore's Personal Data Protection Commission introduced data portability obligation in legislation in 2019. Under the proposed obligation, an organisation must, at the request of the individual, transmit their data in the organisation's possession or under its control to another organisation in a commonly machine-readable format. This would enable greater data flows in the digital economy and encourage business innovation to bring to market innovative products/services. Singapore planned to clarify data portability for white-listed datasets, and to pilot, test and fine-tune the mechanisms and processes with industry to make data porting easy, safe and consistent for consumers.

Increasing data analytic capacities across society is addressed by a minority of governments

Increasing data analytic capacities, either in the public or private sector, was not considered a priority by countries responding to the 2018 EASD Policy Questionnaire. Only 12% of all policy initiatives cited by respondents to the questionnaire addressed data analytic capacities. A quarter of those initiatives focused on establishment of technology centres that support and/or guide in the re-use and analysis of data for public and/or private-sector entities. Some have also supported investments in data-related innovation and R&D.

Supporting the development of data-related skills and infrastructures

Governments have recognised that availability of data-related skills and competences can be a critical bottleneck for the effective re-use and provision of data in both the private and public sectors. Some have established dedicated initiatives to support development of data-related skills and infrastructures.

- The United Kingdom supports skill development in the private and public sectors in several ways. The Digital Skills Partnership, for instance, brings together public, private and charity sector organisations to boost skills for a world-leading, inclusive digital economy. The United Kingdom also has initiatives related to data ethics and AI such as the Data Ethics Framework and the Centre for Data Ethics and Innovation. In addition, it established a Data Skills Taskforce with the help of the Department of Digital, Culture, Media and Sport, Tech Partnership and Accenture to enhance data analytic skills in the workforce.
- Estonia's Digital Solutions seminars target industrial companies keen to improve production efficiency via digital solutions, including use of data. They aim to enhance knowledge and skills on the collection and use of data and information. The initiative was funded with EUR 200 000 between 2017 and 2020.
- The Ministry of Education of the People's Republic of China (hereafter "China") has supported the development of data-related skills through data analytics competitions with the Internet firm, Alibaba. This competition, held every year since 2010, helps partners identify the most talented data scientists in China.

A significant share of initiatives addresses public servants. Slovenia's education and training programmes, for example, increase data-related skills and competencies among public servants. The Ministry of Public Administration has funded these programmes since 2016.

Establishing and collaborating with data (analytics) support centres

Some governments have established data analytic and innovation centres to support their agencies in the sharing and re-use of data. Others have created and strengthened partnerships with such centres.

In 2013, Ireland's Department of Jobs, Enterprise and Innovation, through the state agency Science Foundation Ireland (SFI), established Insight – the SFI Research Centre for Data Analytics. This centre is considered one of Europe's largest data analytics research organisations and involves significant co-funding from and collaboration with industry partners. Insight undertakes high-impact research and seeks to derive value from big data. By enabling better decision making, it provides innovative technology solutions for industry and society. In addition to more than EUR 120 million from SFI, the centre also received cash and in-kind commitments of more than EUR 24 million from close to 90 companies.

Australia's data innovation centre, Data61, which is part of Australia's Commonwealth Scientific and Industrial Research Organisation, has partnered with government agencies. Together, they build new technologies that make high-value government data available to more people, while preserving privacy. In close collaboration with partner agencies, Data61 has developed a suite of new tools and technologies to enhance open data access, data sharing between agencies and managing privacy risks with sensitive data. The Confidential Computing Platform uses distributed machine learning – as well as homomorphic encryption and secure multi-party computing – to provide insights without organisations disclosing any data. This keeps the source data secure, private and up to date (Data61, n.d.^[65]).

The European Commission is working towards a support centre for data sharing under the Connecting Europe Facility Programme. This centre was expected to facilitate sharing of both private and public sector data. "It will offer know-how and assistance on data sharing by providing best practice examples and information on APIs, existing model contracts and other legal and technical aspects" (European

Commission, 2017^[66]). This would include further improving the Guidance on Private Sector Data Sharing (European Commission, 2018^[67]) discussed above.

Supporting innovation and R&D in data analytics and related technologies

A number of countries support innovation and R&D in data analytics and related technologies. Many of these policies are part of broader initiatives to support the digital economy or innovation. Few initiatives are solely dedicated to data analytics and data sharing.

The European Commission has put in place three funding mechanisms for data-related innovation:

- Funding for data innovation incubators connect data providers to data users. Three consortia composed of businesses and research organisations have been funded for three years with EUR 15 million.
- Funding of pan-European aggregators of public sector information (European Data Portal) aims to develop common metadata catalogues of all public sector information published in EU member states, searchable in multiple languages. This initiative, funded with EUR 10 million since 2015, continues until 2020.
- Privacy-enhancing technologies, including five consortia composed of businesses and research organisations, received EUR 65.5 million over three years.

Governments are turning to data-related initiatives to improve the effectiveness of their frontline responses to COVID-19

Governments are turning to a wide array of digital technologies and advanced analytics to collect, analyse and share data for frontline response to the COVID-19 crisis (Dunant, 2020^[16]). Above all, most countries are leveraging the widespread use of mobile phones.

Telecommunication service providers serve substantial portions of the population across entire nations. Thanks to mobile call data records, the movements of millions of people at fine spatial and temporal scales can be measured in near real time. This can provide useful information on trends and fluctuations over time, helping reduce uncertainties attached to outbreak detection and response.

In several OECD countries, telecommunication service providers share geolocation data based on mobile call data records with health officials in an aggregated, anonymised format. For example, the main German telecommunications provider, Deutsche Telekom, is providing anonymised “movement flows” data of its users to the Robert Koch Institute, a government research agency responsible for disease control and prevention (Politik, 2020^[68]). Vodafone Group’s Five Point Plan to address COVID-19 includes providing governments with large anonymised data sets. For example, it has an aggregated and anonymous heat map for the Lombardy region in Italy. These data sets will help authorities better understand population movements (Vodafone, 2020^[69]).

Governments use information from these data sets to track the COVID-19 outbreak, warn vulnerable communities and understand the impact of policies such as social distancing and confinement. The European Commission, for instance, has been liaising with eight European telecommunications operators to obtain anonymised aggregate mobile location data to co-ordinate monitoring of the spread of COVID-19 (European Commission, 2020^[70]). To address privacy concerns, the data were to be deleted at the end of the crisis (Chee, 2020^[71]).

Governments are also fostering the development and use of smart applications to respond to COVID-19, including specific mobile apps for tracking and tracing infections. Some of these applications rely on GPS-based geolocation data or Bluetooth-based proximity data. In Korea, for instance, the government funded a GPS-based Self-quarantine Safety App, which is used by public authorities to effectively support the monitoring of those under self-quarantine. This app has three key features: (i) a self-diagnosis feature for users to conduct a self-assessment of their possible COVID-19 infection and to share the results with their assigned local government officer; (ii) a GPS-based geolocation tracking feature to prevent possible violation of self-quarantine orders; and (iii) an information feature to provide necessary information including self-quarantine guidelines and the contact information of the assigned local government officer. The data collected by the Self-quarantine Safety App is not shared with third parties. In addition, Korea has also deployed an Epidemiological Investigation Support System (EISS) to

trace contacts and movements of confirmed COVID-19 patients. This system, which operates in a strict manner to protect the privacy of individuals, can help public health officials locate possible sources of COVID-19 infections, identify hot beds of infections and warn citizens.

Epidemiologists confirmed these type of applications are crucial in providing detailed information about the movements of infected people, their possible infected contacts and can thus help track and control the pandemic. However, it remains controversial whether making data on e.g. hot beds of infections available to the public is the right choice (Everett, Hudson and Collins, 18 March 2020^[72]). When an individual tests positive for COVID-19, for instance, their city or district might alert people who live nearby about the movements of potentially infected individuals prior to their diagnosis.²⁷ While the World Health Organization praised Korea's extensive tracing measures, some uses by designated local authorities of the data collected through the EISS on the movements of persons with confirmed cases have raised privacy concerns (Zastrow, 2020^[73]; Nemo, 2020^[74]). In response, the Korean government recently published guidance related to the disclosure of the movements of persons with confirmed cases based on the Infectious Disease Control and Prevention Act passed in 2015 which does not allow any information specific to the data subject to be disclosed.

Proximity data collected from contact tracing apps can provide even more granular enriched data on individuals with a potential COVID-19 infection. Singapore, for example, initiated contact tracing for all confirmed cases from the early days of the outbreak. Specifically, it traced contacts of confirmed cases during their infectious period and followed up accordingly. To support the work of contact tracers, Singapore has rolled out digital track and tracing tools and solutions developed by the government. This includes the TraceTogether app, a smartphone app that uses short-distance Bluetooth signals between phones to detect other TraceTogether users in close proximity. To enhance the effectiveness of this initiative, and in particular to include the digitally excluded population, it also introduced a dedicated TraceTogether portable device – TraceTogether Token.²⁸

The developers of TraceTogether (app and token) have put in place a number of privacy safeguards. For instance, TraceTogether does not collect or use location data. As well, data logs are stored in an encrypted form on the device. Authorities can only access the Bluetooth proximity data if a user tests positive for COVID-19. TraceTogether works in tandem with the SafeEntry national digital check-in system that logs the entry and exit of individuals entering and exiting public venues. This reduces the time required for contact tracing to perform activity mapping. This, in turn, allows authorities to alert close contacts of those infected with COVID-19 more quickly. As of June 2020, 35% of the population (more than 2.1 million) in Singapore have downloaded the app according to information provided by the Government of Singapore.

References

- AG Data Transparent (2016), “AG Data’s core principles: The privacy and security principles for farm data”, webpage, <http://www.agdatatransparent.com/principles/> (accessed on 21 October 2020). [59]
- Australian Government (2018), “New Australian government data sharing and release legislation”, *Issues Paper for Consultation*, Department of the Prime Minister and Cabinet, Canberra, <http://www.pmc.gov.au/resource-centre/public-data/issues-paper-data-sharing-release-legislation>. [62]
- Australian Government (2017), *Information about the Data Integration Partnership for Australia*, Brochure, Department of the Prime Minister and Cabinet, Data and Digital Branch, Canberra, <http://www.pmc.gov.au/sites/default/files/publications/DIPA-information.pdf>. [60]
- Australian Government (n.d.), “Designated Datasets — a special class of high-value dataset: Australian government’s response to Productivity Commission Recommendations: 7.1 and 7.2”, webpage, <https://dataavailability.pmc.gov.au/designated-datasets.html> (accessed on 21 October 2020). [61]
- Bajari, P. et al. (2019), “The impact of big data on firm performance: An empirical investigation”, *AEA Papers and Proceedings*, Vol. 109, pp. 33-37, <http://dx.doi.org/10.1257/pandp.20191000>. [6]
- Bakhshi, H., A. Bravo-Biosca and J. Mateos-Garcia (2014), “The analytical firm: Estimating the effect of data and online analytics on firm performance”, *Working Paper*, No. 14/05, Nesta, London, https://media.nesta.org.uk/documents/1405_the_analytical_firm_-_final.pdf. [9]
- Banham, R. (2014), “Who owns farmers’ big data?”, *Forbes*, 8 July, <https://www.forbes.com/sites/emc/2014/07/08/who-owns-farmers-big-data/>. [31]
- Bauer, M., M. Ferracane and E. van der Marel (2016), *Tracing the Economic Impact of Regulations on the Free Flow of Data and Data Localization*, Centre for International Governance Innovation (CIGI), Waterloo, Canada, <https://www.cigionline.org/publications/tracing-economic-impact-regulations-free-flow-data-and-data-localization> (accessed on 21 October 2020). [54]
- Belissent, J. (2017), “Insights services drive data commercialization”, *Featured Insights blog*, 8 March, https://go.forrester.com/blogs/17-03-08-insights_services_drive_data_commercialization/. [20]
- BIS (2013), *Shakespeare Review: An Independent Review of Public Sector Information*, UK Department for Business Innovation & Skills, London, http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198752/13-744-shakespeare-review-of-public-sector-information.pdf. [38]
- BIS (2011), *Better Choices: Better Deals – Consumers Powering Growth*, UK Department for Business Innovation & Skills, London, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/294798/bis-11-749-better-choices-better-deals-consumers-powering-growth.pdf. [23]
- Brynjolfsson, E. and K. McElheran (2019), “Data in action: Data-driven decision making and predictive analytics in U.S. manufacturing”, *Working Paper*, No. 3422397, Rotman School of Management, Toronto, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3422397. [4]
- Brynjolfsson, E. and K. McElheran (2016), “The rapid adoption of data-driven decision-making”, *American Economic Review*, Vol. 106/5, pp. 133-39, <http://dx.doi.org/10.1257/aer.p20161016>. [8]
- Carrière-Swallow, Y. and F. Labbé (2013), “Nowcasting with Google trends in an emerging market”, *Journal of Forecasting*, Vol. 32/4, pp. 289-298. [35]
- Casalini, F. and J. López González (2019), “Trade and Cross-Border Data Flows”, *OECD Trade Policy Papers*, No. 220, OECD Publishing, Paris, <https://dx.doi.org/10.1787/b2023a47-en>. [44]
- Chee, F. (2020), “Vodafone, Deutsche Telekom, 6 other telcos to help EU track virus”, *Reuters Technology News*, 25 March, <https://www.reuters.com/article/us-health-coronavirus-telecoms-eu/vodafone-deutsche-telekom-6-other-telcos-to-help-eu-track-virus-idUSKBN21C36G>. [71]
- Choi, H. and H. Varian (2009), “Predicting the present with Google trends”, *SSRN*, <http://dx.doi.org/10.2139/ssrn.1659302>. [33]
- Chui, M., J. Manyika and S. Kuiken (2014), “What executives should know about open data”, *Our Insights*, McKinsey & Company, New York, 1 January, <http://www.mckinsey.com/industries/high-tech/our-insights/what-executives-should-know-about-open-data> (accessed on 21 October 2020). [30]
- Cisco (2018), *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2017–2022 White Paper* - Cisco, Cisco Systems, San Jose, California, <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-738429.html> (accessed on 21 October 2020). [51]

- Data61 (n.d.), “Confidential Computing – Insights from Data Without Seeing the Data”, webpage, <https://data61.csiro.au/en/Our-Research/Focus-Areas/Privacy-Preserving-Technologies> (accessed on 21 October 2020). [65]
- DCMS (2018), “Guidance Data Ethics Framework”, webpage, <http://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework> (accessed on 21 October 2020). [58]
- Deloitte (2017), “Assessing the value of TfL’s open data and digital partnerships”, report commissioned for Transport for London, <http://content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf>. [41]
- Deloitte (2013), “Market assessment of public sector information”, report commissioned by the UK Department for Business, Innovation & Skills, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf. [37]
- Dunant, R. (2020), “Open letter: Contact tracking and NHSX”, Medium, 23 March, <https://medium.com/@rachelcoldicutt/open-letter-contract-tracking-and-nhsx-e503325b2703?sk=4e6097dea429498f20d5e33b0cfc2436>. [16]
- European Commission (2020), Commission Recommendation (EU) 2020/518 of 8 April 2020 on a common Union toolbox for the use of technology and data to combat and exit from the COVID-19 crisis, in particular concerning mobile applications and the use of anonymised mobility data, European Commission, Brussels, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32020H0518>. [70]
- European Commission (2018), “Guidance on sharing private sector data in the European data economy”, Accompanying the document Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, “Towards a common European data space,” COM(2018), 125, Final, European Commission, Brussels. [67]
- European Commission (2018), “Towards a common European data space”, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, COM(2018), 232, Final, European Commission, Brussels, <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2018:0232:FIN>. [64]
- European Commission (2017), Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions “Building a European Data Economy”, European Commission, Brussels, <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2017%3A9%3AFIN>. [66]
- European Union (2016), Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, European Union, Brussels, <http://data.europa.eu/eli/reg/2016/679/oj>. [19]
- Everett, M., L. Hudson and K. Collins (18 March 2020), “COVID-19: When public health and privacy collide?”, Data Notes blog, <https://hsfnotes.com/data/2020/03/18/covid-19-when-public-health-and-privacy-collide/>. [72]
- Government of France (2016), Loi pour une République numérique, Paris, <http://www.senat.fr/leg/pjl15-744.html>. [63]
- Harris, D. (2011), “Hadoop kills zombies too! Is there anything it can’t solve?”, Gigaom blog, 18 April, <http://gigaom.com/cloud/hadoop-kills-zombies-too-is-there-anything-it-cant-solve/>. [34]
- iClarified (2012), “Goldman Sachs values iPhone/iPad customer base at \$295 billion”, iClarified, 29 June, <https://www.icularified.com/22914/goldman-sachs-values-iphoneipad-customer-base-at-295-billion>. [25]
- IDC and Lisbon Council (2018), “Updating the European data market monitoring tool”, a report commissioned for the European Commission, Brussels, <http://datalandscape.eu/study-reports/first-report-facts-and-figures-updating-european-data-market-monitoring-tool>. [42]
- Information System Authority [Estonia] (2019), “Data Exchange Layer X-tee”, webpage, <http://www.ria.ee/en/state-information-system/x-tee.html> (accessed on 21 October 2020). [57]
- ITU (2020), “Mobile cellular subscriptions”, World Telecommunication/ICT Development Report, (database), International Telecommunication Union, Geneva, <https://data.worldbank.org/indicator/IT.CEL.SETS> (accessed on 21 October 2020). [17]
- King’s College London et al. (2020), C-19 Covid Symptom Tracker, website, <https://covid.joinzoe.com/> (accessed on 21 October 2020). [18]
- Konsynski, B. and F. McFarlan (1990), “Information partnerships – shared data, shared scale”, Harvard Business Review, September-October, <https://hbr.org/1990/09/information-partnerships-shared-data-shared-scale>. [29]
- Lateral Economics (2014), “Open for business: How open data can help achieve the G20 growth target”, report commissioned by Omidyar Network, Redwood City, California. [40]
- Li, W., M. Nirei and K. Yamana (2019), “Value of data: There’s no such thing as a free lunch in the digital economy”, Working Paper, US Bureau of Economic Analysis, Washington, DC. [12]
- Mandel, M. (2012), Beyond Goods and Services: The (Unmeasured) Rise of the Data-Driven Economy | Progressive Policy Institute, Progressive Policy Institute, <https://www.progressivepolicy.org/publication/beyond-goods-and-services-the-unmeasured-rise-of-the-data-driven-economy/> (accessed on 21 October 2020). [77]
- McKinsey & Company (2017), Fueling Growth through Data Monetization, McKinsey & Company, New York, <http://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/fueling-growth-through-data-monetization>. [2]

- McKinsey & Company (2013), *Open Data: Unlocking Innovation and Performance with Liquid Information*, McKinsey & Company, New York, http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information. [39]
- MGI (2016), *Digital Globalisation: The New Era of Global Flows*, McKinsey Global Institute, New York, <https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Digital%20globalization%20The%20new%20era%20of%20global%20flows/MGI-Digital-globalization-Full-report.ashx>. [46]
- Nemo, K. (2020), “More scary than coronavirus’: South Korea’s health alerts expose private lives”, *The Guardian*, 6 March, <https://www.theguardian.com/world/2020/mar/06/more-scary-than-coronavirus-south-koreas-health-alerts-expose-private-lives>. [74]
- Niebel, T., F. Rasel and S. Viete (2018), “BIG data – BIG gains? Understanding the link between big data analytics and innovation”, *Economics of Innovation and New Technology*, July, <http://dx.doi.org/10.1080/10438599.2018.1493075>. [5]
- ODI (2016), *Open Enterprise: How Three Big Businesses Create Value with Open Innovation*, <https://theodi.org/article/open-enterprise-how-three-big-businesses-create-value-with-open-innovation/> (accessed on 21 October 2020). [28]
- OECD (2020), “OECD Competition Assessment Toolkit”, webpage, <https://www.oecd.org/competition/assessment-toolkit.htm> (accessed on 21 October 2020). [15]
- OECD (2019), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris, <https://doi.org/10.1787/276aaca8-en>. [14]
- OECD (2019), *Measuring the Digital Transformation: A Roadmap for the Future*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264311992-en>. [50]
- OECD (2019), “Trade and Cross-Border Data Flows”, *OECD Trade Policy Papers*, No. 220, OECD Publishing, Paris, <https://doi.org/10.1787/b2023a47-en>. [49]
- OECD (2018), “Digital Trade and Market Openness”, *OECD Trade Policy Papers*, No. 217, OECD Publishing, Paris, <https://doi.org/10.1787/1bd89c9a-en>. [48]
- OECD (2017), *OECD Digital Economy Outlook 2017*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264276284-en>. [47]
- OECD (2017), *The Next Production Revolution: Implications for Governments and Business*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264271036-en>. [3]
- OECD (2015), *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264229358-en>. [1]
- OECD (2013), “The App Economy”, *OECD Digital Economy Papers*, No. 230, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5k3ttftlv95k-en>. [24]
- OECD (2008), *Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information*, OECD, Paris, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0362>. [56]
- OECD (2006), *Recommendation of the Council concerning Access to Research Data from Public Funding*, OECD, Paris, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347>. [27]
- OFT (2006), *The Commercial Use of Public Information*, Office of Fair Trading, London, <https://ec.europa.eu/digital-single-market/en/news/commercial-use-public-information-oft-study>. [75]
- Parliament of Australia (2019), *Cth. Parliamentary Debates*, House of Representatives, Canberra, 30 July, pp. 1379, http://www.aph.gov.au/Parliamentary_Business/Hansard/. [78]
- Pepper, R. and J. Garrity (2014), “The Internet of everything: how the network unleashes the benefits of big data”, *The Global Information Technology Report*, pp. 35-42, <https://alln-extcloud-storage.cisco.com/ciscoblogs/GITR-2014-Cisco-Chapter.pdf> (accessed 21 October 2020). [45]
- Politik (2020), *Telekom teilt Daten über „Bewegungsströme“ von Handynutzern mit RKI* [Telekom shares Data about Movement Flows of Mobile Phone Users with the Robert Koch Institute], *Welt*, 8 March, <https://www.welt.de/politik/deutschland/article206624141/Coronavirus-Telekom-teilt-Bewegungstroeme-von-Handynutzern-mit-RKI.html>. [68]
- Productivity Commission (2017), *Productivity Commission Inquiry Report: Data Availability and Use*, Productivity Commission, Government of Australia, Melbourne, <https://www.pc.gov.au/inquiries/completed/data-access/report/data-access.pdf>. [21]
- Reimsbach-Kounatze, C. (2015), “The Proliferation of “Big Data” and Implications for Official Statistics and Statistical Agencies: A Preliminary Analysis”, *OECD Digital Economy Papers*, No. 245, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5js7t9wqzvg8-en>. [32]
- Schimmelpfennig, D. and R. Ebel (2016), “Sequential adoption and cost savings from precision agriculture”, *Journal of Agricultural and Resource Economics*, Vol. 41/1, pp. 97-115, http://dx.doi.org/www.waeaonline.org/UserFiles/file/JARE_January2016Schimmelpfennigpp97-115.pdf. [10]
- Shapiro, R. and S. Aneja (2019), *Who Owns Americans’ Personal Information and What Is It Worth?*, *Future Majority*, <https://www.futuremajority.org/pages/who-owns-americans-personal-information> (accessed on 21 October 2020). [76]

- TeleGeography (n.d.), *Submarine Cable Map*, website, <http://www.submarinecablemap.com> (accessed on 3 August 2020). [55]
- Ubaldi, B. (2013), “Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives”, *OECD Working Papers on Public Governance*, No. 22, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5k46bj4f03s7-en>. [26]
- United Nations Global Pulse (2012), *Big Data for Development: Opportunities & Challenges*, United Nations Global Pulse, New York, <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf>. [36]
- United States Department of Commerce (2016), *Measuring the Value of Cross-Border Data Flows | U.S. Department of Commerce*, <https://www.commerce.gov/news/reports/2016/09/measuring-value-cross-border-data-flows> (accessed on 15 September 2020). [43]
- United States International Trade Commission (2014), *Digital Trade in the U.S. and Global Economies, Part 2*, <https://www.usitc.gov/publications/332/pub4485.pdf> (accessed on 21 October 2020). [52]
- United States Department of Energy (n.d.), “Green button: Open energy data”, webpage, <https://www.energy.gov/data/green-button> (accessed on 21 October 2020). [22]
- van der Marel, E., H. Lee-Makiyama and M. Bauer (2011), *The Costs of Data Localisation: A Friendly Fire on Economic Recovery*, European Centre for International Political Economy (ECIPE), <https://ecipe.org/publications/dataloc/> (accessed on 15 September 2020). [53]
- Vodafone (2020), “An industrial 5G spectrum policy for Europe”, *Public Policy Paper*, Vodafone, Berkshire, United Kingdom, <https://www.vodafone.com/content/dam/vodcom/files/public-policy/5g-report/an-industrial-5g-spectrum-policy-for-europe.pdf>. [69]
- Wamba, S. et al. (2017), “Big data analytics and firm performance: Effects of dynamic capabilities”, *Journal of Business Research*, Vol. 70, pp. 356-365, <http://dx.doi.org/10.1016/j.jbusres.2016.08.009>. [7]
- Waters, R. (2015), “IBM’s latest deal is a new test case for the big data economy”, *Financial Times*, 29 October, <http://www.ft.com/content/0fe3ac2e-7e22-11e5-a1fe-567b37f80b64>. [13]
- Zastrow, M. (2020), “South Korea is reporting intimate details of COVID-19 cases: Has it helped?”, *Nature*, 18 March, <https://www.nature.com/articles/d41586-020-00740-y>. [73]

Notes

- High performers are defined as companies that had annual growth rates of 10% or more over the past three years.
- Shapiro and Aneja (2019^[76]) provide estimates of the value of American personal data based on the digital advertising revenue of the major online platforms. In 2018, based on these companies’ financial statements, the platforms earned USD 111.1 billion from US advertisers targeting American consumers. Moreover, the authors note that, “Google and Facebook dominated this area in 2018, accounting respectively for 37.1 percent (\$41.3 billion) and 20.6 percent (\$22.9 billion) of total digital advertising revenues” (p. 9).
- According to TJI Research, Amazon sells products using 139 private label brands (update April 2019), across different product categories, including clothing, electronics, food, furniture, household goods and health care.
- As Mandel (2012^[77]) highlights: “[...] economic and regulatory policymakers around the world are not getting the data they need to understand the importance of data for the economy. Consider this: The Bureau of Economic Analysis [...] will tell you how much Americans increased their consumption of jewellery and watches in 2011, but offers no information about the growing use of mobile apps or online tax preparation programs. Eurostat [...] reports how much European businesses invested in buildings and equipment in 2010, but not how much those same businesses spent on consumer or business databases. And the World Trade Organization publishes figures on the flow of clothing from Asia to the United States, but no official agency tracks the very valuable flow of data back and forth across the Pacific.”
- While data flow plays a vital role in the digital economy, legal and regulatory frameworks enable such cross-border data flows with trust. This is especially true of privacy and data protection regulation as discussed in Chapter 6.
- In this context, two ideas are debated: i) “data commons”, where some data are shared publicly after adequate anonymisation and aggregation; and ii) “digital smoke signals”, where companies analyse sensitive data but share results with governments.

7. Studies differ significantly in terms of the scope of the sectors (e.g. public sector and/or private sector), types of data (e.g. personal, proprietary or public), and degrees of data openness (and arrangements included, such as open data), as well as the methodologies, including the different level of impact assessed (i.e. organisational, sectoral or macroeconomic).
8. The study was based on the methodology in OfT (2006^[75]) but with a more expanded scope. It focuses particularly on trading funds such as the HM Land Registry, the Registers of Scotland, the Companies House, the Ordnance Survey, the UK Hydrographic Office, the Environment Agency, the Met Office, and the Office of National Statistics.
9. These are based on 2011 data and include around GBP 100 million in revenues generated from sales of public sector information (PSI); GBP 100 million through supply-chain effects from increased jobs and related consumer spending from the production of PSI; and GBP 1.6 billion through consumer surplus from direct use and consumption of PSI-related products.
10. Altogether, it is estimated that consumer and customer surplus generate over half of the total potential value of open data (McKinsey & Company, 2013^[39]). The largest share of the total benefits of open data is attributed to better benchmarking, “an exercise that exposes variability and also promotes transparency within organisations” (McKinsey & Company, 2013^[39]). Better benchmarking would enable “fostering competitiveness by making more information available and creating opportunities to better match supply and demand” as well as “enhancing the accountability of institutions such as governments and businesses [to] raise the quality of decision [making] by giving citizens and consumers more tools to scrutinise business and government” (McKinsey & Company, 2013^[39]).
11. The data market is defined as the marketplace where digital data is exchanged as “products” and “services” as a result of the (re-)processing of raw data. The impact on the data economy is defined more broadly as the overall effects of the data market on the economy, involving generation, collection, storage, processing, distribution, analysis elaboration, delivery and exploitation of data enabled by digital technologies. Therefore, the overall impact is estimated by summing up the direct, indirect and induced impact. For the estimation of data market and the data economy, IDC and the Lisbon Council (2018^[42]) identify data companies that have both data suppliers and data users. Data suppliers have, as their main activity, the production and delivery of digital data-related products, services and technologies, while data users are organisations that generate, exploit, collect and analyse digital data intensively to improve their business activities.
12. See www.volkswagen.com/en/news/stories/2019/03/volkswagen-industrial-cloud.html.
13. That is why impediments to international data transfers can have severe negative economic impacts on businesses and ultimately on complex value chains and trade.
14. The MGI model estimates the contribution of various flows – including data – to approximate their impact on real GDP. Data flows are approximated by cross-border used bandwidth from TeleGeography (sum of capacity for Internet backbones, private networks and switched voice networks). They ran the model for 97 countries for 1995-2013 and found that a 10% increase in cross-border data flows raises GDP by 0.2.
15. The use of data volume is also further complicated by the use of data compression techniques widely applied to flows of data.
16. This section assesses policy trends related to enhancing data access and sharing based on two country surveys, the most recent of which, the EASD Policy Questionnaire, was conducted between June and September 2018 and covered 20 countries plus the European Union. This survey was complemented by the responses to the Digital Economy Policy Questionnaire, which included additional 16 countries, many of which are partner economies. As a result, it analysed 205 policy initiatives across 37 countries.
17. The OECD *Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information* (OECD PSI Recommendation) (OECD, 2008^[56]) defines public sector (government) data as a subset of PSI, which includes not only data but also *digital content*, such as text documents and multimedia files. The terms “public sector data” and “government data” are used as synonyms. The oft-used term “open government data” refers to public sector data made available as *open data*. These data are i) dynamic and continuously generated; ii) often directly produced by the public sector; or iii) associated with the functioning of the public sector (e.g. meteorological data, geo-spatial data, business statistics); and iv) often readily useable in commercial applications with relatively little transformation, as well as being the basis of extensive elaboration.
18. PSI typically includes not only data but also digital content, such as text documents and multimedia files.
19. Government agencies are appointed to acquire, maintain, fuse and distribute quality data securely; centralised infrastructure, with in-built safeguards, is put in place to enable data discoverability, secure access to data and data analytics. The GDA will enable government agencies to access commonly used data within seven working days, and obtain insights more expeditiously and readily. In addition, Singapore introduced the Public Sector

(Governance) Act, which came into effect in April 2018, to formalise the data sharing framework between public sector agencies. The act makes clear that public sector agencies may share data with each other for seven specific purposes. Among these purposes are improving the efficiency or effectiveness of policy planning and service delivery, but this would not overcome confidentiality obligations set out in legislation or contracts. The act also includes safeguards for data protection, including setting out criminal penalties for those who make use of data to benefit themselves, re-identify anonymised data without authorisation and public sector officers who disclose the personal data of Singaporeans without authorisation.

20. The framework includes a Data Ethics Workbook with questions to probe ethical, information assurance and methodological considerations when building or buying new technology.
21. These include a database, geodatabase, shape-file, coverage, raster image or dbf table.
22. The AI section proposes to conclude contracts along with “Exploratory Multi-phased” AI development processes, which consists of assessment, proof of concept, development and retraining.
23. To support the interoperability of ticketing and payment system APIs, the Lippu Network was established.
24. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (OJ L 281, 23.11.1995), Art. 12.
25. Treasury Laws Amendment (Consumer Data Right) Act 2019 (Cth).
26. “Small business” is defined in the Privacy Act 1988 (Cth) as follows: “[a] business is a small business ... in a financial year ... if its annual turnover for the previous financial year is \$3,000,000 or less”, with some exceptions (see section 6D).
27. The alert can provide details about the infected person’s age, gender and a detailed log of their movements – even about the time and names of businesses visited.
28. Functioning in the same way as the TraceTogether app, the TraceTogether Token uses Bluetooth signals to record other nearby TraceTogether apps or tokens. By increasing the overall pool of participants, every user of the app or token would benefit by being informed as early as possible, if/when they have been exposed to COVID-19.



From:
OECD Digital Economy Outlook 2020

Access the complete publication at:
<https://doi.org/10.1787/bb167041-en>

Please cite this chapter as:

OECD (2020), "Enhancing data access, sharing and re-use", in *OECD Digital Economy Outlook 2020*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/3b6a594d-en>

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.