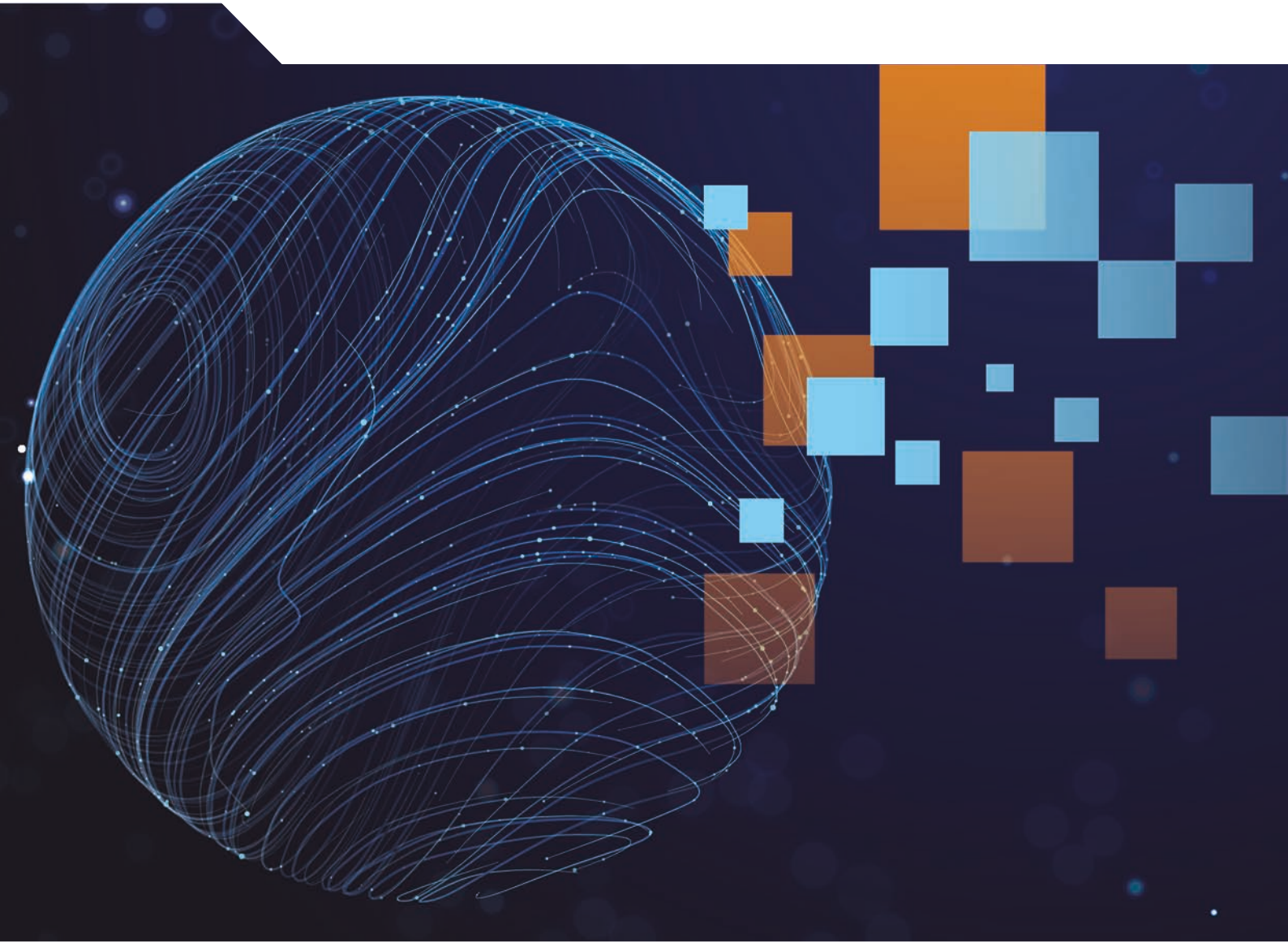
Going Digital to Advance Data Governance for Growth and Well-being



OECD

Going Digital to Advance Data Governance for Growth and Well-being

OECD
BETTER POLICIES FOR BETTER LIVES

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.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Note by the Republic of Türkiye
The information in this document with reference to "Cyprus" relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Türkiye recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Türkiye shall preserve its position concerning the "Cyprus issue".

Note by all the European Union Member States of the OECD and the European Union
The Republic of Cyprus is recognised by all members of the United Nations with the exception of Türkiye. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

Please cite this publication as:
OECD (2022), *Going Digital to Advance Data Governance for Growth and Well-being*, OECD Publishing, Paris, https://doi.org/10.1787/e3d783b0-en.

ISBN 978-92-64-85884-8 (print)
ISBN 978-92-64-64982-8 (pdf)
ISBN 978-92-64-89909-4 (HTML)
ISBN 978-92-64-91575-6 (epub)

Photo credits: Cover © RDVector/Shutterstock

Corrigenda to publications may be found on line at: www.oecd.org/about/publishing/corrigenda.htm.

© OECD 2022

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

Preface

Our increasingly digital world generates enormous volumes of data. Data are the by-product of every tap on a smartphone, scroll through a website or swipe of a credit card. Devices abound in everyday life that collect information about people and phenomena. It has never been easier or cheaper to process these data, store them for future use, or share them with others.

Through the Going Digital project, the OECD identified data as a key driver of economic and social value. We see this across our areas of work: data drive scientific research and fuel artificial intelligence. Firms invest in data, and measuring that investment is a focus of the statistical community. Data can confer competitive advantage and contribute to market power, while trade agreements now often feature provisions on data flows. Finally, as demonstrated by the COVID-19 crisis, data can help track the spread of disease and target health service delivery.

At the same time, data can be misused and abused in ways that can harm individuals and organisations. These harms can include the violation of privacy and personal data protection rights and intellectual property rights, digital security risks and confidentiality breaches. These possibilities and pitfalls illustrate that the way data are governed affects the ability of our societies and economies to develop, grow and respond to global challenges, from future pandemics to climate change.

The outcome of two years of fruitful collaboration across policy areas, *Going Digital to Advance Data Governance for Growth and Well-being* offers an entry point to understanding data and data governance in different policy contexts by providing a comprehensive overview of data, their use across sectors and applications, ensuing data governance challenges, and related policy solutions. It provides new evidence, analysis and insights to inform policies and puts forth concrete approaches to realise the benefits of data governance for growth and well-being.

The stakes for data governance are high – now is the time for robust and mature policies to encourage the wider, yet responsible, use of data.

Ulrik Vestergaard Knudsen

OECD Deputy Secretary-General

Carmine di Noia

Director of Financial and Enterprise Affairs

Andrew Wyckoff

Director of Science, Technology and Innovation

Marion Jansen

Director of Trade and Agriculture

Stefano Scarpetta

Director of Employment, Labour and Social Affairs

Paul Schreyer

Chief Statistician and Director of Statistics and Data

Foreword

This report *Going Digital to Advance Data Governance for Growth and Well-being* presents insights, data and policy messages from the third phase of the OECD Going Digital project, focusing on data governance for growth and well-being. The OECD Going Digital project is led by the OECD Directorate for Science, Technology and Innovation, under the oversight of Deputy Secretary-General Ulrik Vestergaard Knudsen, and the leadership and guidance of Andrew Wyckoff, Director, and Audrey Plonk, Head of the Digital Economy Policy Division. Angela Attrey, Gallia Daor and Christian Reimsbach-Kounatze served as the project co-ordinators for the third phase (2021-22) of the project.

This report was drafted by Angela Attrey with input from Gallia Daor and Christian Reimsbach-Kounatze. It draws on the work of the project's co-leads from five OECD Directorates, including the Directorate for Science, Technology and Innovation (Francesca Casalini, Audrey Plonk, Christian Reimsbach-Kounatze, Vincenzo Spiezia and Jeremy West), the Directorate for Employment, Labour and Social Affairs (Tiago Cravo Oliveira Hashiguchi and Jillian Oderkirk), the Directorate for Financial and Enterprise Affairs (Antonio Capobianco and James Mancini), the Trade and Agriculture Directorate (Javier López González) and the Statistics and Data Directorate (John Mitchell and Jorrit Zwijnenburg).

The report also benefits from the input and research of the project's focal points from across the OECD, including the Directorate for Science, Technology and Innovation (Brigitte Acoca, Luis Aranda, Sara Calligaris, Flavio Calvino, Alessandra Colecchia, David Gierten, Simon Lange, Molly Lesher, Nicholas McSpedden-Brown, Alan Paic and Jan Tscheke), the Centre for Entrepreneurship, SMEs, Regions and Cities (Rudiger Ahrend, Marco Bianchini, Sandrine Kergroach and Lora Pissareva), the Development Co-operation Directorate (Eleanor Carey and Ida McDonnell), the Directorate for Financial and Enterprise Affairs (Oliver Garrett-Jones and Iota Nassr), the Economics Department (Lilas Demmou), the Directorate for Education and Skills (Stéphan Vincent-Lacrin), the Public Governance Directorate (Miguel Amaral, Cecilia Emilsson, Marianna Karttunen, Jacob Arturo Rivera Perez and Barbara Ubaldi), and the International Transport Forum (Philippe Crist). Mark Foss, Sebastian Ordelheide and Angela Gosmann provided editorial support.

This phase of the OECD Going Digital project was led by the OECD Committee on Digital Economy Policy, joined by four co-leading OECD Committees: the Competition Committee, the Health Committee, the Committee on Statistics and Statistical Policy and the Trade Committee. In addition, the project featured the involvement of eleven other OECD bodies: the Committee on Consumer Policy, the Development Assistance Committee, the Economic Policy Committee, the Centre for Education, Research and Innovation Governing Board, the Committee on Financial Markets, the Committee on Industry, Innovation and Entrepreneurship, the Public Governance Committee, the Committee for Scientific and Technological Policy, the Committee on SMEs and Entrepreneurship, the Regional Development Policy Committee and the Regulatory Policy Committee. Input from the stakeholder groups of the Committee on Digital Economy Policy – Business at OECD, the Trade Union Advisory Committee, the Civil Society Information Society Advisory Council and the Internet Technical Advisory Committee – is gratefully acknowledged.

The OECD Committee on Digital Economy Policy approved and declassified *Going Digital to Advance Data Governance for Growth and Well-being* on 27 September 2022. The OECD Secretariat prepared the report for publication.

Table of contents

Preface 3

Foreword 5

Executive summary 8

1 Introduction 11
References 14

2 Data as a strategic asset: The fundamental shift in their use, misuse and appreciation 15
2.1. Data use is transformational for economies and societies 16
2.2. Data-driven business models, including platforms, are transforming markets 17
2.3. The shift in the nature and use of data introduces economic and social risks 18
2.4. The fundamental shift in data use is reflected in public policies at the national and international levels 21
References 23

3 Understanding data: Key characteristics, technological developments and policy challenges 27
3.1. Key characteristics of data 28
3.2. Understanding the value of data 29
3.3. Data and digital technologies are fundamentally linked 30
3.4. Coping with the new data-driven reality 30
3.5. Cross-cutting policy objectives and tensions and promising approaches 34
References 35
Endnote 37

4 Policy considerations 39
4.1. Data governance is a cross-cutting issue that requires a whole-of-government approach 40
4.2. Promoting cross-border data flows with trust requires an international approach 40
4.3. New technologies and governance arrangements can help align incentives and balance data-related benefits and risks 41
4.4. Unlocking the potential of data for digital transformation requires encouraging firms of all sizes to use data 41
4.5. Efforts to measure the value of data, including in the System of National Accounts, should be encouraged 42

4.6. The OECD is well placed to support countries in their data governance policy making and can facilitate international co-ordination 42
References 43

Figures

Figure 1.1. Cost of computer processing power and data storage, 1940-2020 12
Figure 2.1. Open access of COVID-19, diabetes and dementia publications, January-October 2020 16
Figure 2.2. Venture capital deals in big data firms worldwide, 2015-21 18
Figure 2.3. Personal data breaches are on the rise across a number of OECD jurisdictions 19
Figure 2.4. Growing number of regulations affecting cross-border data flows 22
Figure 3.1. Emerging technologies that pose the main challenges for privacy and personal data protection 31
Figure 3.2. OECD Open, Useful and Re-usable data (OURdata) Index: 2019 31
Figure 3.3. Real-time data available to governments as the COVID-19 pandemic hit 32

Boxes

Box 1.1. What is data? What is data governance? 13
Box 2.1. Addressing the risks: Towards a common understanding of data ethics 20
Box 3.1. Few countries had the real-time data they needed when the pandemic hit 32
Box 3.2. Data free flows with “trust”: An essential component of the global economy 33

Executive summary

Data use yields real benefits for individuals, organisations, and economies and societies as a whole – and there remains potential for more

Data and their flow, including across borders, underpin economic activity and well-being in global digital economies and societies. The COVID-19 pandemic highlighted data's potential to underpin crucial, life-saving services. However, it also underscored persistent gaps in the availability of data, notably real-time health data, that diminish this potential and could affect countries' preparedness for future crises. Access to data can empower consumers to make better purchasing and lifestyle decisions, including embracing societal goals such as sustainable consumption patterns. However, this opportunity remains largely untapped. For firms, the use of data can spur productivity and innovation, but uptake of crucial data processing technologies like data analytics and artificial intelligence remains skewed towards larger firms. Governments can use data to improve the design and delivery of public policies and services, but public sector data governance frameworks are often siloed and restricted to specific domains or applications, and thus fail to address policy issues applicable to data governance across the board.

Most data are collected by organisations, who seem to prize its control and use, but efforts to measure the value of data are still evolving

Organisations increasingly seek to collect, access and productively use data, indicating that data are valuable – but measuring that value remains challenging. Approaches explored by national statistical offices focus on measuring investment in, and stocks of, data assets. Existing estimates, which are all based on aggregating the costs associated with data collection, transformation and use, are sizeable: annual investment in total data assets was between 2.2% and 2.9% of value added in Australia (2016), 1.4% and 1.9% in Canada (2018), 2.4% and 3.0% in the Netherlands (2017) and 0.8% in the United States (2020). Other estimates based on a broader definition of data assets range between 3.8% and 6.6% of the market sector's value added in selected European countries. With other international organisations, the OECD is developing guidance to measuring data as an asset in the System of National Accounts. Nonetheless, because most data are generated, collected and used within organisations, and are not shared or traded, there is no single market price of data. Existing metrics do not capture data's strategic and innovative value to organisations, nor to society at large.

The increasingly widespread use of data also creates risks, which should be managed through policy, organisational and technological measures

Data's main potential for growth and well-being relies on increased data openness: the more data can be shared and reused, the more they can drive growth and well-being. However, increased access and sharing can create risks, including concerns related to the violation of privacy and personal data protection rights and intellectual property rights as well as digital security risks. In parallel, increasing data collection

by firms has given rise to concerns about competitive dynamics. These concerns are supported by empirical evidence on slowing productivity growth, rising industry concentration and growing gaps in technology adoption between firms of different sizes across the OECD. Regulatory and policy measures to address these risks and challenges – including conditions on cross-border data flows – can have unintended effects on other policy objectives and undermine the benefits of data use. Leveraging existing commonalities between countries together with informed policy making, including the adoption of organisational and technical measures, can support more calibrated policy approaches that address these risks and promote trust, yet enable responsible data access and sharing.

Data governance is a cross-cutting issue requiring integrated and holistic policy approaches

Data governance is increasingly relevant to different aspects of economies and societies, across policy domains from trade to competition and public governance. While these domains raise different challenges, they share a common set of data governance policy tensions: between openness and control, around incentives for investment in data and complementary resources, and between overlapping interests and frameworks. Data also underpin the functioning and operation of digital technologies, including artificial intelligence and the Internet of Things. As these digital technologies diffuse further across sectors and aspects of life, questions about how data is generated, collected, used and governed have become increasingly salient in today's digital society. Whole-of-government policy approaches to data governance, which integrate cross-cutting economic, social, cultural, technical, and legal governance issues across policy domains, such as through national data strategies, are essential to maximising the benefits of data for economies and societies, while addressing related risks and challenges and protecting important rights and interests.

The OECD can help governments navigate policy tensions and realise the full benefits of data for growth and well-being, while protecting individuals' and organisations' rights and interests

As a trusted forum for evidence-based, multi-disciplinary and multi-stakeholder policy analysis and international dialogue, the OECD can help governments better govern data in the digital age. The OECD's 2021 *Recommendation of the Council on Enhancing Access to and Sharing of Data* is a foundational standard to guide countries in realising the benefits of data while mitigating potential risks. It builds on pioneering advances made by the OECD in developing common principles for health data governance, access to research data and privacy. Other examples include supporting countries in harnessing their commonalities and advancing discussions on promoting cross-border data flows with trust, including through the articulation of commonalities in government access to personal data held by the private sector for law enforcement and national security purposes. The OECD also furthers efforts to open data, including data collected by the public sector. Efforts at the OECD have positive international spillovers, as requirements for data collection, management and use proliferate easily through an interconnected world. The OECD helps develop shared norms, standards and rules, as well as provide a platform for shared learning to accelerate data's use globally. The *Going Digital Guide to Data Governance Policy Making*, which complements this report, compiles and draws on real-life examples from countries, and provides a practical checklist to assist countries in developing their data governance policies.

1 Introduction

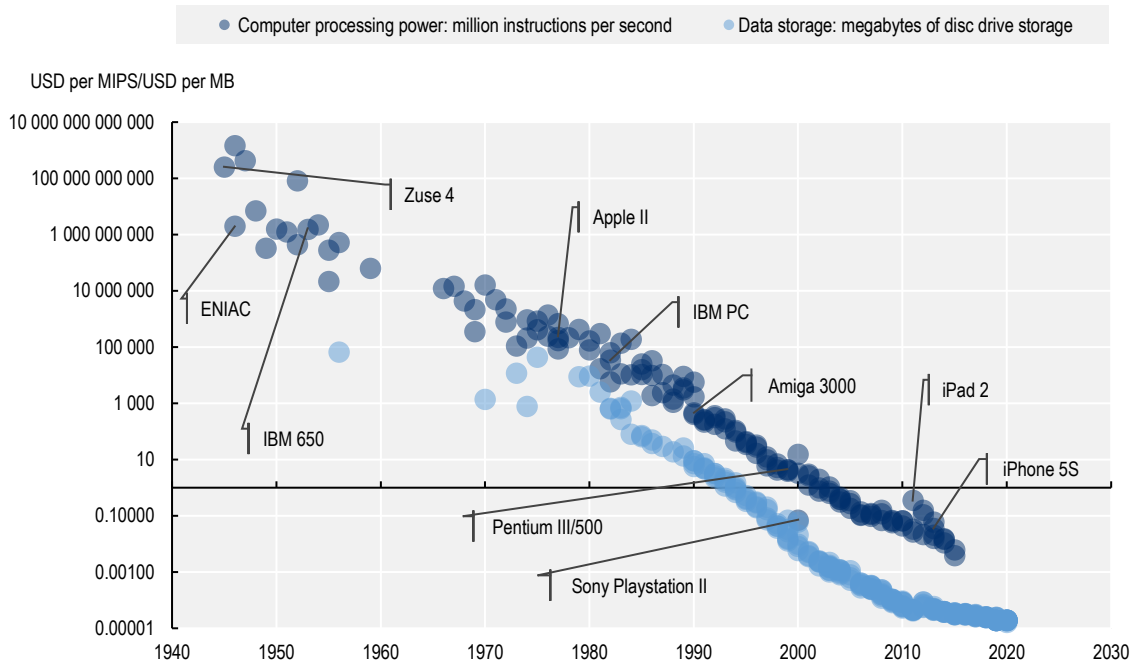
This chapter introduces the main issues and concepts underlying data governance. It explores the roots of “data” – from the first attempts to record information on cave walls to the emergence of computers in the 1940s. In the context of the OECD Project on Data Governance for Growth and Well-being, it defines “data” and “data governance”, and examines ways to balance benefits and risks related to data, including to rights and interests.

Humans have recorded information for millennia. The first known attempts were found in paintings in ochre and dirt, traced onto the walls of caves over 40 000 years ago (Marchant, 2016[1]). About 35 000 years later, our ancestors invented writing (Daniels, 1996[2]). After writing appeared on materials like papyrus and clay, the invention of paper in approximately 1AD enabled books to become the main repository of human knowledge for approximately 2 000 years (Tsien, 1985[3]). This trend was magnified by the development of the printing press in 1436 (Wolf, 1974[4]).

The information landscape changed with the arrival of the transistor in 1947 and the development of the integrated circuit in 1956 (Bell Labs, 2022[5]). Together, these two inventions paved the way for modern digital technologies, which are increasingly ubiquitous across the OECD and beyond. Mobile phones are used to access digital services that let us connect, transact and organise our lives, while digital sensors proliferate through infrastructure and production systems. Each of these digital interactions leaves a trail of data, which can be cheaply stored and processed (see Figure 1.1).

The proliferation of digital technologies has enabled a leap forward in the ability of humans to generate, collect and process information. Before the digital age, recording information took active human effort in some form, whether through observation or recording. In contrast, many sensor-equipped devices, also known as the "Internet of Things", collect data automatically, including as a by-product of an economic or social interaction. Through the Internet and communication infrastructures, data can be easily and quickly shared between other connected users and machines.

Figure 1.1. Cost of computer processing power and data storage, 1940-2020



Notes: MB: megabyte; MIPS: million instructions per second. Costs are expressed in 2020 prices and are deflated using the consumer price index (US Bureau of Economic Analysis, 2022[6]).
Source: OECD (2022[7]), based on data collected by Moravec (2022[8]), "Processor list", https://frc.ri.cmu.edu/~hpm/book97/ch3/processor.list.txt and McCallum (2021[9]), "Memory Prices 1957+", https://jcmit.net/memoryprice.htm.

This report develops an understanding of the main issues and concepts underlying data governance debates to foster a holistic and coherent approach to data governance, domestically and across borders. A common conceptual grasp of data and data governance is key to this understanding (see Box 1.1).

Box 1.1. What is data? What is data governance?

The OECD Project on Data Governance for Growth and Well-being seeks to provide policy guidance to help reap the benefits of data, address related challenges and foster a holistic and coherent approach to data governance. In the context of this report, "data" refer to recorded information in structured or unstructured formats, including text, images, sound and video. Data can be in any format, including analogue formats like paper, or emerging quantum forms like qubits. However, the rise of digital technologies has enabled the growth and policy relevance of digital data, namely information stored by a computer in binary format. Almost every aspect of the online experience, including a website or a banner advertisement, is data. Data in digital formats are characterised by their ability to be processed and analysed by digital technologies. Throughout this report, the term "data" will mean digital data unless otherwise stated.

This working policy definition of data can refer to one data point; several data points in a given dataset; one dataset; or many datasets. Put differently, the term "data" in this report does not carry a specific connotation about their volume. This definition does not refer to how data were collected, namely whether they were inferred, observed or volunteered, or any specific data type.

Finally, it is necessary to make a distinction between the definition of data used in this report, and the evolving understanding of data in the statistical community. Drawing on sources of statistical best practice, the OECD's Statistics Portal defines data as "characteristics or information, usually numerical, that are collected through observation". In collaboration with the international statistical community, the OECD is developing a robust definition of data for statistical and national accounting purposes. For example, the OECD participates in an Intersecretariat Working Group on National Accounts Advisory Expert Group. It recently proposed the following definition of data: "Information content that is produced by accessing and observing phenomena; and recording, organizing and storing information elements from these phenomena in a digital format, which provide an economic benefit when used in productive activities." While the statistical definition of data is evolving, its proposed use is intended to help establish which costs should be captured in determining the value of data and data assets in the System of National Accounts, as explored in Chapter 3.

In the context of the OECD Project on Data Governance for Growth and Well-being, "data governance" refers to diverse arrangements, including technical, policy, regulatory and institutional provisions, that affect data and their creation, collection, storage, use, protection, access, sharing and deletion, including across policy domains and organisational and national borders. Efforts to govern data take many forms. They often seek to maximise the benefits from data, while addressing related risks and challenges, including to rights and interests.

Sources: OECD (2021[10]), *Recommendation of the Council on Enhancing Access to and Sharing of Data*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463; OECD (2020[11]), "Data", https://stats.oecd.org/glossary/detail.asp?ID=532; ISWGNA (2022[12]), *Recording of Data in the National Accounts*, https://unstats.un.org/unsd/nationalaccount/RAdocs/DZ6_GN_Recording_of_Data_in_NA.pdf.

Chapter 2 examines how data have emerged as a strategic asset that can transform lives and markets and confer economic and market power, particularly for the few firms that use data to their full potential. Nevertheless, this opportunity carries risks, including to privacy, data protection rights and intellectual property rights. As the stakes of data use and misuse have increased, data-related policies have emerged in various policy domains and contexts but are rarely cross-cutting or co-ordinated.

A common understanding of data is necessary for effective, integrated policy making. Chapter 3 outlines the characteristics of data that can challenge conventional measurement methodologies. It also highlights how data are underpinned by digital technologies and an evolving technological landscape. These characteristics

make data a uniquely challenging subject for policy makers. Policies and institutions that pre-date the data-driven era must adjust to manage complex trade-offs between openness and control, overlapping interests and misaligned incentives for data collection and use.

Data's main potential for growth and well-being relies on increased data openness, but sharing data also increases the risk of misuse. Better and more co-ordinated policies are needed to help navigate these tensions and maximise the benefits of data governance for growth and well-being, as outlined in Chapter 4 of this report. The OECD is well placed to support governments in these efforts.

References

Bell Labs (2022), *1956 Nobel Prize in Physics*, https://www.bell-labs.com/about/awards/1956-nobel-prize-physics/#gref. [5]

Daniels, P. (1996), *The World's Writing Systems*, Oxford University Press, New York, https://eric.ed.gov/?id=ED418582. [2]

ISWGNA (2022), *Recording of Data in the National Accounts*, Intersecretariat Working Group on National Accounts, https://unstats.un.org/unsd/nationalaccount/RAdocs/DZ6_GN_Recording_of_Data_in_NA.pdf. [12]

Marchant, J. (2016), "A journey to the oldest cave paintings in the world", *Smithsonian Magazine*, No. 1/12, https://www.smithsonianmag.com/history/journey-oldest-cave-paintings-world-180957685/. [1]

McCallum, J. (2021), "Memory Prices 1957+", webpage, https://jcmit.net/memoryprice.htm (accessed on 10 January 2022). [9]

Moravec, H. (2022), "Processor List", webpage, https://frc.ri.cmu.edu/~hpm/book97/ch3/processor.list.txt (accessed on 20 January 2022). [8]

OECD (2022), "Measuring the value of data and data flows", *OECD Digital Economy Papers*, No. 345, OECD Publishing, Paris, https://doi.org/10.1787/923230a6-en. [7]

OECD (2021), *Recommendation of the Council on Enhancing Access to and Sharing of Data*, OECD/LEGAL/0463, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463. [10]

OECD (2020), "Data", *Glossary of Statistical Terms*, https://stats.oecd.org/glossary/detail.asp?ID=532 (accessed on 16 June 2022). [11]

Tsien, T. (1985), "Paper and Printing", in *Science and Civilisation in China, Chemistry and Chemical Technology*, Cambridge University Press. [3]

US Bureau of Economic Analysis (2022), "Personal Consumption Expenditures", *Chain-type Price Index*, Federal Reserve Bank of St Louis (FRED), https://fred.stlouisfed.org/series/DPCERG3A086NBEA. [6]

Wolf, H. (1974), *Geschichte der Druckpressen*, Interprint, Frankfurt/Main. [4]

2 Data as a strategic asset: The fundamental shift in their use, misuse and appreciation

Data have emerged as a strategic asset that can transform lives and markets and confer economic and market power. This chapter outlines why data and their role as a source of value and potential competitive advantage have emerged as a priority for individuals, organisations and nations. At the same time, it highlights risks associated with data collection and use. Noting that the stakes of data use and misuse have increased, this chapter discusses the emergence of data-related policies in various policy domains and contexts.

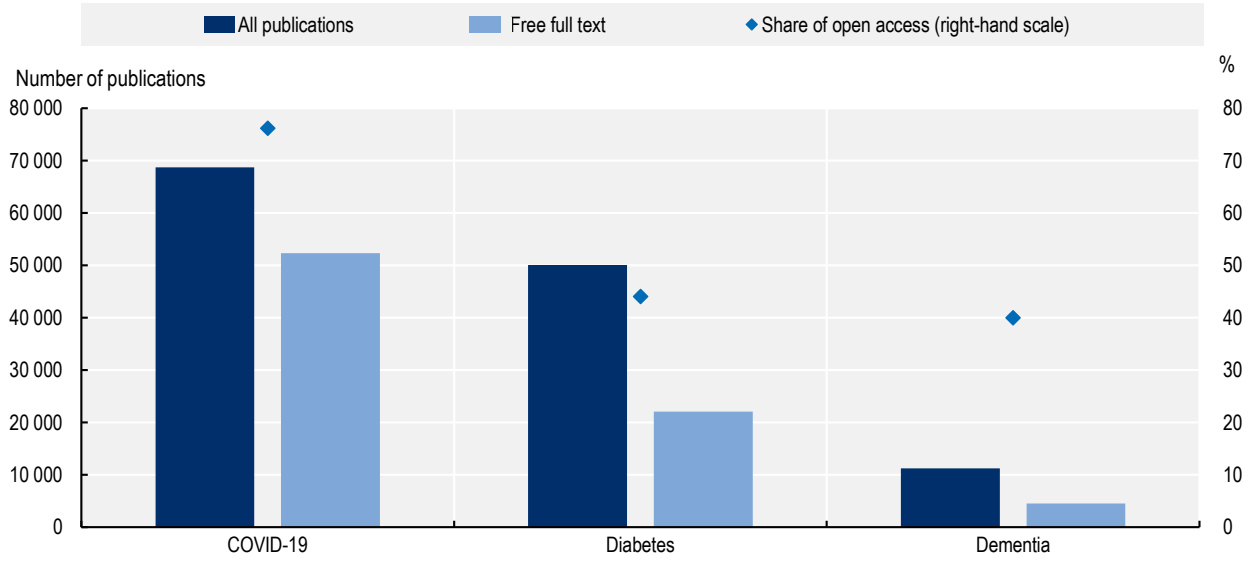
2.1. Data use is transformational for economies and societies

Data are a strategic asset due to their potential benefits for economies and societies. Data use is transformational through two main channels. First, insights gleaned from processing and analysing data can reveal patterns and relationships that enable better, evidence-based decision making. For consumers, this might mean making better and more empowered purchase decisions. For firms, data can be considered an input to production, including in combination with other, more traditional economic factors like labour or land. Second, data can bridge gaps between and among consumers and producers, or governments and citizens, facilitating new transactions and creating new markets. For example, the collection and sharing of data can enable more transparency between unknown third parties online. This can enable them to overcome previous information asymmetries that may have inhibited a successful interaction or transaction.

Through these two channels, the use of data, including their transfer across borders, can improve individual well-being and address societal challenges, as well as boost innovation and productivity. Data hold great potential across many areas of economies and societies. This includes the areas of public service design and delivery; science, research and development; education system monitoring and improvement; territorial management, including for smart cities; consumer empowerment and protection; and global development progress and co-operation (OECD, 2022[1]). Similarly, cross-border data flows play an enabling role for digital trade, including in the context of global value chain co-ordination (OECD, 2022[2]).

The COVID-19 crisis highlighted the growing importance of data to economies and societies, underscoring the role of data in responding to crises and societal challenges. The collection and use of data was crucial to almost all facets of the response to the COVID-19 pandemic. Indeed, an unprecedented amount of real-time and granular data was collected during the crisis (Paic, 2021[3]). For instance, data were fundamental to managing and improving health system performance, including better allocation of limited public health resources to fight the spread of the virus. Also, public bodies published open data and collaborated closely with the private sector. This enabled the co-design of services that citizens could use to better cope with the pandemic on a day-to-day basis (OECD and Govlab, 2021[4]).

Figure 2.1. Open access of COVID-19, diabetes and dementia publications, January-October 2020



Source: OECD (2021[8]), *OECD Science, Technology and Innovation Outlook 2021: Times of Crisis and Opportunity*, https://doi.org/10.1787/75f79015-en.

Data were essential to trace the spread of the virus, including by tracking contacts of confirmed cases and enabling monitoring. In Korea, for example, geolocation data, surveillance camera footage and credit card records were used to trace coronavirus patients (OECD, 2020[5]). In Israel, geolocation data were used to identify people coming into contact with virus carriers. This enabled authorities to inform them to isolate immediately (OECD, 2020[5]).

Data and their governance are also crucial elements of the global science system. Data provide evidence that enables knowledge creation and discovery, including to improve medical treatments and save lives. In particular, access to research data from public funding can help enable basic research and foster innovation networks (OECD, 2020[6]). During the pandemic, open science policies removed obstacles to the free flow of research data and ideas, and accelerated the pace of research critical to combatting the disease (OECD, 2020[8]; 2022[1]; Paic, 2021[3]). An unprecedented number of scientific publications was made openly available, with research databases removing paywalls so the scientific community could quickly share COVID-19 related information (see Figure 2.1). The growing volumes of research data also spurred the uptake of artificial intelligence (AI), including AI and data mining tools, which helped advance efforts to develop vaccines and better understand the virus.

2.2. Data-driven business models, including platforms, are transforming markets

While data are of value in many contexts, firms are at the forefront of realising the potential of data in their operations. Data can drive innovation and efficiency improvements, for instance by enabling real-time monitoring of firm operations. Data can also be used to co-ordinate business operations, including to manage supply chains and strengthen oversight. For example, in agriculture, sensor data from geocoded maps of fields can be linked with historical and real-time data on weather patterns, soil conditions, fertiliser usage and crop features, to optimise production (Jouanjean et al., 2020[9]). In manufacturing, streams of data from connected devices can be used to optimise operations and provide new or better after-sales services. Such data-enhanced business models create room for product differentiation, a key lever of competitiveness and performance for small and medium-sized enterprises (SMEs) (OECD, 2022[10]). Despite this significant potential and the ubiquity of data processing technologies across the OECD, there remains room for SMEs and lagging firms to use data and reap the benefits for economies and societies (OECD, 2021[11]; 2022[12]).

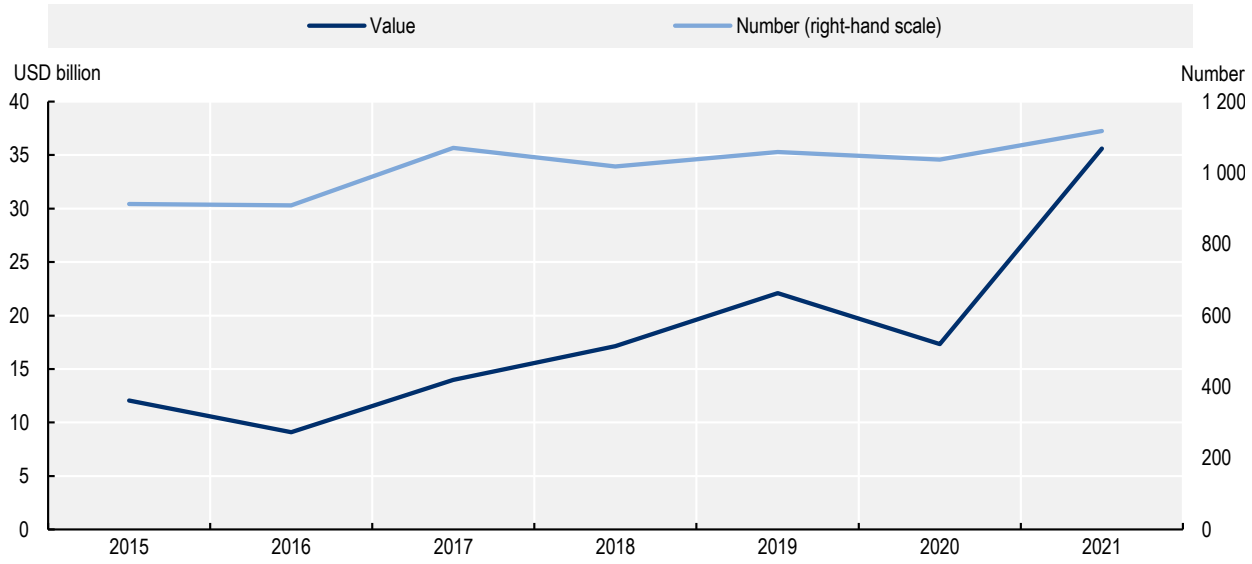
While some firms lag behind, the digital age has facilitated the rise of firms at the cutting edge of the technological frontier. Unlike firms whose operations are simply enhanced by data, some data-enabled firms rely on their ability to generate, collect and analyse data (Nguyen and Paczos, 2020[13]). For data-enabled firms, data are a critical input into their productive activities, and data or data-related tools may be among the most valuable assets they control. The most famous data-enabled firms are now household names across OECD countries, and have been among the largest firms in the world by capitalisation (Ker and Mazzini, 2020[14]). Markets increasingly value firms that can use data. Venture capital investments in “big data” firms, for example, have more than tripled since 2015 to reach USD 35.6 billion in 2021 (see Figure 2.2). Such investments reflect an evaluation of the long-term value of the data assets owned by these firms.

For firms, data have value to investors and strategic potential for innovative and market advantages. However, future returns from their generation and use may be uncertain. This can lead firms to view data systematically as a source of advantage that requires investment and protection. For example, firms may hoard data if sharing them exposes the firms to competition (Jones and Tonetti, 2020[15]). Alternatively, they may have no incentive to share data if they cannot privately capture the benefits of such an arrangement. Maximising the benefits of data to society by enabling and encouraging their wider use is thus a key concern of data governance frameworks (see Chapter 4).

Online platforms are examples of data-enabled firms at the frontier of markets and technological development. The core product of many prominent online platforms is a data-enabled service that connects multiple “sides” of a market, such as consumers, merchants and advertisers. The “matching” service offered by

such online platforms is enabled by the quantity and quality of data available to them, which they often collect from individual interactions. Many popular online platforms offer products at zero price to at least one side of the marketplace. Meanwhile, they generate their primary revenues from advertising (targeted using data) or successful transactions (namely, a commission from a successful data-driven match), incentivising greater data collection (OECD, 2022[17]). Many global companies provide zero-price products and services (to at least one side) using this business model (OECD, 2022[17]).

Figure 2.2. Venture capital deals in big data firms worldwide, 2015-21



Source: OECD (2022[16]), "Measuring the value of data and data flows", https://doi.org/10.1787/923230a6-en, based on Preqin Pro, http://www.pro.preqin.com/ (accessed 14 February 2022).

Platforms enjoy significant direct and indirect network effects where economies of scale benefit users on multiple sides of the market. In other words, as the number of users on one side increases, the value of the product to users on another side also increases. Because the marginal cost of adding an additional user to the platform can be close to zero, online platforms can rapidly scale up and expand their geographic coverage. This enables transactions between far-flung users that were previously impossible (OECD, 2022[17]; 2019[18]). As platforms scale up, so do their opportunities to gather data to improve their services. This potentially drives a feedback loop that could result in few players in a given market (also known as a "winner takes most" or "winner takes all" effect) (OECD, 2022[12]).

2.3. The shift in the nature and use of data introduces economic and social risks

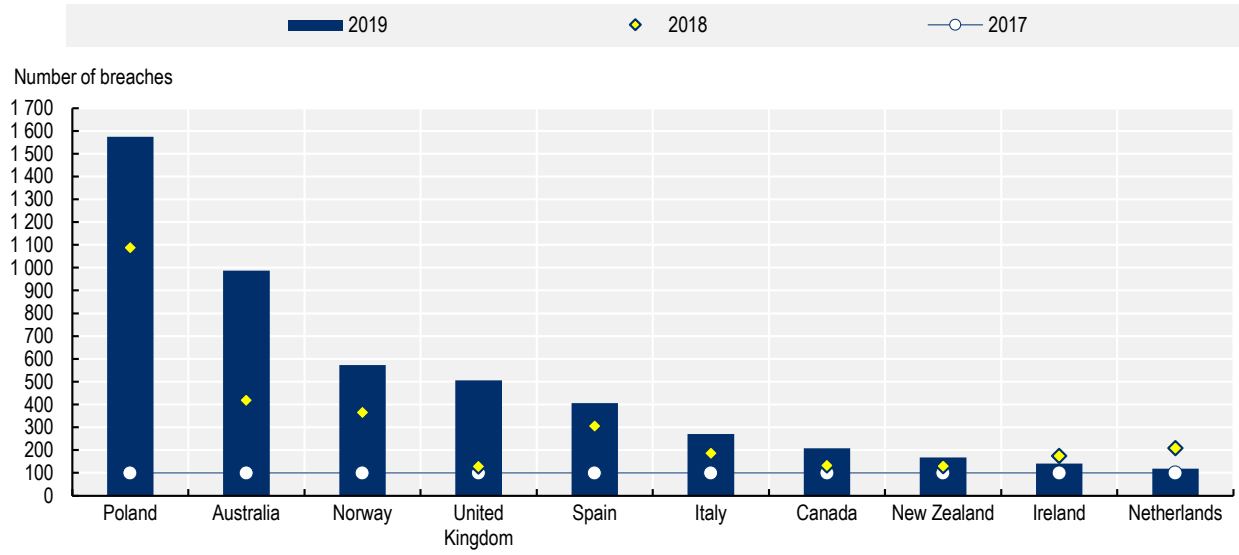
As the significance of data to economies and societies has grown, so too have the potential harms associated with their use and misuse. In particular, the combination of the growing collection of data, and increased access and use of data analytics and related digital technologies, has created the potential for deliberate or accidental misuse of data.

A key risk relates to the violation of privacy and personal data protection rights. Indeed, several OECD jurisdictions reported a general increase in the number of personal data breach notifications over two years (see Figure 2.3). Even where individuals and organisations agree on (and consent to) specific terms for data sharing and data re-use, including on the purposes for which the data should be re-used, there remains a significant level of risk that the data may end up being used differently, including by a third party.

Issues of informed consent become more difficult in an environment where consumers often face information overload, and technologies like AI enable more unanticipated uses for data (OECD, 2022[19]). Similarly, data collection and processing might be less transparent or apparent. The ubiquity and discreteness of connected devices, for example, might sometimes mask that personal data are constantly collected. This collection can often occur without an easily accessible interface for user interaction and control (e.g. to set personal data gathering preferences) (OECD, 2021[20]).

Moreover, online businesses can control the "choice architecture" online to a greater extent than in an offline interaction or transaction. This can result in "dark commercial patterns" – business practices that use elements of digital choice architecture, especially in online user interfaces, that subvert or impair consumer autonomy, decision making or choice (OECD, 2022[21]). Such practices often deceive, coerce or manipulate consumers and are likely to cause direct or indirect consumer detriment in various ways. However, it may frequently be difficult or impossible to measure such detriment. Evidence suggests they are highly prevalent on websites, apps and cookie consent notices (OECD, 2022[21]). In the context of data collection, such practices can lead consumers to give up more personal data than they may have otherwise chosen. They do this, for example, by making privacy-intrusive settings the default or making it hard to opt out of them. In addition, the data that online businesses collect through consumer interactions increasingly allow the construction of fine-grained consumer profiles. Businesses may be able to leverage information asymmetries from such data profiling to exploit consumer vulnerabilities at a highly granular level. However, there is little evidence to suggest that such practices are widespread (OECD, 2022[21]).

Figure 2.3. Personal data breaches are on the rise across a number of OECD jurisdictions



Notes: Australia, New Zealand and United Kingdom data refer to the fiscal year. Canada data combines data on number of breaches for the public and private sectors. Numbers of personal data breach notifications (PDBNs) in 2017 are normalised to 100. Numbers of PDBNs in 2018 and 2019 are compared against this normalised value. To avoid overrepresentation of changes from small numbers, data under 50 were eliminated from the calculation.

Source: Iwaya, Koksal-Oudot and Ronchi (2021[22]), "Promoting comparability in personal data breach notification reporting", https://doi.org/10.1787/88f79eb0-en.

The risk of privacy and personal data protection rights violation also exists outside of the commercial context. Data typically collected by private companies for their business purposes have become increasingly valuable to governments. For instance, government law enforcement and national security agencies have made more requests to private companies to access their user data (Llanos, 2021[23]). While legitimate public interests can justify such access, the elevated number of requests also raises concerns about infringements of privacy and civil liberties, which can erode trust among governments and between governments and

individuals. Furthermore, they can affect personal data flows across borders and create uncertainty and compliance costs for businesses (OECD, 2022[2]). In response, major companies have begun to publish transparency reports on demands from government agencies. However, such disclosures may not be possible in some jurisdictions (Llanos, 2021[23]) (see section 4.2 for OECD policy efforts to address this trust gap).

Regulatory frameworks can be slow to adapt to a fast-moving technological landscape. Consequently, they can be unable to address evolving uses and misuses of data that are viewed as unethical or that may generate undesirable outcomes. Factors that complement regulatory or legal issues, and that are often considered ethical in nature, include issues such as fairness, respect for human dignity, autonomy, self-determination, human rights and the risk of bias and discrimination. These ethical considerations and factors provide additional guidance for data governance and use, beyond mere legal compliance.

A key example concerns the role of data in AI or other algorithmically driven systems, and the possibility of creating feedback loops that reinforce biases. Recent work explored the effect of AI on the lives of working women, for example. It highlights that gendered data used by labour market intermediaries to target online job advertisements can worsen gaps between the number of men and women in some occupations (UNESCO, OECD and IDB, 2022[24]). Moreover, the underlying data used as input into AI systems may not be sufficiently representative. They may also perpetuate historical gender stereotypes around gender roles in labour, care and domestic work. The rise of questionable uses of data underscores the importance of a common understanding of data ethics to guide the use of data where regulatory frameworks may fall short (OECD, 2022[1]). OECD standards such as the OECD Good Practice Principles for Data Ethics in the Public Sector aim at providing further action-oriented guidance in this area (OECD, 2019[25]) (see Box 2.1).

Box 2.1. Addressing the risks: Towards a common understanding of data ethics

Often considered in the frame of broader human rights, data ethics approaches can complement legal approaches and help inform how regulations should be drafted or revised. In particular, taking a data ethical approach can help build trust in responsible organisations. Data ethics audits, impact assessments or certifications can help build overall trusted data transactions. Such schemes should take a multi-stakeholder approach and include civil society organisations, as well as academics and researchers.

A wide range of data ethics definitions and frameworks has been developed across jurisdictions, often encompassing legal, policy and voluntary approaches. Such frameworks tend to provide high-level principles to guide the use of data, including in AI applications. Notably, the private sector and professional and representative bodies, which often self-regulate in data-related fields, increasingly promulgate data ethics frameworks.

In the public sector, ensuring the ethical use of data calls for defining and embedding value-based approaches in the daily management and use of data by public servants and public-sector organisations. The OECD Good Practice Principles for Data Ethics in the Public Sector emerged from observed data ethical practices in OECD countries and partner economies, and their actions to develop data ethical capabilities in the public sector. As their overarching message, the Good Practice Principles state that government data use should serve the public interest and deliver public good. They also discuss the collective and community nature of data governance, the environmental impact of data infrastructures and the abuse of data use during electoral campaigns.

Note: This box is based on an expert workshop on Data Ethics: Balancing Ethical And Innovative Uses of Data, which was co-hosted by the OECD and Danish Business Authority on 9-10 December 2021.
Sources: Jobin, Ienca and Vayena (2019[26]), "The global landscape of AI ethics", https://doi.org/10.1038/s42256-019-0088-2; New Zealand Government (2020[27]), "International Data Ethics Frameworks", https://www.data.govt.nz/assets/Uploads/Discussion-paper-International-data-ethics-frameworks-March-2020.pdf; OECD (2019[25]), "Good Practice Principles for Data Ethics in the Public Sector", https://www.oecd.org/digital/digital-government/good-practice-principles-for-data-ethics-in-the-public-sector.htm.

A different type of risk relates to the lack of incentives for firms to share data. Firms may choose not to share data if they have limited ability to profit from the benefits of the use of data for other firms or society at large. This, in turn, diminishes the utility of data to society. In addition, concerns regarding the intellectual property rights of organisations and the protection of their commercial interests can negatively affect incentives to invest in data collection and data use. For SMEs, for example, identifying which data to share and defining the scope and conditions for access and re-use are perceived as major challenges. Moreover, firms may be reluctant to share data because of potentially significant costs. Sharing data inappropriately could lead to costs associated with privacy violations, as well as opportunity costs of potential innovation. For example, sharing data prematurely can undermine the ability to obtain patent and trade secret protection (OECD, 2021[28]; 2019[29]; 2019[30]).

The use of data by companies is of considerable interest to competition authorities. If data represent a significant barrier to entry to a market, firms that control that data may be able to act in a manner that may have anti-competitive effects. For example, they might exclude competitors from a market if they control access to their data inputs (OECD, 2020[31]). Mergers could strengthen the effect of such behaviour. Similarly, when firms have particularly valuable or complementary datasets, a potential merger could give rise to an advantage that is difficult for competitors to overcome and result in durable market power (OECD, 2020[32]; 2020[33]). Another risk of mergers is that personal data collection in one market could also be leveraged in a related market, which could limit the contestability of both markets. In other words, where data access is limited and essential to compete, challengers would need to offer a matching ecosystem of products to match an incumbent's position (Condorelli and Padilla, 2019[34]). Data and their role in fostering network effects, economies of scale, and scope and feedback loops that can lead to increasing concentration in digital markets, is also an area of concern for competition authorities (OECD, 2022[17]).

The use of data by firms also has wider effects on the economy. Data and other intangibles have taken on increasing importance to knowledge- and service-intensive production. The growing importance of data and other intangibles may have disproportionately benefitted the largest global firms (Bajgar, Criscuolo and Timmis, 2021[35]; Corrado et al., 2021[36]; 2022[37]). Investments in intangible assets, including data, within industries are linked to increasing productivity dispersion in those industries. Namely, as data become important to a given industry and investment in data grows accordingly, gaps in the productivity of firms within those industries grow as well (Corrado et al., 2021[36]; 2022[37]). Intangibles-intensive industries are also becoming increasingly concentrated, with less churning at the top. This indicates that big firms are getting bigger and tend to stay at the top of their industries (Bajgar, Criscuolo and Timmis, 2021[35]).

These dynamics threaten long-term inclusive growth and well-being across the OECD. Productivity growth allows economies and societies to benefit from a wider range of outputs with the same level of input; productivity growth is the most important factor in increases in living standards and economic growth (OECD, 2015[38]). The varying rates of productivity growth between firms has been previously linked to wage inequality (Berlingieri, Blanchenay and Criscuolo, 2017[39]). They may also be related to the trend towards greater income and wealth inequality in advanced economies in recent years. Moreover, reaping the productivity benefits of increased data use is essential to managing forthcoming social challenges, like demographic shifts and climate change. It also enables long-term economic growth and well-being across the OECD. Promoting access to and use of data across all sectors and company sizes, and overcoming related barriers, can enable a wider distribution of the benefits of data use.

2.4. The fundamental shift in data use is reflected in public policies at the national and international levels

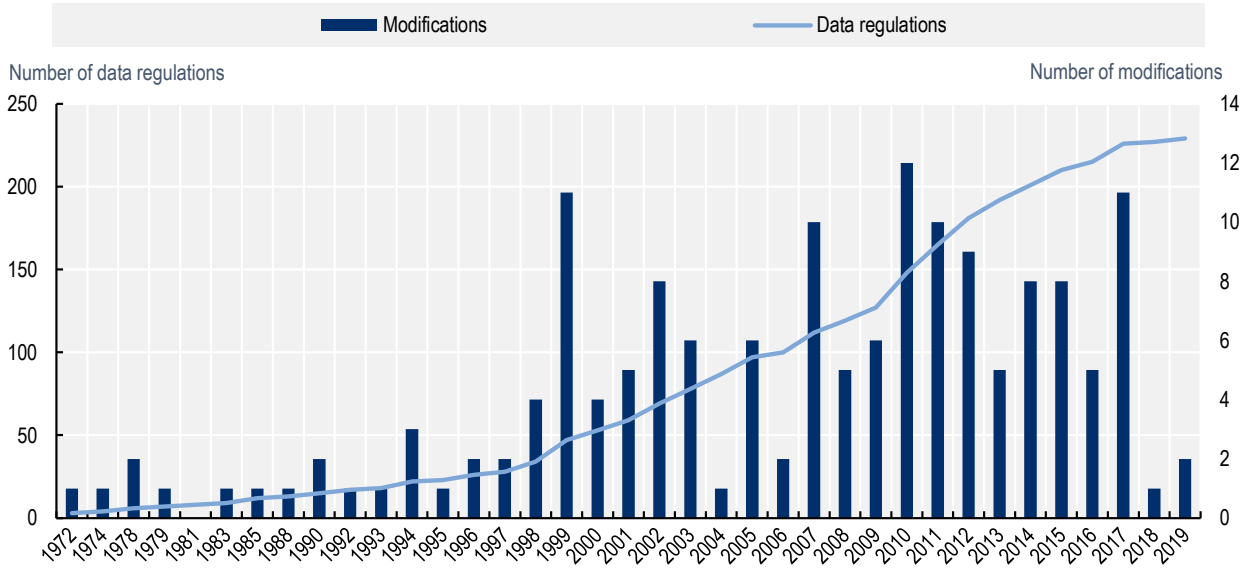
Both the potential and realised transformative effect of data on economies and societies, and the attendant risks, have led to more public policies and regulations that target data. This growth in policy attention also reflects changes resulting from the COVID-19 pandemic and more awareness of the need for policy to

ensure wider yet responsible use of data. Such measures usually respond directly to the opportunities and risks of data use outlined in sections 2.2 and 2.3. However, they rarely attempt to address the interrelated policy issues related to data in a cross-cutting manner.

The COVID-19 crisis, for example, highlighted the need for timely data for decision making (see also section 3.4). Consequently, governments have increased the collection and sharing of data from health systems. In the first half of 2020, few countries scored highly on dataset availability, maturity and use and dataset governance (Oderkirk, 2021[40]). Yet in 2021, 15 of 24 surveyed OECD countries had enacted legal, regulatory or policy reforms to improve health data availability, access or sharing. Meanwhile, 16 of those countries had introduced new technologies to improve health data governance. Governments also introduced measures to improve data linkage and sharing and improve sector-related capacities. As a result of these reforms, most surveyed countries saw significant improvement in the timeliness and quality of key health datasets (de Bienassisi et al., 2022[41]).

With more collection, exchange and use of data, including across borders, over the last two decades, governments have put in place policies to govern the transfer of data across borders (see Figure 2.4). In some cases, this includes the location for storage of such data. Governments may put in place such measures for a number of reasons. These include the desire to preserve competition and contestable markets, data protection, privacy and digital security; to enable regulatory control or audits; or to uphold national security. In addition, because the collection, use and control of data are also often viewed as a competitive advantage at the national level, some such regulations also act as a kind of digital industrial policy, including in the context of economic development (OECD, 2022[2]).

Figure 2.4. Growing number of regulations affecting cross-border data flows



Note: This figure includes different types of regulation relating to data transfers and local storage requirements.
Source: Casalini and López González (2019[42]), "Trade and Cross-Border Data Flows", https://doi.org/10.1787/b2023a47-en.

Another specific form of data-related regulation growing across the OECD includes *ex ante* competition regulation in certain digital markets. This growth reflects concerns that competition frameworks and *ex post* competition enforcement are not adequate for increasingly data-driven and digital markets. These *ex ante* regulations often aim to increase market contestability by imposing requirements on a particular handful of firms. These firms are usually the largest online platforms, which are perceived as having a particular ability to harm competition in digital markets. Notably, the targeted firms, sometimes designated "gatekeepers",

may have a position in markets that might differ from the traditional definition of "dominance" in competition frameworks (OECD, 2022[17]).

In addition to other measures like transparency and business conduct obligations, merger requirements and obligations to limit "self-preferencing" and bundling, many forms of proposed *ex ante* regulation address data-related concerns. Such measures reflect a view that access to and control of data can act as a structural barrier to entry and can carry the risk of exclusionary and exploitative conduct. Indeed, part of the function of these data-related measures is to "[grant] other firms an edge to compete with designated firms", namely those firms that are considered "gatekeepers" (OECD, 2021[43]). Data-related measures in *ex ante* regulations include obligations for data portability and interoperability, prohibitions on combining certain datasets i.e. including after merger activity, and obligations to grant access to certain kinds of datasets to competitors.

References

Bajgar, M., C. Criscuolo and J. Timmis (2021), "Intangibles and industry concentration: Supersize me", *OECD Science, Technology and Industry Working Papers*, No. 2021/12, OECD Publishing, Paris, https://doi.org/10.1787/ce813aa5-en. [35]

Berlingieri, G., P. Blanchenay and C. Criscuolo (2017), "The great divergence(s)", *OECD Science, Technology and Industry Policy Papers*, No. 39, OECD Publishing, Paris, https://doi.org/10.1787/953f3853-en. [39]

Casalini, F. and J. López González (2019), "Trade and Cross-Border Data Flows", *OECD Trade Policy Papers*, No. 220, OECD Publishing, Paris, https://doi.org/10.1787/b2023a47-en. [42]

Condorelli, D. and J. Padilla (2019), *Harnessing Platform Envelopment in the Digital World*, Department of Economics, University of Warwick, https://doi.org/10.2139/ssrn.3504025. [34]

Corrado, C. et al. (2021), "New evidence on intangibles, diffusion and productivity", *OECD Science, Technology and Industry Working Papers*, No. 2021/10, OECD Publishing, Paris, https://doi.org/10.1787/de0378f3-en. [36]

Corrado, C. et al. (2022), "The value of data in digital-based business models: Measurement and economic policy implications", *OECD Economics Department Working Papers*, No. 1723, OECD Publishing, Paris, https://doi.org/10.1787/d960a10c-en. [37]

de Bienassisi, K. et al. (2022), "Health data and governance developments in relation to COVID-19: How OECD countries are adjusting health data systems for the new normal", *OECD Health Working Papers*, No. 138, OECD Publishing, Paris, https://doi.org/10.1787/aec7c409-en. [41]

Iwaya, S., E. Koksal-Oudot and E. Ronchi (2021), "Promoting comparability in personal data breach notification reporting", *OECD Digital Economy Papers*, No. 322, OECD Publishing, Paris, https://doi.org/10.1787/88f79eb0-en. [22]

Jobin, A., M. Ienca and E. Vayena (2019), "The global landscape of AI ethics guidelines", *Nature Machine Intelligence*, Vol. 1, pp. 389-399, https://doi.org/10.1038/s42256-019-0088-2. [26]

Jones, C. and C. Tonetti (2020), "Nonrivalry and the economics of data", *American Economic Review*, Vol. 9, pp. 2018-2058, https://doi.org/10.1257/aer.20191330. [15]

Jouanjean, M. et al. (2020), "Issues around data governance in the digital transformation of agriculture: The farmers' perspective", *OECD Food, Agriculture and Fisheries Papers*, No. 146, OECD Publishing, Paris, https://doi.org/10.1787/53ecf2ab-en. [9]

Ker, D. and E. Mazzini (2020), "Perspectives on the value of data and data flows", *OECD Digital Economy Papers*, No. 299, OECD Publishing, Paris, https://doi.org/10.1787/a2216bc1-en. [14]

Llanos, J. (2021), "Transparency reporting: Considerations for the review of the Privacy Guidelines", *OECD Digital Economy Papers*, No. 309, OECD Publishing, Paris, https://doi.org/10.1787/e90c11b6-en. [23]

New Zealand Government (2020), "Discussion Paper: International Data Ethics Frameworks", prepared for the Government Chief Data Steward for the Data Ethics Advisory Group, https://www.data.govt.nz/assets/Uploads/Discussion-paper-International-data-ethics-frameworks-March-2020.pdf. [27]

Nguyen, D. and M. Paczos (2020), "Measuring the economic value of data and cross-border data flows: A business perspective", *OECD Digital Economy Papers*, No. 297, OECD Publishing, Paris, https://doi.org/10.1787/6345995e-en. [13]

Oderkirk, J. (2021), "Survey results: National health data infrastructure and governance", *OECD Health Working Papers*, No. 127, OECD Publishing, Paris, https://doi.org/10.1787/55d24b5d-en. [40]

OECD (2022), "Dark commercial patterns", *OECD Digital Economy Papers*, No. 336, OECD Publishing, Paris, https://doi.org/10.1787/44f5e846-en. [21]

OECD (2022), "Data shaping firms and markets", *OECD Digital Economy Papers*, No. 345, OECD Publishing, Paris, https://doi.org/10.1787/923230a6-en. [12]

OECD (2022), "Fostering cross-border data flows with trust", *OECD Digital Economy Policy Papers*, No. 343, OECD Publishing, Paris, https://doi.org/10.1787/139b32ad-en. [2]

OECD (2022), *Going Digital Guide to Data Governance Policy Making*, OECD Publishing, Paris, https://doi.org/10.1787/40d53904-en. [19]

OECD (2022), *Japan-OECD Joint Workshop on Data Free Flow with Trust*, OECD, Paris. [10]

OECD (2022), "Measuring the value of data and data flows", *OECD Digital Economy Papers*, No. 345, OECD Publishing, Paris, https://doi.org/10.1787/923230a6-en. [16]

OECD (2022), *OECD Handbook on Competition Policy in the Digital Age*, OECD, Paris, https://www.oecd.org/daf/competition/oecd-handbook-on-competition-policy-in-the-digital-age.pdf. [17]

OECD (2022), "Responding to societal challenges with data: Access, sharing, stewardship and control", *OECD Digital Economy Papers*, No. 342, OECD Publishing, Paris, https://doi.org/10.1787/2182ce9f-en. [1]

OECD (2021), "Ex ante regulation and competition in digital markets", *OECD Competition Background Note*, OECD, Paris, https://www.oecd.org/daf/competition/ex-ante-regulation-and-competition-in-digital-markets-2021.pdf. [43]

OECD (2021), *OECD Science, Technology and Innovation Outlook 2021: Times of Crisis and Opportunity*, OECD Publishing, Paris, https://doi.org/10.1787/75f79015-en. [8]

OECD (2021), *OECD SME and Entrepreneurship Outlook 2021*, OECD Publishing, Paris, https://doi.org/10.1787/97a5bbfe-en. [28]

OECD (2021), *Report on the Implementation of the Recommendation of the Council Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*, OECD, Paris, https://one.oecd.org/document/C(2021)42/en/pdf. [20]

OECD (2021), *The Digital Transformation of SMEs*, OECD Studies on SMEs and Entrepreneurship, OECD Publishing, Paris, https://doi.org/10.1787/bdb9256a-en. [11]

OECD (2020), "Abuse of Dominance in Digital Markets", *OECD Competition Background Note*, OECD, Paris, https://www.oecd.org/daf/competition/abuse-of-dominance-in-digital-markets-2020.pdf. [31]

OECD (2020), "Consumer data rights and competition", *OECD Competition Background Note*, OECD, Paris, https://one.oecd.org/document/DAF/COMP(2020)1/en/pdf. [32]

OECD (2020), *Enhanced Access to Publicly Funded data for Science, Technology and Innovation*, OECD Publishing, Paris, https://doi.org/10.1787/947717bc-en. [6]

OECD (2020), "Merger control in dynamic markets", *OECD Competition Background Note*, OECD, Paris, https://www.oecd.org/daf/competition/merger-control-in-dynamic-markets-2020.pdf. [33]

OECD (2020), "Using artificial intelligence to help combat COVID-19", *OECD Policy Responses to Coronavirus (COVID-19)*, OECD, Paris, https://www.oecd.org/coronavirus/policy-responses/using-artificial-intelligenceto-help-combat-covid-19-ae4c5c21/. [5]

OECD (2020), "Why open science is critical to combatting COVID-19", *OECD Policy Responses to Coronavirus (COVID-19)*, OECD, Paris, https://www.oecd.org/coronavirus/policy-responses/why-open-science-is-critical-to-combatting-covid-19-cd6ab2f9/#:~:text=In%20global%20emergencies%20like%20the,critical%20to%20combating%20the%20disease. [7]

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. [30]

OECD (2019), "Good Practice Principles for Data Ethics in the Public Sector", webpage, https://www.oecd.org/digital/digital-government/good-practice-principles-for-data-ethics-in-the-public-sector.htm (accessed on 14 April 2022). [25]

OECD (2019), *OECD SME and Entrepreneurship Outlook 2019*, OECD Publishing, Paris, https://doi.org/10.1787/34907e9c-en. [29]

OECD (2019), *Unpacking E-commerce: Business Models, Trends and Policies*, OECD Publishing, Paris, https://doi.org/10.1787/23561431-en. [18]

OECD (2015), *The Future of Productivity*, OECD Publishing, Paris, https://doi.org/10.1787/9789264248533-en. [38]

OECD and Govlab (2021), "Open Data in Action: Initiatives During the Initial Stage of the COVID-19 Pandemic – OECD", webpage, https://www.oecd.org/fr/gouvernance/gouvernement-numerique/use-of-open-government-data-to-address-covid19-outbreak.htm (accessed on 14 April 2021). [4]

Paic, A. (2021), "Open Science – Enabling Discovery in the Digital Age", *Going Digital Toolkit Note*, No. 13, https://goingdigital.oecd.org/data/notes/No13_ToolkitNote_OpenScience.pdf. [3]

UNESCO, OECD and IDB (2022), *The Effects of AI on the Working Lives of Women*, UNESCO/OECD/Inter-American Development Bank, https://doi.org/10.18235/0004055. [24]

3 Understanding data: Key characteristics, technological developments and policy challenges

What is specific and unique about data that fuels such economic and social change? This chapter describes what data are, how they can be characterised and how these characteristics challenge the measurement of their value, as well as their relationship with digital technologies. In so doing, it clarifies a policy debate that is complicated by a lack of a common conceptual understanding of data. The chapter concludes by outlining the cross-cutting policy objectives and tensions at the heart of data governance policy making and the potential role of technological and organisational approaches.

3.1. Key characteristics of data

Data are intangible, meaning they do not have a physical or financial embodiment. Like other intangible assets, like software, they are viewed as an asset because they are a source of economic value, can provide future economic returns and are usually under the effective control of an economic actor (Haskel and Westlake, 2018[1]). Data are also non-rival (Jones and Tonetti, 2020[2]). It is impossible for two people to drive the same car at the same time, or for a single drop of oil to be used in two different cars. However, unlike cars or oil, data are theoretically infinitely usable. Data can be used productively and continuously, including simultaneously by multiple actors or machines, without reducing the amount of data available to other actors or applications.

Some non-rival goods are also non-excludable, meaning that access to the good cannot be easily restricted. Take the example of a lighthouse: it would be nearly impossible for a lighthouse keeper to restrict the light to guide some boats and not others. Excludability is a spectrum, however. Thus, use of data can be restricted through a variety of means, including intellectual property rights, data protection regimes or encryption (Ostrom, 2010[3]). Data are not inherently non-excludable, although data that are made available or circulated e.g. on the Internet, cannot easily be controlled, and their use may not be easily restricted.

Data are subject to a range of externalities, which is to say that data sharing and use can generate wider benefits and costs for those who may not be directly implicated in the transaction. Pollution is a classic example of a negative externality: the negative effect of air pollution on society is not usually captured in the private cost to car owners or oil- or car-producers. Data and their sharing and use can be subject to negative externalities. For example, when an individual chooses to share data on line, this could be used to gain information about others who did not consent to the data being shared, and who are not compensated for what might constitute a loss of privacy (Acemoglu et al., 2020[4]; Coyle et al., 2020[5]).

However, data can also exhibit positive externalities, often due to synergies. With synergies, value emerges from comparison, aggregation or processing for information; individual data points often have little to no value. For example, knowing the location of one coronavirus patient may be of little use. However, combining this data point with the location of many other positive patients contributes to a more accurate map of the spread of COVID-19. Importantly, the benefits of this increased accuracy e.g. through better public health management, also flow to those people who did not share their data. A key aspect of the economic and social benefits of data come from such positive spillovers from data's productive use. Chapter 2 examines the benefits of data use for economies and societies, but these gains likely pale in comparison to the potential social benefits from enabling data locked in organisations to be shared more widely.

Externalities can be a source of market failure, indicating that private incentives may not deliver the full range of economic and social benefits possible from greater investment, sharing and use of data. For example, firms may not have an incentive to share data if they cannot capture the benefits of this sharing. There may also be an over-production or collection of some kinds of data, and underproduction of data in contexts with few commercial opportunities (Coyle et al., 2020[5]). In turn, market failure provides a rationale for government intervention (see Chapter 4).

When data are used as a factor of production, they can exhibit increasing returns to scale. This means the output from the use of data increases by a larger proportion than an increase in data volume.[1] A virtuous cycle can emerge where data analysis yields data-driven insights, and the use of such learnings increases product quality and/or scope of operations such that more data are generated. The benefits of this cycle of value creation could be amplified across many actors and across the global economy because of data's non-rival characteristics. However, realising these benefits depends on access to data, as well as significant complementary investments and capacities, like skills or computing (OECD, 2022[6]).

Although data's value can often be a function of their scale, and data are easily generated *en masse* and automatically by digital technologies, data are highly *heterogenous*. As the technical means to record information

expands, data can relate to any observable phenomenon. Policy frameworks often recognise this heterogeneity by identifying different settings for data with different information content or different origins. For example, weather data are usually treated differently in policy frameworks to health data. The latter are usually both personal in nature and subject to ethical standards like medical confidentiality (OECD, 2019[7]).

Data are often co-produced, meaning they are often the result of interactions between many different actors and/or machines. This can mean that many actors can have multiple and overlapping claims to data. Through their choices to share or use data, individuals or organisations might create positive or negative externalities for each other. In addition, the non-rival nature of data and their ability to easily flow and be exploited in multiple applications can challenge conventional notions of ownership (OECD, 2022[8]).

3.2. Understanding the value of data

These characteristics mean that the value of data can be hard to measure. The value of data often relates to their content, namely the information embodied in data after processing, and the context in which the data are collected, stored and used (Mitchell, Ker and Lesher, 2021[9]). Because they can be used in many current and future applications, data may also have considerable option value. This means that incentives exist to collect and store data even without any immediate plans for their use. Similarly, the lack of defined ownership or property rights, and the heterogeneity and other characteristics of data, pose challenges to the emergence of data markets. This is especially true for large-scale, multilateral markets in which data are exchanged under standardised terms (OECD, 2022[10]).

As a result, market statistics can provide an illustrative, but limited, picture of the value of data. For example, while the vast majority of firms do not trade data, firms may indirectly monetise data by selling data-driven services like targeted advertising. The market capitalisation of such firms, or their reported revenues from the sale of data-intensive services, or international trade in such services, can shed light on the value of data. However, data-driven services combine data with other inputs, including digital technologies like AI. This makes it difficult to isolate the value of data. Moreover, classifications in statistical frameworks do not generally help delineate data-driven services, or the firms that provide them. Such estimates would also fail to capture the value of data for the firms that collect data and use them for their internal operations (OECD, 2022[10]).

Nevertheless, data clearly have value as an input into productive activities for organisations. Moreover, better capture of that value in macroeconomic statistics through the System of National Accounts (SNA) would help enable better and more comparable estimates of investments in, and the stock of, data across economies.

With other international institutions, the OECD is working to explicitly incorporate business production and use of data into the planned 2025 update of the SNA statistical framework. In view of the issues with market-based valuations of data, a consensus has emerged on the use of the “sum-of-cost” approach. This measures the value of data and data assets based on the incurred costs of production, like compensation of employees and use of fixed capital.

Many practical challenges persist for the sum-of-cost approach. These include understanding which costs to include, identifying assumptions that should inform estimates of costs and avoiding overlap with the costs included in estimates of other intangible assets, Developing guidance on these issues is a key part of the work of the OECD and the international statistical community for the next update of the SNA (OECD, 2022[10]).

National statistical institutes across the OECD have developed experimental estimates for the value of data assets using the sum-of-cost approach, although with varying underlying assumptions. These initial estimates are sizeable: annual investment in total data assets was between 2.2% and 2.9% of total value added in Australia (2016), 1.4% and 1.9% in Canada (2018), 2.4% and 3.0% in the Netherlands (2017) and 0.8% in the United States (2020). Estimates by academia, based on a broader definition of data assets, range between 3.8% and 6.6% of the market sector’s value added in selected EU countries. Nevertheless,

empirical studies suggest other parts of the SNA already capture these expenditures. This implies that implementation of these efforts in macroeconomic statistics will help better separate and understand the value of data and data assets. However, these efforts will have limited effect on macroeconomic statistics like gross domestic product and productivity (OECD, 2022[10]).

3.3. Data and digital technologies are fundamentally linked

Recorded information could theoretically be in any format, including analogue formats like paper, or emerging quantum forms like qubits. However, policy efforts have largely focused on the governance of digital data, namely information stored by a computer in binary format. Unlike some other inputs into production, like air or water, digital data are not naturally occurring or managed. Instead, digital data depend on digital technologies for their generation, collection, storage, transfer and use.

The emergence of vast amounts of digital data stems from the deployment of broadband networks. These enable more networked users and machines in an increasing number of contexts, sectors and applications to generate and transmit digital data. The proliferation and increasing frequency of use of connected devices like the smartphone enables the observation and recording of a range of primary data. It also permits the collection of inferred or secondary data, such as with respect to user behaviour in the digital environment or previous searching or purchasing patterns.

These digital data are put to productive use through the use of digital technologies like AI and data analytics. Increasingly statistical and probabilistic AI systems use data to “train” models to make accurate predictions, recommendations or decisions (OECD, 2019[11]). In turn, such data processing can enable the extraction of valuable insights from datasets, which may have otherwise remained unexploited. The magnitude of possible opportunities and risks of data collection has increased as a result of advances in the capacities of these digital technologies.

More generally, when organisations use data in their activities and operations, they often adopt digital technologies in the process of generating, collecting and storing data, in addition to analysing them. Cloud computing is a digital technology that enables scalable access to computing resources, including software, storage and processing infrastructures. Current cloud computing business models often involve the transfer of data to data centres that centralise computing services and resources. Across the OECD, the uptake of data analytics appears complementary to the uptake of cloud computing (OECD, 2022[6]). This highlights that cloud computing may enable more actors, including SMEs, to use data.

Data governance policies should seek to consider the digital technologies that underpin the generation, collection, transfer and productive use of data. For example, a current concern of data governance involves regulations that mandate that some kinds of data are processed, or at least one copy is stored, within a given jurisdiction. But evolutions in digital technologies, including the virtualisation of broadband networks and the increasing bandwidth requirements of the Internet of Things, may mean that more data are processed and stored at the “edge” of the network, including within intelligent devices themselves. This could reduce the need for data to cross borders in some applications (OECD, 2022[12]).

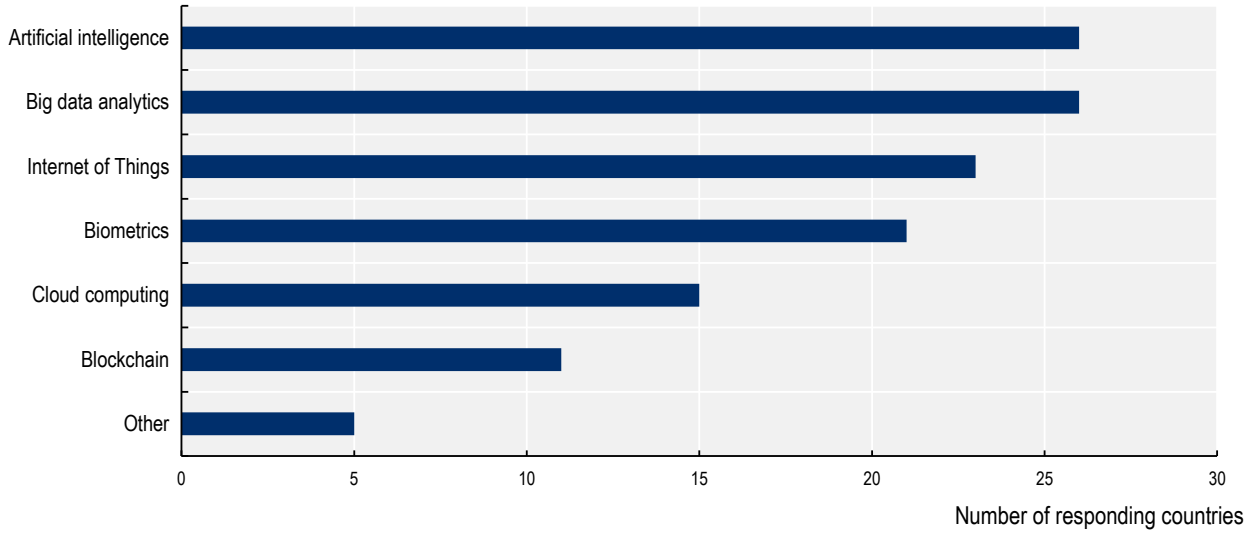
3.4. Coping with the new data-driven reality

With many policies pre-dating the data-driven era, existing institutions, organisations and legal frameworks must adjust to the growing importance of data and the shifts in their use. Many policy frameworks have struggled to keep up with a shifting technological landscape. Meanwhile, institutions have been slow to collect, share and use data to their full potential.

Privacy enforcement authorities increasingly acknowledge the difficulties with safeguarding privacy in the digital age. A recent review of the *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* (hereafter the *OECD Privacy Guidelines*) (OECD, 2021[13]) acknowledged that technological

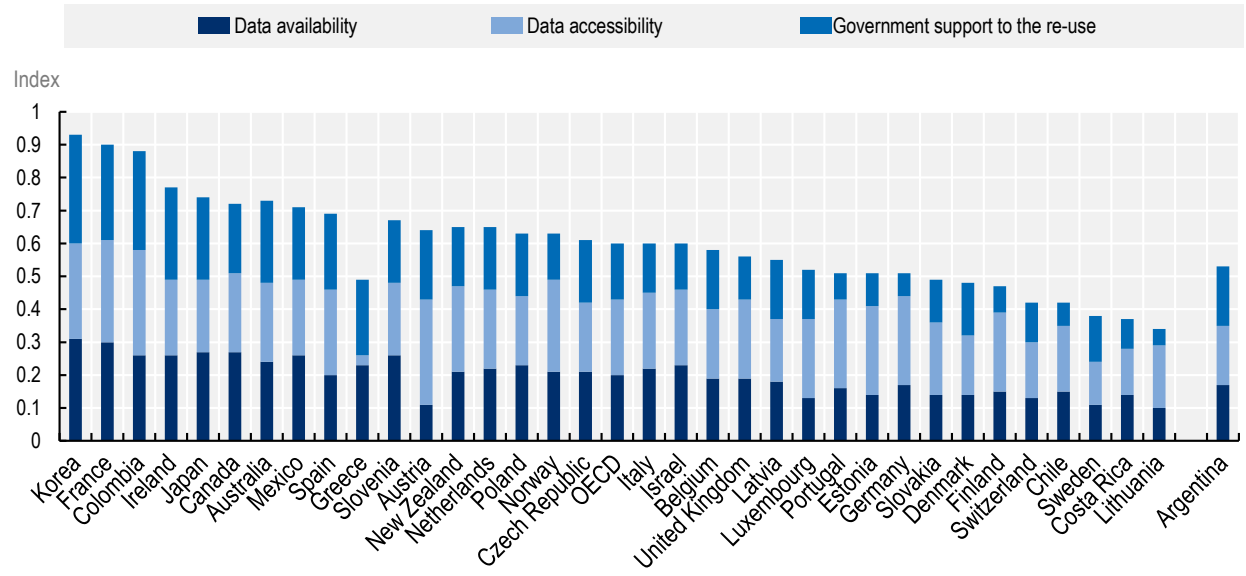
advances are expanding the methods and ease with which individuals may be identified by their data or re-identified from apparently anonymised data. The rapid pace of such advances means that data that are considered private today may not be private tomorrow (OECD, 2022[14]). For example, in 2019, all but four of the surveyed privacy enforcement authorities noted that technological developments were the main challenge to regulatory frameworks. Most frequently, they cited big data analytics and AI as the greatest risk (see Figure 3.1).

Figure 3.1. Emerging technologies that pose the main challenges for privacy and personal data protection



Source: OECD (2021[13]), *Report on the Implementation of the Recommendation of the Council Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data,* https://one.oecd.org/document/C(2021)42/en/pdf.

Figure 3.2. OECD Open, Useful and Re-usable data (OURdata) Index: 2019



Note: Data are not available for Hungary, Iceland, Türkiye and the United States. Data for Costa Rica were collected from the IDB-OECD Open Government Data Survey 2018.
Source: OECD (2020[20]), "Open, Useful and Re-usable data (OURdata) Index: 2019", https://doi.org/10.1787/45f6de2d-en.

Governments are among the largest producers and consumers of data, and making these data available can be a cornerstone of innovation for other actors (OECD, 2015[15]; 2019[16]). In particular, governments can aim to make data more available through stronger open data agendas. Potential exists to further enable access to public data. For example, the quantity of available datasets only slightly increased from 2017 to 2019 (OECD, 2019[17]).

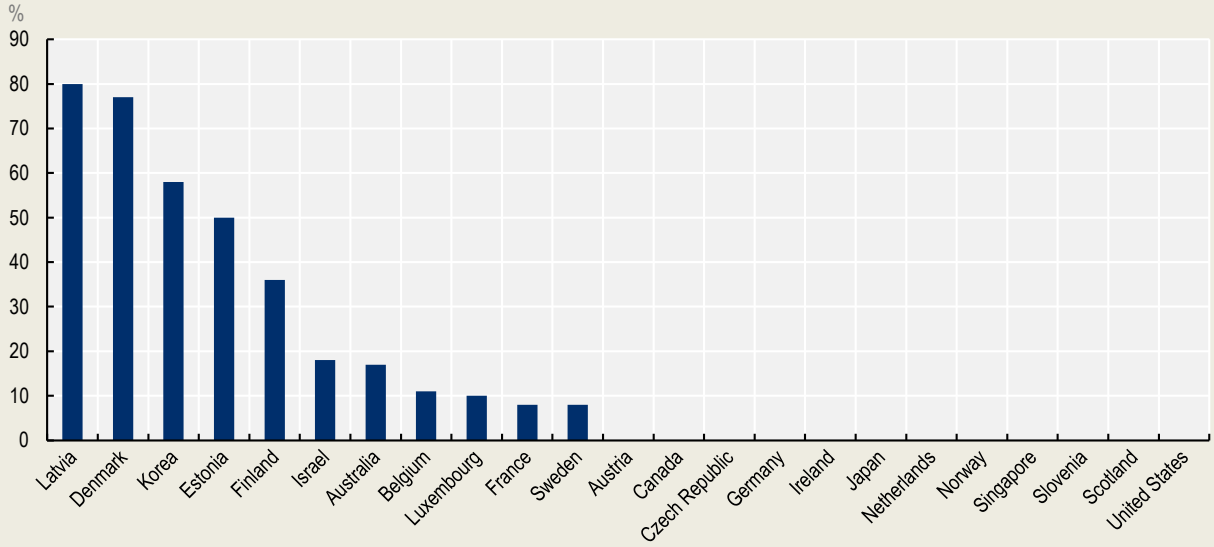
Of policy measures to foster and enhance access to data, most (65%) still focus on public sector data (OECD, 2019[18]). In this regard, open government data has been a priority for countries for over a decade to support business innovation, social value creation and government transparency. In 2018, as many as 8 out of 10 OECD countries had a strategy for open government data, and 9 out of 10 had requirements for public-sector organisations to publish open data in a machine-readable format (OECD, 2020[19]). Also, results from the 2019 Open, Useful and Re-usable data (OURdata Index) show that "the OECD average increased from 0.54 in 2017 to 0.60 in 2019, indicating a greater general maturity of open data policies at the central level" (OECD, 2020[20]).

Box 3.1. Few countries had the real-time data they needed when the pandemic hit

The lack of timely data is one of the key areas of concern for policy makers combatting the COVID-19 pandemic. Figure 3.3 shows that just before the pandemic hit, few countries had real-time data about key parts of the health care system. Denmark, Estonia, Korea and Latvia stood out for having a short time lapse, of one week or less, between when a data record was first created and when it was included in the national dataset used for analysis across all or most key national health datasets. The main barriers cited by countries related to legal restrictions or policy barriers to data sharing among policy authorities or undertaking data linkages. Since the pandemic, several countries have acted to improve timely information for decision making (see section 2.4).

Figure 3.3. Real-time data available to governments as the COVID-19 pandemic hit

Share of available public health datasets where the time between record creation and inclusion in the dataset was one week or less



Source: Oderkirk (2021[23]), "Survey results: National health data infrastructure and governance", https://doi.org/10.1787/55d24b5d-en.

In contrast, only 1 out of 10 OECD countries had a dedicated, comprehensive public sector data strategy covering a broader spectrum of data access and sharing arrangements, and their enablers. Examples of such strategies include Ireland's Public Service Data Strategy 2019-23 (Ireland DPER, 2019[21]) and the Danish government's Basic Data Strategy, published as early as 2012 (Local Government Denmark, 2012[22]). Few countries have policy initiatives to facilitate data sharing within the private sector, although sharing and re-use of private sector data is a frequently cited challenge (OECD, 2019[18]).

Similarly, governments have been slow to realise the benefits of data in informing better policies and decision making. For example, although governments invested in health data and governance in response to COVID-19 (section 2.4), the pandemic revealed the weakness of data collection and sharing frameworks (see Box 3.1).

In a globalised and digital economy, enabling cross-border flows of data remains essential for many individuals, firms and organisations. However, data crossing jurisdictional boundaries can raise concerns about the application and enforcement of domestic policies. Governments are increasingly placing conditions on cross-border data flows. This results in a complex landscape for many actors, including regulators. For example, in 2019, privacy enforcement authorities most often noted "uncertainty regarding legal privacy regimes" and "incompatibility of legal regimes" as the main challenges to transborder data flows (OECD, 2021[13]). Despite this uncertainty, recent OECD work highlights that many frameworks show a degree of commonalities, convergence or complementarities. These can be used to foster a foundation of trust for data flows (Casalini, López González and Nemoto, 2021[24]) (see Box 3.2).

Box 3.2. Data free flows with "trust": An essential component of the global economy

Data have become integral to international trade, and cross-border data flows are necessary for enabling social interactions, facilitating trade in digital products, co-ordinating production along global value chains and enabling firms to access global markets. However, despite the international relevance of data, countries make policies at the national level. This includes policies that aim to protect privacy, digital security, national security and other public policy interests. These data governance policies can nonetheless have international impacts as international firms are forced to comply with multiple, intersecting policy frameworks. Such a patchwork can carry a risk of international fragmentation in the trade of data and digital products across borders. This fragmented landscape also shifts frequently, as new arrangements proliferate and others are struck down. This, in turn, contributes to regulatory uncertainty that may preclude the ability of firms to make long-term plans and investments.

The challenge for policy makers is promoting international discussions that create a global and interoperable digital environment where data can move across borders with trust. Despite some differences between countries, there remain areas of complementarity, convergence and commonality among national approaches to cross-border data flows. This indicates the potential for architectures that seek to combine the benefits of data flows and the achievement of public policy objectives. A promising approach focuses on interoperability of privacy regimes, understood as the "ability of different privacy and data protection regimes, or legal frameworks, to work together at multiple levels through policy and practical arrangements and thereby bridge any differences in approaches and systems of privacy and personal data protection to facilitate transborder flows of personal data".

Sources: OECD (2022[14]), "Fostering cross-border data flows with trust", https://doi.org/10.1787/139b32ad-en; Robinson, Kizawa and Ronchi (2021[25]), "Interoperability of privacy and data protection frameworks", https://doi.org/10.1787/64923d53-en.

3.5. Cross-cutting policy objectives and tensions and promising approaches

Data governance is characterised by overlapping policy tensions and objectives, and different intersections across policy domains. These overlaps call for whole-of-government approaches and new organisational and technical mechanisms to data governance. In particular, data intermediaries and privacy-enhancing technologies show promise in navigating the new data governance landscape.

The *Going Digital Guide to Data Governance Policy Making* (OECD, 2022[26]) introduces a conceptual foundation of policy objectives and tensions related to data governance that apply across policy domains and a checklist to put it into practice. These policy tensions result from the underlying characteristics of data (section 3.1), their benefits (section 2.1) and potential risks (section 2.3):

1. **Balancing data openness and control while maximising trust:** Data can be characterised by different levels of openness on a spectrum ranging from full and unrestricted openness to arrangements that condition or limit the access, sharing, transfer and/or use of data to specific users, destinations or use cases. Higher levels of openness imply an increased ability to reap the benefits of data. However, they are often associated with a higher degree of risk, often linked to loss of control.
2. **Managing overlapping and potentially conflicting interests and regulations related to data:** Multiple stakeholders are often involved in data generation and may have competing or overlapping rights, interests or legal control over these data. Data governance structures may need to evolve to accommodate these interests or enable them to be prioritised.
3. **Incentivising investments in data and their effective re-use:** Investments are often required to generate, collect and use data. However, the original data holder may not have an incentive to make such investments or share data if the benefits of data use accrue to other actors or society.

These tensions cut across policy domains and need to be addressed holistically. For example, the increasing importance of data collected from or about individuals calls for cross-regulatory co-operation between privacy enforcement authorities (and frameworks) and other regulators, such as competition authorities. As outlined in section 2.2, some firms may be in a better position to collect and generate data, and this position may be self-reinforcing. Where a firm's control of non-replicable datasets can inhibit competition and distort the level playing field, authorities may feel compelled to intervene in the market to enhance access to data. Yet, when the data in question are personal and under the purview of data protection and privacy laws, it may conflict with privacy frameworks to provide third parties with access to such data (OECD, 2022[6]).

On the other hand, data protection requirements may be interpreted as distorting competition if equivalent transfers of personal data are possible within large incumbents i.e. between subsidiaries but not between independently owned businesses (CMA/ICO, 2021[27]). Some OECD jurisdictions have interpreted the personal data collection practices of firms, and the resulting reduction in privacy, as an abuse of market power and a violation of competition law (OECD, 2022[28]; 2020[29]).

The relationship between competition and privacy frameworks highlights several of the cross-cutting policy opportunities and tensions outlined above. Namely, privacy considerations can be conceptualised as a legal compulsion to keep data more closed and under the effective control of relatively few users. Meanwhile, some efforts to increase competition involve opening access to data. Efforts to mediate this conflict include data portability measures, which enable controlled data transfers between parties upon request (OECD, 2021[30]).

However, firms may not be incentivised to invest in collecting and using data if the expected returns to such investments are reduced by legally compelled data sharing. Similarly, effective sharing of data to enable their wider use relies on technical standards and interoperability of data across use cases, which is a sub-dimension of openness. This relationship underscores that formerly disparate policy frameworks now require a whole-of-government approach and cross-regulatory co-operation to address data governance challenges, adapt policies and realise the potential of data in the digital age.

New organisational and technical approaches may complement traditional policy making to address some of the tensions. "Data intermediaries" – "service providers that facilitate data access and sharing under commercial or non-commercial agreements between data holders, data producers, and/or users" (OECD, 2021[31]) – are expected to play a key role in the data ecosystem. Added-value services provided by data intermediaries could include data processing services (including data aggregation), payment and clearing services, as well as legal services including the provision of standard licence and certification schemes (OECD, 2019[18]; forthcoming[32]).

Privacy-enhancing technologies (PETs) also show promise to enable wider re-use of data while managing potential risks to privacy and security. PETs enable data to be processed without disclosing their inputs, or preventing their (re-)identification. As such, they might fundamentally alter the way organisations gather, access and process data, including personal data.

In general, PETs include data accountability tools, data obfuscation tools, encrypted data processing tools and distributed analytics that enable more control for data subjects. PETs are increasingly cited in a variety of policy domains as a means of moving closer to the goal of "privacy by design" (OECD, forthcoming[33]). However, developments in data analytics and AI combined with the increasing volume and variety of available data sets and the capacity to link these different data sets have altered the technical landscape. Essentially, these changes have made it easier to infer and relate seemingly non-personal or anonymised data to an identified or identifiable entity (OECD, 2022[12]).

References

Acemoglu, D. et al. (2020), "Too much data: Prices and inefficiencies in data markets", *NBER Working Paper*, No. 26296, https://www.nber.org/papers/w26296. [4]

Bajari, P. et al. (2019), "The impact of big data on firm performance: An empirical investigation", *AEA Papers and Proceedings*, No. 109, https://www.aeaweb.org/articles?id=10.1257/pandp.20191000. [34]

Casalini, F., J. López González and T. Nemoto (2021), "Mapping commonalities in regulatory approaches to cross-border data transfers", *OECD Trade Policy Papers*, No. 248, OECD Publishing, Paris, https://doi.org/10.1787/ca9f974e-en. [24]

CMA/ICO (2021), "Competition and data protection in digital markets: A joint statement between the the CMA and the ICO", Competition & Markets Authority and Information Commissioner's Office, London, https://ico.org.uk/media/about-the-ico/documents/2619797/cma-ico-public-statement-20210518.pdf. [27]

Coyle, D. et al. (2020), *The Value of Data – Policy Implications*, Bennett Institute for Public Policy, Cambridge in partnership with the Open Data Institute, https://www.bennettinstitute.cam.ac.uk/publications/value-data-policy-implications/. [5]

Haskel, J. and S. Westlake (2018), *Capitalism without Capital: The Rise of the Intangible Economy*, Princeton University Press, Princeton. [1]

Iansiti, M. (2021), "The value of data and its impact on competition", *Harvard Business School NOM Unit Working Paper*, No. 22-002, https://doi.org/10.2139/ssrn.3890387. [35]

Ireland DPER (2019), *Public Service Data Strategy 2019-2023*, Department of Public Expendiure and Reform, Ireland, https://www.gov.ie/en/publication/1d6bc7-public-service-data-strategy-2019-2023/. [21]

Jones, C. and C. Tonetti (2020), "Nonrivalry and the economics of data", *American Economic Review*, Vol. 9, pp. 2018-2058, https://doi.org/10.1257/aer.20191330. [2]

Local Government Denmark (2012), *Good Basic Data For Everyone - A Driver for Growth and Efficiency*, https://en.digst.dk/media/18773/good-basic-data-for-everyone-a-driver-for-growth-and-efficiency.pdf. [22]

Mitchell, J., D. Ker and M. Lesher (2021), "Measuring the economic value of data", *OECD Going Digital Toolkit Notes*, No. 20, https://doi.org/10.1787/f46b3691-en. [9]

Oderkirk, J. (2021), "Survey results: National health data infrastructure and governance", *OECD Health Working Papers*, No. 127, OECD Publishing, Paris, https://doi.org/10.1787/55d24b5d-en. [23]

OECD (2022), "Data in an evolving technological landscape: The case of connected and automated vehicles", *OECD Digital Economy Papers*, No. 346, OECD Publishing, Paris, https://doi.org/10.1787/ec7d2f6b-en. [12]

OECD (2022), "Data shaping firms and markets", *OECD Digital Economy Papers*, No. 344, OECD Publishing, Paris, https://doi.org/10.1787/7b1a2d70-en. [6]

OECD (2022), "Fostering cross-border data flows with trust", *OECD Digital Economy Policy Papers*, No. 343, OECD Publishing, Paris, https://doi.org/10.1787/139b32ad-en. [14]

OECD (2022), *Going Digital Guide to Data Governance Policy Making*, OECD Publishing, Paris, https://doi.org/10.1787/40d53904-en. [26]

OECD (2022), "Measuring the value of data and data flows", *OECD Digital Economy Papers*, OECD Publishing, Paris, https://doi.org/10.1787/923230a6-en. [10]

OECD (2022), *OECD Handbook on Competition Policy in the Digital Age*, OECD, Paris, https://www.oecd.org/daf/competition/oecd-handbook-on-competition-policy-in-the-digital-age.pdf. [28]

OECD (2022), "Responding to societal challenges with data: Access, sharing, stewardship and control", *OECD Digital Economy Papers*, No. 342, OECD Publishing, Paris, https://doi.org/10.1787/2182ce9f-en. [8]

OECD (2021), "Mapping data portability initiatives, opportunities and challenges", *OECD Digital Economy Papers*, No. 321, OECD Publishing, Paris, https://doi.org/10.1787/a6edfab2-en. [30]

OECD (2021), *Recommendation of the Council on Enhancing Access to and Sharing of Data*, OECD/LEGAL/0463, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463. [31]

OECD (2021), *Report on the Implementation of the Recommendation of the Council Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*, OECD, Paris, https://one.oecd.org/document/C(2021)42/en/pdf. [13]

OECD (2020), "Consumer data rights and competition", *OECD Competition Background Note*, OECD, Paris, http://www.oecd.org/daf/competition/consumer-data-rights-and-competition.htm. [29]

OECD (2020), "Digital Government Index: 2019 results", *OECD Public Governance Policy Papers*, No. 03, OECD Publishing, Paris, https://doi.org/10.1787/4de9f5bb-en. [19]

OECD (2020), "Open, Useful and Re-usable data (OURdata) Index: 2019", *OECD Public Governance Policy Papers, No. 01*, OECD Publishing, Paris, https://doi.org/10.1787/45f6de2d-en. [20]

OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris, https://doi.org/10.1787/eedfee77-en. [11]

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. [18]

OECD (2019), *Going Digital: Shaping Policies, Improving Lives*, OECD Publishing, Paris, https://doi.org/10.1787/9789264312012-en. [16]

OECD (2019), *Government at a Glance 2019*, OECD Publishing, Paris, https://doi.org/10.1787/8ccf5c38-en. [17]

OECD (2019), *Recommendation of the Council on Health Data Governance*, OECD/LEGAL/0433, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0433. [7]

OECD (2015), *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Paris, https://doi.org/10.1787/9789264229358-en. [15]

OECD (forthcoming), *Companion document to the OECD Council Recommendation on Enhancing Access to and Sharing of Data*, OECD Publishing, Paris. [32]

OECD (forthcoming), "Emerging privacy enhancing technologies: Maturity, opportunities and challenges", *OECD Digital Economy Papers*, OECD Publishing, Paris. [33]

Ostrom, E. (2010), "Beyond markets and states: Polycentric governance of complex economic systems", *American Economic Review*, Vol. 100/3, pp. 641-672, https://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.100.3.641. [3]

Robinson, L., K. Kizawa and E. Ronchi (2021), "Interoperability of privacy and data protection frameworks", *Going Digital Toolkit Note*, No. 21, OECD Publishing, Paris, https://doi.org/10.1787/64923d53-en. [25]

Endnote

[1] Data used in a single application might exhibit diminishing returns to scale over time. Additional data can have a large initial effect on model precision, but the effect may decrease with increasing sample size (Iansiti, 2021[35]). For example, Bajari et al. (2019[34]) find that increasing data sample size over time improves the accuracy of forecasts but at a diminishing rate.

4 Policy considerations

Data are a building block of the digital economy. Drawing on the body of work realised across the OECD, this chapter provides considerations for policy makers when developing or revising data governance policies. It highlights that policy choices need to promote growth, well-being, innovation and competition through increased access to and use of data. However, policies must ensure that incentives and safeguards exist for data to be collected and processed in a responsible manner that appropriately manages the risks of increasing data openness, including to individual privacy.

4.1. Data governance is a cross-cutting issue that requires a whole-of-government approach

The main characteristics of data, including their non-rivalry, the presence of externalities and their co-generation, as well as the spectrum of excludability, imply that the generation and use of data can involve multiple actors and applications. In addition, they can have effects that ripple across interrelated issues and concerns. Use of data also allows products and services to cut across previously well-established boundaries between markets, jurisdictions and regulations. This indicates that policies to govern data should be developed in a whole-of-government manner, with due account to potential spillovers.

Navigating these common policy tensions may require co-operation between parts of the government and sectors that may not have interacted beforehand or that may have different objectives. Departments and authorities concerned with issues such as consumer protection, privacy enforcement, competition, health, finance and economics, financial stability and other regulators may all also need to collaborate. In this way, they can ensure that data governance arrangements (from policies to regulation and technical platforms) are appropriately holistic to manage risks and benefit from these opportunities. The *Going Digital Guide to Data Governance Policy Making* provides a conceptual foundation drawing on real-life examples from countries, and a checklist to support governments in doing so. National data strategies may help ensure that policy objectives and tensions underlying data governance are evaluated in a holistic manner, foster co-ordinated data governance policies and enable increased and more responsible data access and sharing (OECD, 2022[1]; 2022[2]).

4.2. Promoting cross-border data flows with trust requires an international approach

The digital economy is increasingly global and interconnected. The Internet enables data to flow easily across boundaries and forges new connections between individuals, firms and machines. Data are a highly mobile factor of production. Their non-rival nature implies the benefits of enabling their greater use could drive greater economic and social returns at the global level (see section 2.1). Nevertheless, although governments consider the international context, the ways in which countries enable data flows vary. In view of the spillovers underlying the use of data across the world, a growing number of international bodies and forums has called for international co-operation on data governance issues. These include the United Nations Conference on Trade and Development (2021[3]), the International Monetary Fund (2021[4]), the World Bank (2021[5]), the G7 (2022[6]) and the G20 (2019[7]). A key role for such international collaboration is the promotion of "trust" between a diverse set of stakeholders as a basis for enabling cross-border data flows and realising their benefits. In addition, efforts to further cross-border co-operation in regulatory matters, including for privacy enforcement, are also necessary for a global approach to data governance.

International dialogue in forums like the OECD can help promote data free flow with trust (OECD, 2022[8]), building on countries' shared values and commonalities, complementarities and areas of convergence in instruments used to enable data to flow across borders (Casalini, López González and Nemoto, 2021[9]). The *OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data* (hereafter the *OECD Privacy Guidelines*) (OECD, 2013[10]) have played a foundational role in this respect. Ongoing work to develop principles on government access to personal data held by the private sector (OECD, 2022[11]) seeks to fill gaps at the international level. Interoperability of privacy regimes, introduced in the 2013 revision of the *OECD Privacy Guidelines*, has been increasingly considered a means to promote data free flow with trust (Robinson, Kizawa and Ronchi, 2021[12]; OECD, 2022[8]).

4.3. New technologies and governance arrangements can help align incentives and balance data-related benefits and risks

As outlined in the *Going Digital Guide to Data Governance Policy Making* (OECD, 2022[1]), data governance issues are characterised by policy tensions with respect to overlapping interests related to data governance. These, in turn, can affect incentives and disincentives to invest in data, and policy decisions related to the access, use and control of data. New organisational and technical mechanisms, including data intermediaries and privacy-enhancing technologies (PETs), as discussed in section 3.5, can complement other means of policy making in resolving some of these tensions.

Organisational measures also show some promise in enabling control, access to and use of data in such a way that would allow only approved users (including data subjects), within and across organisational borders, to access and use data. For example, data intermediaries can enable data access and sharing under commercial or non-commercial agreements (OECD, 2021[13]) by helping manage control over data. Some notable examples, including trusted data-sharing platforms or data trusts, often also feature the application of PETs.

PETs can help increase user control and confidentiality of data while enabling their use by other actors. Data accountability tools, for example, enable data access, while ensuring limitations and restrictions are attached to, and follow, data as they are shared. For their part, data obfuscation tools limit the need for sensitive data to leave a data subject's device or organisation, or alter the underlying data to preclude identification. Encrypted data processing tools enable computation of data without disclosure, while distributed analytics enable remote processing of data within common networks. However, developments in data analytics and artificial intelligence, alongside the increasing volume and variety of available data sets and the capacity to link these different data sets, have changed the landscape. The two factors have combined to make it easier to infer and relate seemingly non-personal or anonymised data to an identified or identifiable entity.

As digital technologies continue to evolve, so too will the policy concerns of data governance policies. However, technological development is inherently unpredictable. Digital technologies are deployed in an ever-increasing array of applications, which may raise new or different policy questions and considerations. Policies governing data, in turn, should seek to remain flexible, technologically neutral and sufficiently broad to remain resilient to technological change (OECD, 2022[14]).

4.4. Unlocking the potential of data for digital transformation requires encouraging firms of all sizes to use data

Firms increasingly use data as part of their business, for innovation and productivity, but this uptake is uneven. Due to the qualities of data as a highly scalable, non-rival input, larger firms may be in a better position to make the investments needed to generate, collect and use data and reap productivity benefits. Recent OECD work highlights how such dynamics are affecting productivity dispersion and industry concentration, with implications for the inclusiveness and dynamism of the broader economy (OECD, 2022[15]).

It is therefore essential to encourage the use of data in a wider share of the firm population. Data portability mechanisms, for example, can enable firms to access data, while strengthening consumer control. Other mechanisms include policies to encourage firms to make complementary investments necessary to use data productively, as well as to address market failures in financing systems that may discourage investments in data and other intangible assets. Due to their specific characteristics and the widely acknowledged importance of building their resilience (OECD, 2022[16]), small and medium-sized enterprises (SMEs) may merit more targeted assistance from policy makers. However, the extent of policy efforts in this area vary greatly across countries. Based on a cross-country analysis of 485 policies and 209 institutions, recent OECD work finds a variety of different institutions are in charge of digital and data policy design and implementation. Most do not have SME and entrepreneurship policy as a core mandate (OECD, 2022[17]).

At the same time, amid concerns that data are largely collected and used by a small number of firms, policy interest on the role of data in affecting competitive dynamics has also grown. To assist policy makers, the OECD has identified considerations for assessing whether data contributes to market power (OECD, 2022[18]). It has also outlined how competition authorities may need to adapt analytical approaches to incorporate challenges related to data (OECD, 2022[19]; 2022[15]). In addition, data portability and interoperability mechanisms hold promise for realising pro-competitive data sharing. In recent years, they have attracted interest from competition policy makers. However, co-operation between competition and privacy enforcement authorities may be key to the success of potentially overlapping regulations with different objectives, requirements and scope.

4.5. Efforts to measure the value of data, including in the System of National Accounts, should be encouraged

Data clearly have value to firms, individuals and societies at large, and efforts to better measure their value should be encouraged. However, as outlined in Chapter 3, the characteristics of data often preclude the emergence of a market price. This can therefore stymy market- and income-based approaches (OECD, 2022[20]), requiring development of new or more inventive models.

With other international organisations and national statistical institutes, the OECD is developing guidance on implementation of the sum-of-cost approach. With such an approach, the value of data is derived on the basis of the costs incurred to produce them. As with other own-account intellectual property products, like software and research and development, sum-of-cost appears to be the most promising approach to estimate the value of data produced by a firm for its own use rather than for sale (ISWGNA, 2022[21]; OECD, 2022[20]). Developing international statistical guidelines for this approach in the measurement of data investments and assets is a key focus of the work of the OECD and the international statistical community in the coming years.

Similarly, adapting statistical frameworks and nomenclatures to better identify data products, and the firms that produce them, could help shed light on the value of data. In addition, advancing efforts to measure cross-border data flows is crucial in view of increasing policy efforts in this domain. Measurement agendas should also seek to better understand and quantify the value of data and their use not captured in economic statistics, including for society at large (OECD, 2022[20]).

4.6. The OECD is well placed to support countries in their data governance policy making and can facilitate international co-ordination

The OECD provides multidisciplinary, evidence-based advice across policy domains transformed by data and on the role of data and data governance in economies and societies. The OECD has been at the forefront of developing policy guidance for technology and data governance, including *Recommendations of the Council on Access to Research Data From Public Funding* (OECD, 2021[22]), *Enhancing Access to and Sharing of Data* (OECD, 2021[13]), *Health Data Governance* (OECD, 2019[23]) and *Artificial Intelligence* (OECD, 2019[24]). The OECD is also developing evidence-based policy analysis in these areas, including furthering fundamental improvements in measurement frameworks, as well as promoting international regulatory co-operation (OECD, 2021[25]).

These efforts are essential to shaping global governance frameworks, particularly as developments in OECD countries proliferate easily in an interconnected world. The OECD provides a platform for shared learning among countries to help develop policies to realise the potential of data for social and economic prosperity around the world. As a trusted forum for evidence-based, multi-disciplinary and multi-stakeholder policy analysis and international dialogue, the OECD can help governments better govern data in the digital age.

References

Casalini, F., J. López González and T. Nemoto (2021), "Mapping commonalities in regulatory approaches to cross-border data transfers", *OECD Trade Policy Papers*, No. 248, OECD Publishing, Paris, https://doi.org/10.1787/ca9f974e-en. [9]

G20 (2019), *G20 Osaka Leaders' Declaration*, G20 2019 Japan, https://www.mofa.go.jp/policy/economy/g20_summit/osaka19/en/documents/final_g20_osaka_leaders_declaration.html. [7]

G7 (2022), *G7 Action Plan for Promoting Data Free Flow With Trust*, G7 Germany 2022, https://bmdv.bund.de/SharedDocs/DE/Anlage/K/g7-praesidentschaft-final-declaration-annex-1.pdf?__blob=publicationFile. [6]

Haksar, V. et al. (2021), "Toward a global approach to data in the digital age", *IMF Staff Discussion Note*, International Monetary Fund, Washington, DC, https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2021/10/06/Towards-a-Global-Approach-to-Data-in-the-Digital-Age-466264. [4]

ISWGNA (2022), *Recording of data in the National Accounts*, Intersecretariat Working Group on National Accounts, https://unstats.un.org/unsd/nationalaccount/RAdocs/DZ6_GN_Recording_of_Data_in_NA.pdf. [21]

OECD (2022), "Data in an evolving technological landscape: The case of connected and automated vehicles", *OECD Digital Economy Papers*, No. 346, OECD Publishing, Paris, https://doi.org/10.1787/ec7d2f6b-en. [14]

OECD (2022), "Data shaping firms and markets", *OECD Digital Economy Papers*, No. 344, OECD Publishing, Paris, https://doi.org/10.1787/7b1a2d70-en. [15]

OECD (2022), *Financing Growth and Turning Data into Business: Helping SMEs Scale Up*, OECD Studies on SMEs and Entrepreneurship, OECD Publishing, Paris, https://doi.org/10.1787/81c738f0-en. [17]

OECD (2022), "Fostering cross-border data flows with trust", *OECD Digital Economy Policy Papers*, No. 343, OECD Publishing, Paris, https://doi.org/10.1787/139b32ad-en. [8]

OECD (2022), *Going Digital Guide to Data Governance Policy Making*, OECD Publishing, Paris, https://doi.org/10.1787/40d53904-en. [1]

OECD (2022), "Government access to personal data held by the private sector", *Statement by the OECD Committee on Digital Economy Policy*, https://www.oecd.org/digital/trusted-government-access-personal-data-private-sector.htm. [11]

OECD (2022), "Measuring the value of data and data flows", *OECD Digital Economy Papers*, No. 345, OECD Publishing, Paris, https://doi.org/10.1787/923230a6-en. [20]

OECD (2022), *OECD Handbook on Competition Policy in the Digital Age*, OECD, Paris, https://www.oecd.org/daf/competition/oecd-handbook-on-competition-policy-in-the-digital-age.pdf. [19]

OECD (2022), *Recommendation of the Council on SME and Entrepreneurship Policy*, OECD/LEGAL/0473, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0473. [16]

OECD (2022), "Responding to societal challenges with data: Access, sharing, stewardship and control", *OECD Digital Economy Papers*, No. 342, OECD Publishing, Paris, https://doi.org/10.1787/2182ce9f-en. [2]

OECD (2022), "The evolving concept of market power in the digital economy", *OECD Competition Background Note*, OECD, Paris, https://www.oecd.org/daf/competition/the-evolving-concept-of-market-power-in-the-digital-economy-2022.pdf. [18]

OECD (2021), *International Regulatory Co-operation*, OECD Best Practice Principles for Regulatory Policy, OECD Publishing, Paris, https://doi.org/10.1787/5b28b589-en. [25]

OECD (2021), *Recommendation of the Council concerning Research Data from Public Funding*, OECD/LEGAL/0347, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347. [22]

OECD (2021), *Recommendation of the Council on Enhancing Access to and Sharing of Data*, OECD/LEGAL/0463, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463. [13]

OECD (2019), *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449. [24]

OECD (2019), *Recommendation of the Council on Health Data Governance*, OECD/LEGAL/0433, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0433. [23]

OECD (2013), *OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*, OECD/LEGAL/0188, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0188. [10]

Robinson, L., K. Kizawa and E. Ronchi (2021), "Interoperabiltiy of privacy and data protection frameworks", *OECD Going Digital Toolkit Notes*, No. 21, OECD Publishing, Paris, https://doi.org/10.1787/64923d53-en. [12]

UNCTAD (2021), *Digital Economy Report 2021*, United Nations Conference on Trade and Development, Geneva, https://unctad.org/page/digital-economy-report-2021. [3]

World Bank (2021), *World Development Report*, World Bank, Washington, DC, https://www.worldbank.org/en/publication/wdr2021. [5]

Going Digital to Advance Data Governance for Growth and Well-being

Data are generated wherever digital technologies are deployed namely, in almost every part of modern life. Using these data can empower individuals, drive innovation, enable new digital products and improve policy making and public service delivery. But as data become more widely used across sectors and applications, the potential for misuse and harm also grows. To advance data governance for growth and well-being, this report advocates a holistic and coherent approach to data governance, domestically and across borders. It examines how data have emerged as a strategic asset, with the ability to transform lives and confer economic advantage. It explains how the unique characteristics of data can pose complex trade-offs and challenge policies that pre-date the data-driven era. This report provides new insights, evidence and analysis and outlines considerations for better data governance policies in the digital age.

This publication is a contribution to Phase III of the OECD Going Digital Project, which aims to provide policymakers with the tools they need to design and implement better data policies to promote growth and well-being.

For more information, visit ***www.oecd.org/going-digital/project.***

#GoingDigital

PRINT ISBN 978-92-64-85884-8
PDF ISBN 978-92-64-64982-8